ANNALES

UNIVERSITATIS SCIENTIARUM BUDAPESTINENSIS DE ROLANDO EÖTVÖS NOMINATAE

SECTIO MATHEMATICA

TOMUS XLIV.

2001

ANNALES

UNIVERSITATIS SCIENTIARUM BUDAPESTINENSIS DE ROLANDO EÖTVÖS NOMINATAE

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

ANNALES UNIV. SCI. BUDAPEST., **44** (2001), 3–9

THE INFLUENCE OF S-QUASINORMALITY OF SOME SUBGROUPS OF PRIME POWER ORDER ON THE STRUCTURE OF FINITE GROUPS

By

M. RAMADAN

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

(Received September 23, 1999)

1. Introduction

A subgroup of a group G which permutes with every subgroup of G is called a quasinormal subgroup of G. We say, following KEGEL [7] that a subgroup of G is S-quasinormal in G if it permutes with every Sylow subgroup of G. Several authors have investigated the structure of a finite group when some subgroups of prime power order of the group are well-situated in the group. ITO [6] proved that a finite group G of odd order is nilpotent provided that all minimal subgroups of G lie in the center of G. BUCKLEY [3] proved that if all minimal subgroups of an odd order group are normal, then the group is supersolvable. SHAALAN [10] proved that if every subgroup of G of prime order or 4 is S-quasinormal in G, then G is supersolvable. Recently, the authors [2, 8] proved the following: Put $\pi(G) = \{p_1, p_2, \ldots, p_n\}$, where $p_1 > p_2 > \ldots > p_n$. Let P_i be a Sylow p_i-subgroup of G and let the exponent of $\Omega(P_i)$ be $p_i^{e_i}$, where $i = 1, 2, \ldots, n$. Suppose that all members of the family $\{H \mid H \leq \Omega(P_i), H' = 1, \operatorname{Exp} H = p_i^{e_i}, i = 1, 2, \ldots, n\}$ are normal (quasinormal) in G. Then G is supersolvable. The object of this paper is to get: Put $\pi(G) = \{p_1, p_2, \ldots, p_n\}$, where $p_1 > p_2 > \ldots > p_n$. Let P_i be a Sylow p_i-subgroup of G and let the exponent of $\Omega(P_i)$ be $p_i^{e_i}$, where $i = 1, 2, \ldots, n$. Suppose that all members of the family $\{H \mid H \leq \Omega(P_i), H' = 1, \operatorname{Exp} H = p_i^{e_i}, i = 1, 2, \ldots, n\}$ are S-quasinormal in G. Then G is supersolvable. Throughout this paper the term group always means a group of finite order. Our notation is standard and taken mainly from [4].

2. Main results

We prove the following result:

THEOREM 2.1. *Let p be the smallest prime dividing $|G|$ and let P be a Sylow p-subgroup of G of exponent p^e, where $e \geq 1$. Suppose that all members of the family $\{H \mid \leq P,\ H' = 1, \operatorname{Exp} H = p^e\}$ are S-quasinormal in G. Then G has a normal p-complement.*

PROOF. We prove the Theorem by induction on $|G|$. Let H be a cyclic subgroup of P of order p^e. Our hypothesis implies that H is S-quasinormal in G. So it follows that HQ is a subgroup of G, where Q is a Sylow q-subgroup of G and $q > p$. Then HQ has a normal p-complement by [5, Satz 2.8, p.420]. We have that $H \leq P_G = \bigcap_{x \in G} P^x$ is normal in G and so $H = P_G \cap HQ$ is normal in HQ. It follows that $HQ = H \times Q$. Thus $O^p(G) = \langle Q \mid Q$ is a Sylow q-subgroup of G, $q \neq p \rangle \leq C_G(H)$. If $C_G(H) = G$ for all cyclic subgroups of order p^e in P, then it is easy to see that $P \leq Z(G)$ and so G has a normal p-complement by [4, Theorem 4.3, p.252]. Let $C_G(H) < G$ for some cyclic subgroup H of order p^e. Then $C_G(H)$ has a normal p-complement by induction on $|G|$. Since $O^p(G) \leq C_H(G)$, we have that $O^p(G)$ has a normal p-complement and so also does G.

As an immediate consequence of Theorem 2.1, we have:

COROLLARY 2.2. *Put $\pi(G) = \{p_1, p_2, \ldots, p_n\}$, where $p_1 > p_2 > \ldots > p_n$. Let P_i be a Sylow p_i-subgroup of G of exponent $p_i^{e_i}$, where $i = 2, 3, \ldots, n$. Suppose that all members of the family $\{H \mid H \leq P_i,\ H' = 1, \operatorname{Exp} H = p_i^{e_i},$ $i = 2, 3, \ldots, n\}$ are S-quasinormal in G. Then G possesses an ordered Sylow tower.*

PROOF. By Theorem 2.1, G has a normal p_n-complement. Let P_n be a Sylow p_n-subgroup of G and K be a normal p_n-complement of P_n in G. By induction, K possesses an ordered Sylow tower. Therefore, G possesses an ordered Sylow tower too.

We need the following Lemma:

LEMMA 2.3. *Let P be a normal Sylow p-subgroup of G of exponent p^e, where $e \geq 1$ such that G/P is supersolvable. Suppose that all members of the family $\mathcal{H} = \{H \mid H \leq P,\ H' = 1, \operatorname{Exp} H = p^e\}$ are S-quasinormal in G. Then G is supersolvable.*

PROOF. We prove the lemma by induction on $|G|$. If every element in $\mathcal{H}$ is normal in G, then by [2, Theorem 4(b), p.253], $P \leq Q_\infty(G)$, where $Q_\infty(G)$ is the largest supersolvably embedded subgroup of G (see[12]), and hence $\langle x \mid x \in P, |x| \text{ is a prime or } 4\rangle \leq Q_\infty(G)$. Therefore, G is supersolvable by Yokoyama [11]. Thus we may assume that there exists an element H in $\mathcal{H}$ such that H is not normal in G. Our hypothesis implies that H is S-quasinormal in G. So it follows that HQ is a subgroup of G and $q \neq p$. Clearly, H is a subnormal Hall subgroup of HQ. Thus H is normal in HQ and so $Q \leq N_G(H)$. So $O^p(G) \leq N_G(H) < G$. Let $L = HO^p(G) \leq N_G(H)$. Then $G = PL$. Since $G/P \cong L/(L \cap P)$ is supersolvable, it follows that L is supersolvable by induction on $|G|$. Then $O^p(G)$ is a normal supersolvable subgroup of G. Since $O^p(G)$ is a normal supersolvable subgroup of G, it follows by [9, Exercise 7.2.23, p.159] that $PO^p(G) = G$ is supersolvable.

THEOREM 2.4. *Put $\pi(G) = \{p_1, p_2, \ldots, p_n\}$, where $p_1 > p_2 > \ldots > p_n$. Let P_i be a Sylow p_i-subgroup of G of exponent $p_i^{e_i}$, where $i = 1, 2, \ldots, n$. Suppose that all members of the family $\{H \mid H \leq P_i,\ H' = 1,\ \operatorname{Exp} H = p_i^{e_i},\ i = 1, 2, \ldots, n\}$ are S-quasinormal in G. Then G is supersolvable.*

PROOF. By corollary 2.2, G possesses an ordered Sylow tower. Then P_1 is normal in G. By Schur–Zassenhaus' theorem, G possesses a p_i'-Hall subgroup K which is a complement to P_1 in G. Hence K is supersolvable by induction on $|G|$. Now it follows from lemma 2.3 that G is supersolvable.

THEOREM 2.5. *Let P be a normal p-subgroup of G of exponent p^e, where $e \geq 1$ such that G/P is supersolvable. Suppose that all members of the family $\{H \mid H \leq P,\ H' = 1,\ \operatorname{Exp} H = p^e\}$ are S-quasinormal in G. Then G is supersolvable.*

PROOF. We prove the Theorem by induction on $|G|$. Let P_1 be a Sylow p-subgroup of G. We treat the following two cases:

CASE 1. $P = P_1$. Then by lemma 2.3, G is supersolvable.

CASE 2. $P < P_1$. Put $\pi(G) = \{p_1, p_2, \ldots, p_n\}$, where $p_1 > p_2 > \ldots > p_n$. Since G/P is supersolvable, it follows by [1, Theorem 5, p.5] that G/P possesses supersolvable subgroups H/P and K/P such that $|G/P : H/P| = p_1$ and $|G/P : K/P| = p_n$. Since H/P and K/P are supersolvable, it follows that H and K are supersolvable by induction on $|G|$. Since $|G : H| = |G/P : H/P| = p_1$ and $|G : K| = |G/P : K/P| = p_n$, it follows by [1, Theorem 5, p.5] that G is supersolvable.

COROLLARY 2.6. *Let K be a normal subgroup of G such that G/K is supersolvable. Put $\pi(K) = \{p_1, p_2, \ldots, p_s\}$, where $p_1 > p_2 > \ldots > p_s$ and let P_i be a Sylow p_i-subgroup of K of exponent $p_i^{e_i}$, where $i = 1, 2, \ldots, s$. Suppose that all members of the family $\{H \mid H \leq P_i,\ H' = 1, \operatorname{Exp} H = p_i^{e_i},\ i = 1, 2, \ldots, s\}$ are S-quasinormal in G. Then G is supersolvable.*

PROOF. We prove the corollary by induction on $|G|$. Theorem 2.4 implies that K is supersolvable and so P_1 is normal in K, where P_1 is a Sylow p_1-subgroup of K and p_1 is the largest prime dividing $|K|$. Clearly, P_1 is normal in G. Also, $(G/P_1)/(K/P_1) \cong G/K$ is supersolvable. Now we conclude that G/P_1 is supersolvable by induction on $|G|$. Now it follows from Theorem 2.5 that G is supersolvable. The corollary is proved.

For a finite group P, we write

$$\Omega(P) = \begin{cases} \Omega_1(P) & \text{if } p > 2 \\ \Omega_2(P) & \text{if } p = 2 \end{cases}$$

where, as usual,

$$\Omega_i(P) = \langle x \in P \mid |x| \mid p^i \rangle.$$

We are now in a position to prove the following results:

THEOREM 2.7. *Let p be the smallest prime dividing $|G|$, P be a Sylow p-subgroup of G and let the exponent of $\Omega(P)$ be p^e, where $e \geq 1$. Suppose that all members of the family $\{H \mid H \leq \Omega(P),\ H' = 1, \operatorname{Exp} H = p^e\}$ are S-quasinormal in G. Then G has a normal p-complement.*

PROOF. Let H be an abelian subgroup of $\Omega(P)$ of exponent p^e. Our hypothesis implies that H is S-quasinormal in G. So it follows that HQ is a subgroup of G, where Q is a Sylow q-subgroup of G and $q \neq p$. Clearly, H is normal in HQ and so $Q \leq N_G(H)$. Thus $O^p(G) \leq N_G(H) \leq G$. Clearly, $HO^p(G) \leq N_G(H) \leq G$. If $HO^p(G) \leq N_G(H) < G$, then $HO^p(G)$ has a normal p-complement, say, K by induction. Thus K is a normal p'-Hall subgroup of $O^p(G)$. Since K char $O^p(G)$ and $O^p(G)$ is normal in G, it follows that K is normal in G. Since $G/O^p(G)$ is a p-group, we have that K is a normal p'-Hall subgroup of G and so G has a normal p-complement. Thus we may assume that $N_G(H) = G$. In particular, H is normal in G. If G has no normal p-complement, then by Frobenius' theorem, there exists a nontrivial p-subgroup L of G such that $N_G(L)/C_G(L)$ is not a p-group. Clearly, we can assume that $L \leq P$. Let r be any prime dividing $|N_G(L)|$ with $r \neq p$ and let R be a Sylow r-subgroup of $N_G(L)$. Then R normalizes

L and so $\Omega(L)R$ is a subgroup of $N_G(L)$. Since H is normal in G, hence Theorem 2.1 implies that $(H\Omega(L))R$ has a normal p-complement and so also does $\Omega(L)R$. By [5, Satz 5.12, p.437], R centralizes L. Thus for each prime r dividing $|N_G(L)|$ with $r \neq p$, each Sylow r-subgroup R of $N_G(L)$ centralizes L and hence $N_G(L)/C_G(L)$ is a p-group; a contradiction. Therefore G has a normal p-complement.

As an immediate consequence of Theorem 2.7 we have:

COROLLARY 2.8. *Put* $\pi(G) = \{p_1, p_2, \ldots, p_n\}$, *where* $p_1 > p_2 > \ldots > p_n$. *Let* P_i *be a Sylow* p_i*-subgroup of* G *and let the exponent of* $\Omega(P_i)$ *be* $p_i^{e_i}$, *where* $i = 2, 3, \ldots, n$. *Suppose that all members of the family* $\{H \mid H \leq \Omega(P_i),\ H' = 1,\ \operatorname{Exp} H = p_i^{e_i},\ i = 2, 3, \ldots, n\}$ *are* S*-quasinormal in* G. *Then* G *possesses an ordered Sylow tower.*

We need the following lemmas:

LEMMA 2.9. [M. Ezzat, Finite groups in which some subgroups of prime power order are normal, M. SC. Thesis, Cairo University (1995)]. *Suppose that* P *is a normal Sylow* p*-subgroup of* G *and that* $\Omega(P)K$ *is supersolvable, where* K *is a* p'*-Hall subgroup of* G. *Then* G *is supersolvable.*

LEMMA 2.10. *Let* P *be a normal* p*-subgroup of* G *such that* G/P *is supersolvable and let the exponent of* $\Omega(P)$ *be* p^e, *where* $e \geq 1$. *Suppose that all members of the family* $\{H \mid H \leq \Omega(P),\ H' = 1,\ \operatorname{Exp} H = p^e\}$ *are* S*-quasinormal in* G. *Then* G *is supersolvable.*

PROOF. We prove the lemma by induction on $|G|$. Let P_1 be a Sylow p-subgroup of G. We treat the following two cases:

CASE 1. $P = P_1$. Then by Schur–Zassenhaus' theorem, G possesses a p'-Hall subgroup K which is a complement to P in G. Thus $G/P \cong K$ is supersolvable. Since $\Omega(P)$ char P and P is normal in G, it follows that $\Omega(P)$ is normal in G. Then $\Omega(P)K$ is a subgroup of G. If $\Omega(P)K = G$, then $G/\Omega(P)$ is supersolvable. Therefore G is supersolvable by Theorem 2.5. Thus we may assume that $\Omega(P)K < G$. Since $\Omega(P)K/\Omega(P) \cong K$ is supersolvable, it follows by Theorem 2.5 that $\Omega(P)K$ is supersolvable. Applying lemma 2.9, we conclude the supersolvability of G.

CASE 2. $P < P_1$ put $\pi(G) = \{p_1, p_2, \ldots, p_n\}$, where $p_1 > p_2 > \ldots > p_n$. Since G/P is supersolvable, it follows by [1, Theorem 5, p.5] that G/P possesses supersolvable subgroups H/P and K/P such that $|G/P : H/P| = p_1$ and $|G/P : K/P| = p_n$. Since H/P and K/P are supersolvable, it follows that H and K are supersolvable by induction on $|G|$. Since $|G : H| = |G/P : H/P| = p_1$ and $|G : K| = |G/P : K/P| = p_n$, it follows by [1, Theorem 5, p.5] that G is supersolvable.

As an immediate consequence of corollary 2.8 and lemma 2.10, we have:

THEOREM 2.11. *Put $\pi(G) = \{p_1, p_2, \ldots, p_n\}$, where $p_1 > p_2 > \ldots > p_n$. Let P_i be a Sylow p_i-subgroup of G and let the exponent of $\Omega(P_i)$ be $p_i^{e_i}$, where $i = 1, 2, \ldots, n$. Suppose that all members of the family $\{H \mid H \leq \Omega(P_i),\ H' = 1,\ \mathrm{Exp}\, H = p_i^{e_i},\ i = 1, 2, \ldots, n\}$ are S-quasinormal in G. Then G is supersolvable.*

PROOF. We prove the Theorem by induction on $|G|$. By corollary 2.8, we have that G possesses an ordered Sylow tower. Then P_1 is normal in G. By Schur–Zassenhaus' theorem, G possesses a p_1'-Hall subgroup K which is a complement to P_1 in G. Hence K is supersolvable by induction. Now it follows from lemma 2.10 that G is supersolvable.

We now obtain at once:

COROLLARY 2.12. *Put $\pi(G) = \{p_1, p_2, \ldots, p_n\}$, where $p_1 > p_2 > \ldots > p_n$. Let P_i be a Sylow p_i-subgroup of G and let the exponent of $\Omega(P_i)$ be $p_i^{e_i}$, where $i = 1, 2, \ldots, n$. Suppose that all members of the family $\{H \mid H \leq \Omega(P),\ H' = 1,\ \mathrm{Exp}\, H = p_i^{e_i},\ i = 2, 3, \ldots, n\}$ are S-quasinormal in G. Then*

(i) *G possesses an ordered Sylow tower.*

(ii) *G/P_1 is supersolvable.*

We can now prove:

COROLLARY 2.13. *Let K be a normal subgroup of G such that G/K is supersolvable. Put $\pi(K) = \{p_1, p_2, \ldots, p_s\}$, where $p_1 > p_2 > \ldots > p_s$ and let P_i be a Sylow p_i-subgroup of K and let the exponent of $\Omega(P_i)$ be $p_i^{e_i}$, where $i = 1, 2, \ldots, s$. Suppose that all members of the family $\{H \mid H \leq \Omega(P_i),\ H' = 1,\ \mathrm{Exp}\, H = p_i^{e_i}\ i = 1, 2, \ldots, s\}$ are S-quasinormal in G. Then G is supersolvable.*

PROOF. We prove the corollary by induction on $|G|$. Theorem 2.11 implies that K is supersolvable and so P_1 is normal in K, where P_1 is a Sylow p_1-subgroup of K and p_1 is the largest prime dividing the order of K. Clearly, P_1 char K and since K is normal in G, it follows that P_1 is normal in G. Since $(G/P_1)/(K/P_1) \cong G/K$ is supersolvable, it follows that G/P_1 is supersolvable by induction on $|G|$. Therefore G is supersolvable by lemma 2.10. The corollary is proved.

References

[1] M. ASAAD, On the supersolvability of finite groups I, *Annales Univ. Sci. Budapest.,* **18** (1975), 3–7.

[2] M. ASAAD, M. RAMADAN and M. EZZAT, Finite groups with normal subgroups of prime power orders, *Pure Math. & Appl.,* **5** (1994), 251–256.

[3] J. BUCKLEY, Finite groups whose minimal subgroups are normal, *Math. Z.,* **116** (1970), 15–17.

[4] D. GORENSTEIN, *Finite groups,* Harper and Row, New York, 1968.

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

[6] N. ITO, Über eine zur Frattini-gruppe duale Bildung, *Nagoya Math. J.,* **9** (1955), 123–127.

[7] O. H. KEGEL, Sylow-gruppen and Subnormalteiler endlicher gruppen, *Math. Z.,* **78** (1962), 205–221.

[8] M. RAMADAN, Finite groups in which some subgroups of prime power order are quasinormal, *J. Egypt Math. Soc.,* **1** (1997), 1–7.

[9] W. R. SCOTT, *Group theory,* Prentice-Hall, Englewood Cliffs, New Jersey, USA, 1964.

[10] E. SHAALAN, The influence of π-quasinormality of some subgroups on the structure of a finite group, *Acta Math. Hung.,* **56** (1990), 287–293.

[11] A. YOKOYAMA, Finite solvable groups whose J-hypercenter contains all minimal subgroups, *Arch. Math.,* **26** (1975), 123–130.

[12] M. WEINSTEIN (editor), *Between nilpotent and solvable,* Polygonal Publishing House, Passaic, 1982.

ANNALES UNIV. SCI. BUDAPEST., **44** (2001), 11–26

TWO-DIMENSIONAL LANDSBERG MANIFOLDS WITH VANISHING DOUGLAS TENSOR

By

SZ. VATTAMÁNY and CS. VINCZE

Institute of Mathematics and Informatics, University of Debrecen, Debrecen

(Received March 23, 2002)

Introduction

In his outstanding work [2], LUDWIG BERWALD showed that a two-dimensional Landsberg manifold reduces to a Berwald manifold if its Douglas tensor vanishes. In his own words: *"The Landsberg spaces, the extremals of which form a quasigeodesic system of curves, are identical with the two-dimensional affinely connected Finsler spaces."* ([2], p. 110.) The notion of a "quasigeodesic system of curves" was introduced by W. BLASCHKE and G. BOL in their book [3]. In modern language, the condition on the extremals states that the geodesics of a Landsberg manifold coincide with those of a linear connection on the underlying manifold. Since by the Douglas–Shen theorem ([7], 6.6) this property is equivalent to the vanishing of the Douglas tensor, our formulation is indeed a translation of Berwald's theorem.

The analogous result in the higher dimensional case was proved in [9] with the help of the projection of the Douglas tensor onto the indicatrix bundle. Unfortunately, this elegant method does not work in two dimensions, since the Douglas tensor then has components only along the Liouville vector field by formula (26a) of Remark 2.4 below. Since the Liouville vector field is orthogonal to the unit sphere bundle, we infer immediately that *the projected Douglas tensor of a two-dimensional Finsler manifold vanishes identically.* So we have to search for a completely different plan of attack in the two-dimensional case. Our approach is based on Berwald's original ideas, and it may be considered as an intrinsic version of them.

We shall adopt throughout the notations, terminology and conventions of [9] with one restriction. In this paper (M, E) will denote a *positive definite two-dimensional Finsler manifold.* (There is a generalization of the theorem for two-dimensional Finsler manifolds with nondegenerate Riemann–Finsler metric. It also needs a little modification of Berwald's method as we can see in [1] using the machinery of classical tensor calculus.)

1. Two-dimensional Finsler manifolds

1.1. THE BERWALD FRAME. Starting from the Liouville vector field C and the canonical spray S of (M, E), let us first consider the unit vector fields

$$C_0 := \frac{1}{\sqrt{2E}} C \qquad \text{and} \qquad S_0 := \frac{1}{\sqrt{2E}} S.$$

Next, using the Gram–Schmidt process we construct, at least locally, a g-orthonormal basis (C_0, X_0) of $\mathfrak{X}^v(\mathcal{T}M)$, where $g \in \mathcal{T}^0_2(\mathcal{T}M)$ is the metric tensor given by (20) in [9]. Applying the almost complex structure F associated with the Barthel endomorphism of (M, E), we obtain a local g-orthonormal basis (FX_0, S_0) of $\mathfrak{X}^h(\mathcal{T}M)$. The quadruple

$$(C_0, X_0, FX_0, S_0)$$

constructed in this way is a (local) orthonormal basis of $\mathfrak{X}(\mathcal{T}M)$; it is called the *Berwald frame* of the Finsler manifold (M, E).

Note that

$$X_0 E = 0. \tag{1}$$

This can be seen by a straightforward calculation:

$$0 = g(C, X_0) \overset{(20)/[9]}{=} \omega(C, FX_0) \overset{(15)/[9]}{=} = d_J E(FX_0) =$$
$$= dE[JF(X_0)] \overset{(13)/[6]}{=} dE[vX_0)] = dE(X_0) = X_0 E.$$

1.2. PROPOSITION. *The members of the Berwald frame have the following homogeneity properties:*

$$[C, C_0] = -C_0, \tag{2}$$
$$[C, S_0] = 0, \tag{3}$$
$$[C, X_0] = -X_0, \tag{4}$$
$$[C, FX_0] = 0. \tag{5}$$

PROOF. Taking into account the fact that the function $\frac{1}{\sqrt{2E}}$ is homogeneous of degree -1, i.e., that

$$C\left(\frac{1}{\sqrt{2E}}\right) = -\frac{1}{\sqrt{2E}}. \tag{6}$$

we obtain

$$[C, C_0] = \left[C, \frac{1}{\sqrt{2E}}C\right] = C\left(\frac{1}{\sqrt{2E}}\right)C = -\frac{1}{\sqrt{2E}}C = -C_0,$$

which shows (2). Similarly, since S is a spray and hence $[C, S] = S$,

$$[C, S_0] = \frac{1}{\sqrt{2E}}[C, S] + C\left(\frac{1}{\sqrt{2E}}\right)S = \frac{1}{\sqrt{2E}}S + C\left(\frac{1}{\sqrt{2E}}\right)S \overset{(6)}{=} 0,$$

so (3) is also true.

In view of the definition (21) and the property (27) in [9] of the first Cartan tensor $\mathscr{C}$,

$$0 = 2g(\mathscr{C}(S, FX_0), X_0) = \mathscr{L}_C(J^* g)(FX_0, FX_0) =$$
$$= Cg(X_0, X_0) - 2g(J[C, FX_0], X_0] \overset{1.1}{=} -2g(J[C, FX_0], X_0).$$

In the same way we find that

$$0 = 2g(\mathscr{C}(S, FX_0), C) = Cg(X_0, C) - g(J[C, FX_0], C) - g(X_0, J[C, S]) =$$
$$\overset{(1.1)}{=} -g(J[C, FX_0], C) - g(X_0, C) \overset{(1.1)}{=} -g(J[C, FX_0], C).$$

It follows from the last two relations that $J[C, FX_0] = 0$, and so $[C, FX_0]$ is vertical. Thus,

$$X_0 = vX_0 \overset{(13)/[6]}{=} JFX_0 \overset{(9)/[6]}{=} [J, C](FX_0) = [X_0, C] - J[FX_0, C] = -[C, X_0]$$

which proves the relation (4).

Taking into account again the fact that the vector field $[FX_0, C]$ is vertical, from the homogeneity of the Barthel endomorphism (see (18) in [9]) we obtain

$$0 = [h, C](FX_0) = [h(FX_0), C] - h[FX_0, C] =$$
$$= [h(FX_0), C] \overset{(13)/[6]}{=} [F(vX_0), C] = [FX_0, C]$$

whence (5). ∎

1.3 PROPOSITION. *With the hypothesis above, the following relations hold:*

(7) $\mathscr{C}'(FX_0, FX_0) = v[X_0, FX_0] \qquad (v := 1 - h),$

(8) $v[X_0, S] = 0,$

(9) $\forall Y \in \mathfrak{X}(\mathcal{T}M) : D_{hY} X_0 = 0,$ *where D is the Cartan connection.*

PROOF. According to the definition of the second Cartan tensor $\mathscr{C}'$ (see e.g. [9]. formula (23)),

$$\begin{aligned}
2g(\mathscr{C}'(FX_0, FX_0), X_0) &= (\mathcal{L}_{h(FX_0)} g)(J(FX_0), J(FX_0)) = \\
&\overset{(13)/[6]}{=} (\mathcal{L}_{FX_0} g)(X_0, X_0) = FX_0 g(X_0, X_0) - 2g([FX_0, X_0], X_0) = \\
&\overset{1.1}{=} 2g([X_0, FX_0], X_0) = \\
&= 2g(v[X_0, FX_0], X_0) + 2g(h[X_0, FX_0], X_0) = 2g(v[X_0, FX_0], X_0),
\end{aligned}$$

using the g-orthogonality of the vertical and horizontal subbundle in the last step. From this we obtain the relation

$$g(\mathscr{C}'(FX_0, FX_0) - v[X_0, FX_0], X_0) = 0. \tag{10}$$

Furthermore, by the properties (26), (27) in [9] of the second Cartan tensor, we can write

$$\begin{aligned}
0 &= 2g(\mathscr{C}'(FX_0, FX_0), C) = \\
&= FX_0 g(X_0, C) - g([FX_0, X_0], C) - g(X_0, [FX_0, C]) = \\
&\overset{1.1,\,(5)}{=} g([X_0, FX_0], C) = \\
&= g(v[X_0, FX_0], C) + g(h[X_0, FX_0], C) = g(v[X_0, FX_0], C).
\end{aligned}$$

This means that the equality

$$g(\mathscr{C}'(FX_0, FX_0) - v[X_0, FX_0], C) = 0 \tag{11}$$

is valid automatically. The relations (10) and (11) show that the vertical vector field $\mathscr{C}'(FX_0, FX_0) - v[X_0, FX_0]$ is orthogonal to two nowhere vanishing vertical vector fields. Hence it must be the zero vector field, which proves (7).

Similarly, on the one hand

$$\begin{aligned}
0 &= 2g(\mathscr{C}'(S, FX_0), X_0) = Sg(X_0, X_0) - 2g([S, X_0], X_0) = \\
&= 2g([X_0, S], X_0) = \\
&= 2g(v[X_0, S], X_0) + 2g(h[X_0, S], X_0) = 2g(v[X_0, S], X_0),
\end{aligned}$$

on the other hand

$$\begin{aligned}0 &= 2g(\mathscr{C}'(S, FX_0), C) = Sg(X_0, C) - g([S, X_0], C) - g(X_0, [S, C]) = \\ &= g([X_0, S], C) + g(X_0, S) = g([X_0, S], C) = \\ &= g(v[X_0, S], C) + g(h[X_0, S], C) = g(v[X_0, S], C),\end{aligned}$$

which imply the relation (8).

To prove the formula (9) it is sufficient to check that $D_{FX_0}X_0$ and $D_S X_0$ vanish. But this is immediate:

$$D_{FX_0}X_0 \overset{1.6/[9]}{=} v[FX_0, X_0] + \mathscr{C}'(FX_0, FX_0) \overset{(7)}{=} 0,$$

$$D_S X_0 \overset{1.6/[9]}{=} v[S, X_0] \overset{(8)}{=} 0. \qquad \blacksquare$$

1.4. Definition and Remark. The function

$$\lambda := g(\mathscr{C}(FX_0, FX_0), X_0) \tag{12}$$

is said to be the *main scalar* of (M, E) with respect to the Berwald frame (C_0, X_0, FX_0, S_0). – Actually, λ depends only on the choice of X_0 and it is uniquely determined up to sign.

1.5. Lemma. *With the help of the main scalar and the vector field X_0, the first Cartan tensors can be represented in the form*

$$\mathscr{C} = \lambda i_{X_0}\omega \otimes i_{X_0}\omega \otimes X_0 \qquad \textit{and} \qquad \mathscr{C}_b = \lambda i_{X_0}\omega \otimes i_{X_0}\omega \otimes i_{X_0}\omega, \tag{13}$$

where ω is the fundamental two-form.

Proof. The vertical vector field $\mathscr{C}(FX_0, FX_0)$ can be uniquely represented as a linear combination

$$\mathscr{C}(FX_0, FX_0) = \lambda_1 X_0 + \lambda_2 C_0, \qquad \lambda_1,\ \lambda_2 \in C^\infty(\mathcal{T}M).$$

Since on the one hand

$$g(\mathscr{C}(FX_0, FX_0), C_0) = \frac{1}{\sqrt{2E}}\mathscr{C}_b(FX_0, FX_0, S) \overset{(27)/[8]}{=} 0,$$

on the other hand

$$g(\mathscr{C}(FX_0, FX_0), C_0) = g(\lambda_1 X_0 + \lambda_2 C_0, C_0) = \lambda_2,$$

it follows that $\lambda_2 = 0$. So

$$\lambda := g(\mathscr{C}(FX_0, FX_0), X_0) = g(\lambda_1 X_0, X_0) = \lambda_1,$$

hence $\mathscr{C}(FX_0, FX_0) = \lambda X_0$, where λ is the main scalar. From this observation we infer immediately (13). $\blacksquare$

1.6. COROLLARY. *A positive definite two-dimensional Finsler manifold is a Riemannian manifold if and only if its main scalar vanishes.*

PROOF. This is an immediate consequence of (13), because (M, E) is a Riemannian manifold if and only if $\mathscr{C} = 0$. ∎

1.7. PROPOSITION. *A positive definite two-dimensional Finsler manifold is a Berwald manifold if and only if the horizontal vector fields annihilate the main scalar, i.e., $d_h\lambda = 0$.*

PROOF. We first recall that a Finsler manifold is a Berwald manifold if and only if the h-covariant derivative, with respect to the Cartan connection, of the first Cartan tensor vanishes. (An intrinsic proof of this well-known fact is available in [8].) Thus we can argue as follows:

$$\begin{gathered}(M,E) \textit{ is a Berwald manifold} \Leftrightarrow \forall Y \in \mathfrak{X}(\mathcal{T}M) : D_{hY}\mathscr{C} = 0 \Leftrightarrow \\ \Leftrightarrow \ \forall Y \in \mathfrak{X}(\mathcal{T}M) : 0 = (D_{hY}\mathscr{C})(FX_0, FX_0) = \\ = D_{hY}[\mathscr{C})(FX_0, FX_0)] - 2\mathscr{C}(D_{hY}FX_0, FX_0) \overset{(\mathrm{F{\scriptsize INS}2})/[6]}{=} \\ = D_{hY}[\mathscr{C}(FX_0, FX_0)] - 2\mathscr{C}(FD_{hY}X_0, FX_0) \overset{(9)}{=} D_{hY}[\mathscr{C}(FX_0, FX_0)] \overset{(13)}{=} \\ = D_{hY}(\lambda X_0) = [(hY)\lambda]X_0 + \lambda D_{hY}X_0 \overset{(9)}{=} [(hY)\lambda]X_0 = [d_h\lambda(Y)]X_0 \ \Leftrightarrow \\ \Leftrightarrow \ d_h\lambda = 0.\end{gathered}$$

∎

1.8. COROLLARY. *The second Cartan tensor of (M, E) is completely determined by the formula*

$$\mathscr{C}'(FX_0, FX_0) = -S(\lambda)X_0. \tag{14}$$

PROOF. Since $D_S\mathscr{C} = -\mathscr{C}'$ (see [5], Prop. (A.12)), taking into account the previous proof we obtain

$$\begin{aligned}\mathscr{C}'(FX_0, FX_0) &= -(D_S\mathscr{C})(FX_0, FX_0) = -[d_h\lambda(S)]X_0 = \\ &= -[d\lambda(hS)]X_0 = [-(d\lambda)S]X_0 = -S(\lambda)X_0.\end{aligned}$$

∎

1.9. PROPOSITION AND DEFINITION. *Let us consider the curvature tensor $R = -\frac{1}{2}[h, h]$ of (M, E). Then*

$$R(FX_0, S_0) = g(R(FX_0, S_0), X_0)X_0, \tag{15}$$

and R is uniquely determined by this formula on the domain of the Berwald frame constructed in **1.1**. *The function*

$$\kappa := g(R(FX_0, S_0), X_0) \tag{16}$$

is said to be the Gauss curvature of (M, E).

PROOF. Since R is a semibasic tensor of type $(1,2)$, it is uniquely determined by its value on the local basis (FX_0, S_0) of $\mathfrak{X}^h(\mathcal{T}M)$. So our only task is to verify the equality (15). Starting with the definition of the Nijenhuis torsion, we obtain

$$R(FX_0, S_0) = -([hFX_0, hS_0] + h^2[FX_0, S_0] - h[FX_0, hS_0) - h[hFX_0, S_0]) = \\ = -[FX_0, S_0] - h[FX_0, S_0] + 2h[FX_0, S_0] = -v[FX_0, S_0].$$

Now, if

$$R(FX_0, S_0) = f_1 X_0 + f_2 C_0 \qquad (f_1, f_2 \in C^\infty(\mathcal{T}M)),$$

then on the one hand

$$g(R(FX_0, S_0), C) = g(f_1 X_0 + f_2 C_0, C) = f_2 g(C_0, C) = \\ = \frac{1}{\sqrt{2E}} f_2 g(C, C) = \sqrt{2E} f_2$$

on the other hand, using repeatedly the fact that $d_h E = 0$,

$$\begin{aligned} g(R(FX_0, S_0), C) &= -g(v[FX_0, S_0], C) = -\omega(v[FX_0, S_0], FC) = \\ &= -\omega(v[FX_0, S_0], S) = \\ &= i_S\omega(v[FX_0, S_0]) \overset{(16)/[9]}{=} -v[FX_0, S_0](E) = \\ &= h[FX_0, S_0](E) - [FX_0, S_0](E) = [S_0, FX_0](E) = \\ &= S_0[FX_0(E)] - FX_0[S_0(E)] = 0. \end{aligned}$$

These imply that $f_2 = 0$, $R(FX_0, S_0) = f_1 X_0$, and

$$f_1 = g(f_1 X_0, X_0) = g(R(FX_0, S_0), X_0)$$

whence (15). ∎

1.10. THEOREM (E. CARTAN's "permutation formulas"). *For the Lie brackets of the members of the Berwald frame we have*

$$\text{(17a–c)} \qquad \boxed{\begin{aligned} [X_0, FX_0] &= -\frac{1}{\sqrt{2E}} S_0 - \lambda FX_0 - S(\lambda) X_0 \\ [S_0, X_0] &= -\frac{1}{\sqrt{2E}} FX_0 \\ [FX_0, S_0] &= -\kappa X_0 \end{aligned}}$$

where λ *is the main scalar,* κ *is the Gauss curvature of* (M, E).

PROOF. Since

$$[X_0, FX_0] = v[X_0, FX_0] + h[X_0, FX_0] \overset{(7),(14)}{=} -S(\lambda)X_0 + h[X_0, FX_0],$$

to prove the relation (17a) it remains to be shown that

$$h[X_0, FX_0] = -\frac{1}{\sqrt{2E}}S_0 - \lambda FX_0. \tag{18}$$

First we observe that

$$2\lambda \overset{(12)}{=} 2g(\mathscr{C}(FX_0, FX_0), X_0) = X_0 g(X_0, X_0)$$
$$-2g(J[X_0, FX_0], X_0) = -2g(J[X_0, FX_0], X_0)$$

whence

$$g(J[X_0, FX_0], X_0) = -\lambda. \tag{19}$$

Similarly,

$$\begin{aligned} 0 &= 2g(\mathscr{C}(FX_0, FX_0), C) = \\ &= X_0 g(X_0, C) - g(J[X_0, FX_0], C) - g(X_0, J[X_0, S]) = \\ &= -g(J[X_0, FX_0], C) - g(X_0, J[X_0, S]) = \\ &\overset{\text{Prop. I.7/[4]}}{=} -g(J[X_0, FX_0], C) - g(X_0, X_0). \end{aligned}$$

Hence it follows that

$$g(J[X_0, FX_0], C_0) = -\frac{1}{\sqrt{2E}}. \tag{20}$$

Now we use orthonormal expansion to express $J[X_0, FX_0]$ in terms of the local basis (X_0, C_0) of $\mathfrak{X}^v(\mathcal{T}M)$:

$$\begin{aligned} J[X_0, FX_0] &= g(J[X_0, FX_0], X_0)X_0 + g(J[X_0, FX_0], C_0)C_0 = \\ &\overset{(19),(20)}{=} -\lambda X_0 - \frac{1}{\sqrt{2E}}C_0. \end{aligned}$$

In view of the identity $F \circ J = h$, from this we obtain (18) and hence (17a).

For the second formula (17b) we have

$$[S_0, X_0] = h[S_0, X_0] + v[S_0, X_0] \overset{(8)}{=} h[S_0, X_0] = -FJ\left[X_0, \frac{1}{\sqrt{2E}}S\right] =$$
$$\overset{(1)}{=} -\frac{1}{\sqrt{2E}}FJ[X_0, S] \overset{\text{Prop. I.7/[3]}}{=} -\frac{1}{\sqrt{2E}}FX_0.$$

To prove (17c), first it will be shown that the vector field $[FX_0, S_0]$ is vertical, i.e.,

$$h[FX_0, S_0] = 0. \tag{21}$$

The vanishing of the h-horizontal torsion of the Cartan connection (see [9], **1.7** or [6], (M3)) yields

$$h[FX_0, S_0] = D_{FX_0}S_0 - D_{S_0}FX_0 = D_{FX_0}S_0 - FD_{S_0}X_0 \overset{(9)}{=} D_{FX_0}S_0.$$

Using the fact that $d_h E = 0$ and the vanishing of the h-deflection of the Cartan connection ([6], (M4)) we obtain

$$D_{FX_0}S_0 = D_{FX_0}\left(\frac{1}{\sqrt{2E}}S\right) = FX_0\left(\frac{1}{\sqrt{2E}}\right)S + \frac{1}{\sqrt{2E}}D_{FX_0}S =$$
$$= \frac{1}{\sqrt{2E}}D_{FX_0}FC = \frac{1}{\sqrt{2E}}FD_{FX_0}C = 0,$$

thus (21) is true. By this observation and taking into account **1.9** it follows that

$$[FX_0, S_0] = v[FX_0, S_0] = -R(FX_0, S_0) = -\kappa X_0,$$

completing the proof. ∎

1.11. PROPOSITION („Bianchi identity”).

$$\boxed{\lambda\kappa + X_0(\kappa) + S_0(S\lambda) = 0.} \tag{22}$$

PROOF. Starting with the Jacobi identity, taking into account that both FX_0 and S_0 are horizontal vector fields and $d_h E = 0$, finally using (17a–c) we obtain

$$0 = [X_0, [S_0, FX_0]] + [FX_0, [X_0, S_0]] + [S_0, [FX_0, X_0]] =$$
$$\overset{(17a\text{–}c)}{=} [X_0, \kappa X_0] + \left[FX_0, \frac{1}{\sqrt{2E}}FX_0\right] + \left[S_0, \frac{1}{\sqrt{2E}}S_0 + \lambda FX_0 + S(\lambda)X_0\right] =$$
$$= (X_0\kappa)X_0 + \lambda[S_0, FX_0] + S_0(\lambda)FX_0 + S(\lambda)[S_0, X_0] + S_0(S\lambda)X_0 =$$
$$\overset{(17b\text{–}c)}{=} (X_0\kappa)X_0 + \lambda\kappa X_0 + S_0(\lambda)FX_0 - \frac{1}{\sqrt{2E}}S(\lambda)FX_0 + S_0(S\lambda)X_0 =$$
$$= [X_0(\kappa) + \lambda\kappa + S_0(S\lambda)]X_0$$

whence (22). ∎

2. Two-dimensional Landsberg manifolds with vanishing Douglas tensor

2.1. PROPOSITION. *A positive definite two-dimensional Finsler manifold is a Landsberg manifold if and only if the main scalar is a first integral of the canonical spray, i.e.,*

$$S(\lambda) = 0. \tag{23}$$

PROOF. *This is an immediate consequence of (14) and* **2.1** *in [9].* ∎

2.2. LEMMA. *Suppose that (M, E) is a (positive definite, two-dimensional) Landsberg manifold. Then the mixed curvature and the mixed Ricci tensor of the Berwald connection are completely determined by the formulas*

$$\overset{\circ}{\mathbb{P}}(FX_0, FX_0)FX_0 = -FX_0(\lambda)X_0, \tag{24}$$

$$\overset{\tilde{\circ}}{\mathbb{P}}(FX_0, FX_0) = -FX_0(\lambda), \tag{25}$$

where λ is the main scalar of (M, E).

PROOF. (25) is a trivial consequence of (24). To prove (24), we first recall that the relation

$$\overset{\circ}{\mathbb{P}}(X,\ Y)Z = -(D_{hX}\mathscr{C})(Y, Z) \quad (X,\ Y,\ Z \in \mathfrak{X}(\mathcal{T}M))$$

holds in any Landsberg manifold (see e.g. [9], **2.1**/(iv)). Thus, in our case

$$\begin{aligned}\overset{\circ}{\mathbb{P}}(FX_0, FX_0)FX_0 &= -(D_{FX_0}\mathscr{C})(FX_0, FX_0) = -D_{FX_0}\mathscr{C}(FX_0, FX_0) + \\ &\quad + 2\mathscr{C}(D_{FX_0}FX_0, FX_0) \overset{(13)}{=} -D_{FX_0}\lambda X_0 + \\ &\quad + 2\mathscr{C}(FD_{FX_0}X_0, FX_0) \overset{(9)}{=} -FX_0(\lambda)X_0\end{aligned}$$

whence (25). ∎

2.3. PROPOSITION. *The Douglas tensor of a positive definite two-dimensional Landsberg manifold is completely determined by*

$$\boxed{\mathbb{D}(FX_0, FX_0)FX_0 = \frac{1}{3}[X_0(FX_0(\lambda) + 2\lambda FX_0(\lambda)]C.} \tag{26}$$

PROOF. As we know from 6.2 and 6.3 in [7], $\mathbb{D}$ is semibasic, symmetric, and for any semispray S_0, $i_{S_0}\ \mathbb{D} = 0$. Thus, in two dimensions $\mathbb{D}$ is completely determined by its value on the triplet (FX_0, FX_0, FX_0). According to 6.2/(b) in [7],

$$\mathbb{D}(FX_0, FX_0)FX_0 = \overset{\circ}{\mathbb{P}}(FX_0, FX_0)FX_0 - \frac{1}{3}\left(\overset{\circ}{D}_J \overset{\tilde{\circ}}{\mathbb{P}}\right)(FX_0, FX_0, FX_0)C-$$

$$-\overset{\tilde{\circ}}{\mathbb{P}}(FX_0, FX_0)X_0 \overset{(24),(25)}{=} -\frac{1}{3}\left(\overset{\circ}{D}_{X_0} \overset{\tilde{\circ}}{\mathbb{P}}\right)(FX_0, FX_0)C =$$

$$= -\frac{1}{3}\left[X_0\left(\overset{\tilde{\circ}}{\mathbb{P}}(FX_0, FX_0)\right) - 2\overset{\tilde{\circ}}{\mathbb{P}}\left(\overset{\circ}{D}_{X_0} FX_0, FX_0\right)\right] C \overset{(25)}{=}$$

$$= \frac{1}{3}\left[X_0(FX_0(\lambda)) + 2\overset{\tilde{\circ}}{\mathbb{P}}\left(\overset{\circ}{D}_{X_0} FX_0, FX_0\right)\right] C.$$

There remains only to calculate the second member of the right hand side. Applying the rules of calculation of the Berwald connection ((27) and (BRW1–4) in [6]) we obtain

$$\overset{\circ}{D}_{X_0} FX_0 = F\,\overset{\circ}{D}_{JFX_0} JFX_0 = FJ[X_0, FX_0] = h[X_0, FX_0] \overset{(17a)}{=}$$

$$= -\frac{1}{\sqrt{2E}}S_0 - \lambda(h \circ F)X_0 - S(\lambda)hX_0 = -\frac{1}{\sqrt{2E}}S_0 - \lambda FX_0.$$

Hence

$$\overset{\circ}{\mathbb{P}}\left(\overset{\circ}{D}_{X_0} FX_0, FX_0\right) = -\frac{1}{\sqrt{2E}}\overset{\tilde{\circ}}{\mathbb{P}}(S_0, FX_0) - \lambda\overset{\tilde{\circ}}{\mathbb{P}}(FX_0, FX_0) =$$

$$\overset{(4.4a)/[7]}{=} -\lambda\overset{\tilde{\circ}}{\mathbb{P}}(FX_0, FX_0) \overset{(25)}{=} \lambda FX_0(\lambda).$$

and the result follows.

2.4. REMARK. Using the same technique, it may be proved that the effect of the tensors $\overset{\circ}{\mathbb{P}}$, $\overset{\tilde{\circ}}{\mathbb{P}}$ and $\mathbb{D}$ can be described analogously in *any* (positive definite, two-dimensional) Finsler manifold. More precisely, the following relations are fulfilled in the general case:

$$\text{(24a)} \qquad \overset{\circ}{\mathbb{P}}(FX_0, FX_0)FX_0 = -[FX_0(\lambda) + X_0(S\lambda)]X_0 + 2\frac{S(\lambda)}{\sqrt{2E}}C_0,$$

$$\text{(25a)} \qquad \overset{\tilde{\circ}}{\mathbb{P}}(FX_0, FX_0) = -FX_0(\lambda) - X_0(S\lambda),$$

$$\text{(26a)} \qquad \mathbb{D}(FX_0, FX_0)FX_0 = \frac{1}{3}\Big[X_0(FX_0(\lambda)) + X_0(X_0(S\lambda)) +$$

$$+ 2\lambda FX_0(\lambda) + 2X_0(S\lambda) + 3\frac{S(\lambda)}{E}\Big] C.$$

2.5. Lemma. *Suppose that (M, E) is a positive definite two-dimensional Landsberg manifold with vanishing Douglas tensor. The iterated Lie derivatives of the main scalar with respect to the vector fields X_0, FX_0, S_0 (up to fifth order) can be expressed as follows:*

$$\mathcal{L}_{FX_0}\mathcal{L}_{X_0}\lambda = -\lambda\mathcal{L}_{FX_0}\lambda, \tag{27}$$

$$\mathcal{L}_{S_0}\mathcal{L}_{X_0}\lambda = -\frac{1}{\sqrt{2E}}\mathcal{L}_{FX_0}\lambda, \tag{28}$$

$$\mathcal{L}_{FX_0}\mathcal{L}^2_{X_0}\lambda = \left(\lambda^2 - \mathcal{L}_{X_0}\lambda - \frac{1}{2E}\right)\mathcal{L}_{FX_0}\lambda, \tag{29}$$

$$\mathcal{L}_{S_0}\mathcal{L}^2_{X_0}\lambda = \frac{3}{\sqrt{2E}}\lambda\mathcal{L}_{FX_0}\lambda, \tag{30}$$

$$\mathcal{L}_{FX_0}\mathcal{L}^3_{X_0}\lambda = \left(-\lambda^3 + 3\lambda\mathcal{L}_{X_0}\lambda - \mathcal{L}^2_{X_0}\lambda + \frac{2}{E}\lambda\right)\mathcal{L}_{FX_0}\lambda, \tag{31}$$

$$\mathcal{L}_{S_0}\mathcal{L}^3_{X_0}\lambda = \frac{1}{\sqrt{2E}}\left(4\mathcal{L}_{X_0}\lambda - 7\lambda^2 + \frac{1}{2E}\right)\mathcal{L}_{FX_0}\lambda, \tag{32}$$

$$\begin{aligned}\mathcal{L}_{FX_0}\mathcal{L}^4_{X_0}\lambda = \Big(&-\mathcal{L}^3_{X_0}\lambda + 4\lambda\mathcal{L}^2_{X_0}\lambda + 3(\mathcal{L}_{X_0}\lambda)^2 - 6\lambda^2\mathcal{L}_{X_0}\lambda + \\ &+ \frac{4}{E}\mathcal{L}_{X_0}\lambda - \frac{11}{2E}\lambda^2 + \lambda^4 + \frac{1}{(2E)^2}\Big)\mathcal{L}_{FX_0}\lambda,\end{aligned} \tag{33}$$

$\mathcal{L}^n_{X_0} := \mathcal{L}_{X_0} \circ \ldots \circ \mathcal{L}_{X_0}$ *(n times).*

Proof. We shall verify only the first three formulas, the remaining ones can be handled in the same way. First we observe that the vanishing of the Douglas tensor implies by (26) the relation

$$\boxed{X_0(FX_0(\lambda)) = -2\lambda FX_0(\lambda).} \tag{34}$$

From now on we calculate.

(a) $\mathcal{L}_{FX_0}\mathcal{L}_{X_0}\lambda = [FX_0, X_0]\lambda + \mathcal{L}_{X_0}\mathcal{L}_{FX_0}\lambda \overset{(34)}{=} [FX_0, X_0]\lambda - 2\lambda FX_0(\lambda) \overset{(17a),(23)}{=}$
$= \frac{1}{\sqrt{2E}}(S_0\lambda) + \lambda(FX_0)\lambda - 2\lambda(FX_0)\lambda \overset{(23)}{=} -\lambda(FX_0)\lambda = -\lambda\mathcal{L}_{FX_0}\lambda,$
thus (27) is proved.

(b) $\mathcal{L}_{S_0}\mathcal{L}_{X_0}\lambda = [S_0, X_0]\lambda + X_0(S_0\lambda) \overset{(17b),\,(23)}{=} -\frac{1}{\sqrt{2E}}(FX_0)\lambda = -\frac{1}{\sqrt{2E}}\mathcal{L}_{FX_0}\lambda,$
so we have obtained (28).

c) $\mathcal{L}_{FX_0}\mathcal{L}^2_{X_0}\lambda = \mathcal{L}_{FX_0}\mathcal{L}_{X_0}(X_0\lambda) = [FX_0, X_0](X_0\lambda) + X_0[FX_0(X_0\lambda)] \overset{(17a),(23)}{=}$
$= \frac{1}{\sqrt{2E}} S_0(X_0\lambda) + \lambda(FX_0)(X_0\lambda) + X_0[FX_0(X_0\lambda)] \overset{(28),(27)}{=} -\frac{1}{2E}\mathcal{L}_{FX_0}\lambda -$
$-\lambda^2\mathcal{L}_{FX_0}\lambda + X_0[FX_0(X_0\lambda)] \overset{(27)}{=} \left(-\frac{1}{2E} - \lambda^2\right)\mathcal{L}_{FX_0}\lambda + X_0(-\lambda(FX_0)\lambda) =$
$= \left(-\frac{1}{2E} - \lambda^2\right)\mathcal{L}_{FX_0}\lambda - (X_0\lambda)\mathcal{L}_{FX_0}\lambda - \lambda\mathcal{L}_{X_0}\mathcal{L}_{FX_0}\lambda \overset{(34)}{=}$
$= \left(-\frac{1}{2E} - \lambda^2 - \mathcal{L}_{X_0}\lambda\right)\mathcal{L}_{FX_0}\lambda + 2\lambda^2\mathcal{L}_{FX_0}\lambda = \left(\lambda^2 - \mathcal{L}_{X_0}\lambda - \frac{1}{2E}\right)\mathcal{L}_{FX_0}\lambda,$
showing that (29) is also valid. ∎

2.6. THEOREM. *If a positive definite two-dimensional Landsberg manifold has a vanishing Douglas tensor, then it is a Berwald manifold.*

PROOF. **(A)** In the next, quite tedious calculations our aim is to show that

$$\mathcal{L}_{FX_0}\lambda = 0. \tag{35}$$

Then on the one hand

$$0 = (FX_0)\lambda = [(F \circ v)X_0]\lambda = [h(FX_0)]\lambda = (d_h\lambda)(FX_0)$$

on the other hand

$$(d_h\lambda)(S) = (d\lambda)(hS) = (d\lambda)S = S(\lambda) \overset{(23)}{=} 0,$$

so it follows that $d_h\lambda = 0$ and, in view of Proposition 1.7, (M, E) is a Berwald manifold.

Notice that our subsequent reasoning relies heavily on the fact that

$$\lambda\kappa = -X_0(\kappa). \tag{36}$$

This relation is an immediate consequence of the Bianchi identity (22) and the property (23).

(B) We start with the "permutation formula" (17c) and apply both its sides to the main scalar. Taking into account (23), we obtain

$$\mathcal{L}_{S_0}\mathcal{L}_{FX_0}\lambda = \kappa X_0(\lambda). \tag{37}$$

Now we evaluate the vector field $[S_0, X_0]$ on the function $FX_0(\lambda)$.

$[S_0, X_0](FX_0(\lambda)) = S_0[X_0(FX_0(\lambda))] - X_0[S_0(FX_0(\lambda))] \overset{(34),(37)}{=}$

$= -2S_0[\lambda(FX_0)(\lambda)] - X_0(\kappa X_0(\lambda)) \overset{(23)}{=}$

$= -2\lambda S_0[(FX_0)\lambda] - X_0(\kappa)X_0(\lambda) - \kappa X_0(X_0\lambda) \overset{(37),(36)}{=}$
$= -2\lambda\kappa X_0(\lambda) + \lambda\kappa X_0(\lambda) - \kappa X_0(X_0\lambda) = -\kappa(\lambda X_0(\lambda) + X_0(X_0\lambda)).$

Since, on the other side, $[S_0, X_0] \overset{(17b)}{=} -\frac{1}{\sqrt{2E}} FX_0$, it follows that

$$\frac{1}{\sqrt{2E}} \mathcal{L}^2_{FX_0} \lambda = \kappa(\lambda X_0(\lambda) + X_0(X_0\lambda)). \tag{38}$$

Applying the vector field X_0 to both sides of (38), owing to (1) we obtain

$$\frac{1}{\sqrt{2E}} \mathcal{L}_{X_0} \mathcal{L}^2_{FX_0} \lambda = X_0(\kappa)(\lambda X_0(\lambda) + X_0(X_0\lambda)) +$$

$$+\kappa \left[(\mathcal{L}_{X_0}\lambda)^2 + \lambda \mathcal{L}^2_{X_0}\lambda + \mathcal{L}^3_{X_0}\lambda \right] \overset{(36)}{=} \kappa(-\lambda^2 \mathcal{L}_{X_0}\lambda + (\mathcal{L}_{X_0}\lambda)^2 + \mathcal{L}^3_{X_0}\lambda),$$

i.e.

$$\frac{1}{\sqrt{2E}} \mathcal{L}_{X_0} \mathcal{L}^2_{FX_0} \lambda = \kappa(-\lambda^2 \mathcal{L}_{X_0}\lambda + (\mathcal{L}_{X_0}\lambda)^2 + \mathcal{L}^3_{X_0}\lambda). \tag{39}$$

The Lie derivatives of the two sides of (34) with respect to FX_0 yield

$$\frac{1}{\sqrt{2E}} \mathcal{L}_{FX_0} \mathcal{L}_{X_0} \mathcal{L}_{FX_0} \lambda = \frac{1}{\sqrt{2E}} [-2(\mathcal{L}_{FX_0}\lambda)^2 - 2\lambda \mathcal{L}^2_{FX_0}\lambda]. \tag{40}$$

Taking the difference of (39) and (40), and then substituting the term $\frac{1}{\sqrt{2E}} \mathcal{L}^2_{FX_0}\lambda$ from the right hand side of (38) we obtain

$$\frac{1}{\sqrt{2E}} [X_0, FX_0](FX_0(\lambda)) =$$

$$= \kappa(\lambda^2 \mathcal{L}_{X_0}\lambda + (\mathcal{L}_{X_0}\lambda)^2 + \mathcal{L}^3_{X_0}\lambda + 2\lambda \mathcal{L}^2_{X_0}\lambda) + \frac{2}{\sqrt{2E}} (\mathcal{L}_{FX_0}\lambda)^2.$$

The left hand side of this equality can also be written in the form

$$\frac{1}{\sqrt{2E}} [X_0, FX_0](FX_0(\lambda)) \overset{(17a),(23)}{=}$$

$$= -\frac{1}{2E} S_0[FX_0(\lambda)] - \frac{1}{\sqrt{2E}} \lambda(FX_0)(FX_0(\lambda)) =$$

$$\overset{(37),(38)}{=} -\kappa \left[\frac{1}{2E} X_0(\lambda) + \lambda^2 X_0(\lambda) + \lambda X_0(X_0\lambda) \right],$$

so it follows that

$$\frac{2}{\sqrt{2E}} (\mathcal{L}_{FX_0}\lambda)^2 + \kappa(\mathcal{L}^3_{X_0}\lambda + 3\lambda \mathcal{L}^2_{X_0}\lambda + (\mathcal{L}_{X_0}\lambda)^2 + 2\lambda^2 \mathcal{L}_{X_0}\lambda + \frac{1}{2E} \mathcal{L}_{X_0}\lambda) = 0. \tag{41}$$

Now we apply the vector field X_0 to (41). Taking into account that

$$X_0\left[\frac{2}{\sqrt{2E}}(\mathcal{L}_{FX_0}\lambda)^2\right] \overset{(1)}{=} \frac{2}{\sqrt{2E}}\mathcal{L}_{X_0}(\mathcal{L}_{FX_0}\lambda)^2 = \frac{4}{\sqrt{2E}}(\mathcal{L}_{FX_0}\lambda)\mathcal{L}_{X_0}\mathcal{L}_{FX_0}\lambda =$$

$$\overset{(34)}{=} -\frac{8\lambda}{\sqrt{2E}}(\mathcal{L}_{FX_0}\lambda)^2 =$$

$$\overset{(41)}{=} 4\cdot\kappa\lambda\left(\mathcal{L}^3_{X_0}\lambda + 3\lambda\mathcal{L}^2_{X_0}\lambda + (\mathcal{L}_{X_0}\lambda)^2 + 2\lambda^2\mathcal{L}_{X_0}\lambda + \frac{1}{2E}\mathcal{L}_{X_0}\lambda\right),$$

we obtain the relation

$$(X_0\kappa)\left(\mathcal{L}^3_{X_0}\lambda + 3\lambda\mathcal{L}^2_{X_0}\lambda + (\mathcal{L}_{X_0}\lambda)^2 + 2\lambda^2\mathcal{L}_{X_0}\lambda + \frac{1}{2E}\mathcal{L}_{X_0}\lambda\right) +$$

$$+\kappa\left(\mathcal{L}^4_{X_0}\lambda + 5(\mathcal{L}_{X_0}\lambda)\mathcal{L}^2_{X_0}\lambda + 7\lambda\mathcal{L}^3_{X_0}\lambda + 8\lambda(\mathcal{L}_{X_0}\lambda)^2 +\right.$$

$$\left.+14\lambda^2\mathcal{L}^2_{X_0}\lambda + \frac{1}{2E}\mathcal{L}^2_{X_0}\lambda + 8\lambda^3\mathcal{L}_{X_0}\lambda + \frac{2\lambda}{E}\mathcal{L}_{X_0}\lambda\right) = 0.$$

Using (36), this takes the form

$$\kappa\left[\mathcal{L}^4_{X_0}\lambda + 6\lambda\mathcal{L}^3_{X_0}\lambda + \left(5\mathcal{L}_{X_0}\lambda + 11\lambda^2 + \frac{1}{2E}\right)\mathcal{L}^2_{X_0}\lambda +\right.$$
$$\left. + \left(7\mathcal{L}_{X_0}\lambda + 6\lambda^2 + \frac{3}{2E}\right)\lambda\mathcal{L}_{X_0}\lambda\right] = 0. \tag{42}$$

(C) To conclude the proof, we finally discuss the relation (42).

(a) If $\kappa = 0$, then we see from (41) that $\mathcal{L}_{FX_0}\lambda = 0$. This means by **(A)** that (M, E) is a Berwald manifold.

(b) In the case $\kappa \neq 0$ the second factor has to vanish on the left hand side of (42). Then we take the Lie derivative of this factor with respect to the vector field FX_0. Applying the relations (27)–(33), after a somewhat lengthy but quite straightforward calculation we obtain

$$\left[\mathcal{L}^3_{X_0}\lambda + 3\lambda\mathcal{L}^2_{X_0}\lambda + (\mathcal{L}_{X_0}\lambda)^2 + 2\lambda^2\mathcal{L}_{X_0}\lambda + \frac{1}{2E}\mathcal{L}_{X_0}\lambda\right]\mathcal{L}_{FX_0}\lambda = 0. \tag{43}$$

If $\mathcal{L}_{FX_0}\lambda = 0$, then the process ends. Otherwise the first factor on the left hand side of (43) is zero, but, owing to (41), this also yields the desired relation $\mathcal{L}_{FX_0}\lambda = 0$. ∎

References

[1] S. BÁCSÓ and M. MATSUMOTO, Reduction theorems of certain Landsberg spaces to Berwald spaces, *Publ. Math. Debrecen,* **48** (1996), 337–366.

[2] L. BERWALD, On Finsler and Cartan geometries, III, Two-dimensional Finsler spaces with rectilinear extremals, *Ann. of Math.*, **42** (19–11), 84–112.

[3] W. BLASCHKE and G. BOL, *Geometrie der Gewebe,* Springer Verlag, Berlin, 1935.

[4] J. GRIFONE, Structure presque tangente et connexions I, *Ann. Inst. Fourier, Grenoble* **22**, 1 (1972), 287–334.

[5] J. GRIFONE, Structure presque tangente et connexions II, *Ann. Inst. Fourier, Grenoble* **22**, 3 (1972), 291–338.

[6] J. SZILASI, Notable Finsler connections on a Finsler manifold, *Lect. Mat.,* **10** (1998), 7–34.

[7] J. SZILASI and SZ. VATTAMÁNY, On the projective geometry of sprays, *Differ. Geom. Appl.,* **12** (2000), 183–208.

[8] J. SZILASI and CS. VINCZE, A new look at Finsler connections and special Finsler manifolds, *Acta Math. Acad. Paedagog. Nyházi.,* (N.S.) **16** (2000), 33–63.

[9] SZ. VATTAMÁNY, Projection onto the indicatrix bundle of a Finsler manifold, *Publ. Math. Debrecen,* accepted for publication.

ANNALES UNIV. SCI. BUDAPEST., **44** (2001), 27–31

EXTREMAL FROBENIUS NUMBERS IN SOME SPECIAL CASES

By

G. KISS

Kisújszállás, Hungary

(Received August 20, 2001)

1. Introduction

Let $0 < a_1 < a_2 < \ldots < a_n$ be integers with $\gcd(a_1, \ldots, a_n) = 1$. It is well-known that the equation $N = \sum_{i=1}^{n} x_i a_i$ has a solution in non-negative integers x_i provided N is sufficiently large. Denote $G(a_1, a_2, \ldots, a_n)$ the greatest integer N for which the preceding equation has no such solution. $G(a_1, a_2, \ldots, a_n)$ is called a *Frobenius number.* The computation and estimate of G have given rise to many papers. The question of the estimation of G naturally suggests the following extremal problem [2]. For integers n and t, define $g(n,t)$ by

$$g(n,t) = \max G(a_1, a_2, \ldots, a_n)$$

where the max is taken over all a_i satisfying $1 < a_1 < \ldots < a_n \leq t$, $\gcd(a_1; \ldots; a_n) = 1$. $g(n,t)$ is called an *extremal Frobenius number.*

ERDŐS and GRAHAM proved in [2] that, for $n \geq 2$,

$$\frac{t^2}{n-1} - 5t \leq g(n,t) < 2\frac{t^2}{n}.$$

They found the exact value of $g(n, 2n+k)$ for fixed k, if n is sufficiently large ($n > 20k^2$):

$$(1) \qquad g(n, 2n+k) = \begin{cases} 2n + 4k - 1, & \text{for } k \geq 1 \text{ and } n - k \equiv 1 \pmod 3 \\ 2n + 4k + 1, & \text{for } k \geq 1 \text{ and } n - k \not\equiv 1 \pmod 3). \end{cases}$$

DIXMIER has proved [1, Thm.1] that, for $2 \leq n < t$

$$\left\lfloor \frac{t-2}{n-1} \right\rfloor (t - n + 1) - 1 \leq g(n,t) \leq \left(\left\lceil \frac{t-1}{n-1} \right\rceil - 1 \right) t - 1,$$

thus proving a conjecture by ERDŐS and GRAHAM [3, page 86] stating that

$$g(n,t)\leq \frac{t^2}{(n-1)}.$$

In the same paper DIXMIER improved the upper bound as follows [1, Thm.3]. For $2 \leq n < t$

$$g(n,t)\leq (v-1)(t-r-1)-1, \tag{2}$$

$$\text{where } t-1=v(n-1)-r \text{ and } 0\leq r<n-1.$$

DIXMIER gave the exact value of $g(n,t)$ for some special cases.

In this paper we find the exact value of some further extremal Frobenius numbers and extend the validity of Erdős–Graham's formula (1) for any $n \geq k+2$ using Dixmier's upper bound.

2. Main result

THEOREM. *Let d, n, k be integers such that $2\leq d<n,\ \ 0 \leq k \leq n-d$. If $n-k \equiv 0 \pmod{d+1}$ or $n-k \equiv -1 \pmod{d+1}$ then*

$$g(n, dn+k) = d(d-1)n + 2dk + d^2 - d - 1.$$

PROOF. With the notations of (2), we obtain

$$dn+k-1 = (d+1)(n-1) - n + k + d = (d+1)(n-1) - (n-k-d),$$

if $k \leq n-d$. We have $v = d+1$ and $r = n-k-d$. We can apply formula (2):

$$\begin{aligned} g(n,t) &\leq (v-1)(t-r-1)-1 = d[dn+k-(n-k-d)-1]-1 = \\ &= d(dn+k-n+k+d-1)-1 = \\ &= d[(d-1)n+2k+d-1]-1 = d(d-1)n+2dk+d^2-d-1. \end{aligned}$$

The proof will be complete if we find integers $0 < a_1 < a_2 < \ldots < a_n \leq t$ such that

$$G(a_1, a_2, \ldots, a_n) = d(d-1)n + 2dk + d^2 - d - 1.$$

We consider the cases $n-k \equiv 0$ and $1 \pmod{d+1}$ separately.

Case (i). Let $n-k \equiv 0 \pmod{d+1}$. Write $n = l(d+1)+k$, then

$$dn+k = ld(d+1) + dk + k = (d+1)(ld+k).$$

Let $A = \{a_1; a_2; \ldots; a_n\}$ consist of all multiples of $(d+1)$ and the l largest elements of the residue class (-1) modulo $(d+1)$ up to t:

$$A = \{d+1; 2(d+1); 3(d+1) \ldots; (ld+k-1)(d+1); (ld+k)(d+1);$$
$$dn+k-1; dn+k-1-(d+1); dn+k-1-2(d+1); \ldots; dn+k-1-(l-1)(d+1)\}.$$

It is clear that $|A| = (ld+k)+l = l(d+1)+k = n$. Let $z = dn+k-1-(l-1)(d+1)$ be the smallest element of A, which is not a multiple of $(d+1)$. Let us write z in the form

$$z = dn+k-1-(l-1)(d+1) = (ld+k)(d+1)-1-(l-1)(d+1) =$$
$$= (d+1)(ld+k-l+1)-1.$$

We know that $0; z; 2z; \ldots; (d-1)z; dz$ is a complete residue system mod $(d+1)$. Hence the largest integer, which has no representation by A is (see e.g.[4])

$$G(A) = dz-(d+1) = d(d+1)(ld+k-l+1)-d-d-1 =$$
$$= d(d+1)[(d-1)l+k+1]-2d-1 =$$
$$= d(d+1)(d-1)l+d(d+1)k+d^2+d-2d-1.$$

Substituting $n-k = l(d+1)$, we obtain the desired

$$G(A) = d(d-1)n+2dk+d^2-d-1.$$

Case (ii). Suppose $n-k \equiv -1 \pmod{d+1}$. Then $n-k = l(d+1)-1$, or $n = l(d+1)+k-1$. We see that $dn+k = (d+1)dl+dk-d+k = (d+1)(dl+k)-d$, hence $dn+k-1 = (d+1)(dl+k)-d-1 = (d+1)(dl+k-1)$ is a multiple of $(d+1)$. Let $A = \{a_1; a_2; \ldots; a_n\}$ consist of all multiples of $(d+1)$ and the l largest elements of the residue class (1) modulo $(d+1)$ up to t:

$$A = \{d+1; 2(d+1); 3(d+1) \ldots; (ld+k-1)(d+1);$$
$$dn+k; dn+k-(d+1); dn+k-2(d+1); \ldots; dn+k-(l-1)(d+1)\}.$$

It is obvious that $|A| = (ld+k-1)+l = l(d+1)+k-1 = n$. Let $x = dn+k-(l-1)(d+1)$ be the smallest element of A, which is in the residue class (1) mod $(d+1)$. We write x in the form

$$x = dn+k-(l-1)(d+1) = (ld+k-1)(d+1)+1-(l-1)(d+1) =$$
$$= (d+1)(ld+k-1-l+1)+1 = (d+1)[(d-1)l+k]+1.$$

The proof is carried out analogously to Case (i), since $0; x; 2x; \ldots; (d-1)x; dx$ is a complete system of residues mod $(d+1)$. The largest integer, which has no representation by A is

$$G(A) = d(d+1)(d-1)l+d(d+1)k+d-d-1 =$$
$$= d(d-1)(n-k+1)+d(d+1)k-1 = d(d-1)n+2dk+d^2-d-1.$$

3. Some corollaries

First we apply our Theorem for $d = 2$.

COROLLARY 1. *Let* n, k *be integers such that* $2<n, \quad 0 \le k \le n-2$. *If* $n-k \equiv 0 \pmod 3$ *or* $n-k \equiv -1 \pmod 3$ *then*

$$g(n, 2n+k) = 2n + 4k + 1.$$

This improves the case $n - k \not\equiv 1 \pmod 3$ of the Erdős–Graham result (1) by omitting the premise "n is sufficiently large". For the greatest permissible value $k = n - 2$, we have

$$g(n, 3n-2) = 2n + 4(n-2) + 1 = 6n - 7.$$

The next exact value is obtained by DIXMIER [1, Thm.4]. Since $n - 1$ divides $t - 2 = 3n - 3$, we have

$$g(n, 3n-1) = \frac{(3n-1)(3n-3)}{n-1} - (3n-1) + 1 = 3(3n-1) - 3n + 2 = 6n - 1.$$

Now, take $d = 3$ in the Theorem. We obtain:

COROLLARY 2. *Let* n, k *be integers such that* $2<n, \quad 0 \le k \le n-3$. *If* $n-k \equiv 0 \pmod 4$ *or* $n-k \equiv -1 \pmod 4$ *then*

$$g(n, 3n+k) = 6n + 6k + 5.$$

We can continue this specification and get exact formulae for further *extremal Frobenius numbers.*

COROLLARY 3. *Let* d, n *be integers such that* $d < n$. *If* $n \equiv 0 \pmod{d+1}$ *or* $n \equiv -1 \pmod{d+1}$ *then*

$$g(n, dn) = d(d-1)n + d^2 - d - 1.$$

For example: $g(n, 2n) = 2n + 1$; $g(n, 3n) = 6n + 5$; $g(n, 4n) = 12n + 11$.

Acknowledgement

I would like to thank Professor Róbert Freud for helpful discussions, suggestions.

References

[1] J. DIXMIER, Proof of a conjecture by Erdős and Graham concerning the problem of Frobenius, *J. Number Theory*, **34** (1990), 198–209.

[2] P. ERDŐS and R. L. GRAHAM, On a linear diophantine problem of Frobenius, *Acta Arithmetica,* **21** (1972), 399–408.

[3] P. ERDŐS and R. L. GRAHAM, Old and New Problems and Results in Combinatorial Number Theory, *Monographies de l'Enseignement Mathématiques,* **28** Université de Genève, (1980).

[4] E. S. SELMER, On the linear diophantine problem of Frobenius, *Journal für reine und angewandte Mathematik,* **293/294** (1) (1977), 1–17.

ANNALES UNIV. SCI. BUDAPEST., **44** (2001), 33–38

ON THE "GOOD" NODES OF WEIGHTED LAGRANGE INTERPOLATION FOR EXPONENTIAL WEIGHTS

By

L. SZILI

Department of Numerical Analysis L. Eötvös University, Budapest

(Received October 17, 2001)

In [8] J. SZABADOS established a relation between the "good" nodes of weighted Lagrange interpolation and the weighted Lebesgue constants. In this note we give a generalization of his results for other exponential weights.

Let $X_n := \{x_{n,n} < \ldots < x_{1,n}\} \subset I := \mathbb{R}$ or $(-1,1)$ $(n \in \mathbb{N})$ be an interpolatory matrix and $w : I \to \mathbb{R}$ be a given weight function. It is known that (see e.g. [7] and [10]) the weighted Lebesgue constants $\Lambda_n(w, X_n)$ $(n \in \mathbb{N})$ play a fundamental role in the convergence-divergence behaviour of sequences of weighted Lagrange interpolatory polynomials. $\Lambda_n(w, X_n)$ is defined as the supremum norm on I of the weighted Lebesgue function

$$\lambda_n(w, X_n, x) := \sum_{k=1}^{n} \left| \frac{(\omega_n w)(x)}{(\omega_n w)'(x_{k,n})(x - x_{k,n})} \right| =: \sum_{k=1}^{n} |q_{k,n}(x)| \tag{1}$$

$$(x \in I, n \in \mathbb{N}),$$

where $\omega_n(x) := c_n \prod_{k=1}^{n} (x - x_{k,n})$.

Throughout this paper we shall assume that our weight has the form $w := e^{-Q}$, where $Q : I \to \mathbb{R}$ is even, continuous and convex. The nth Mhaskar–Rahmanov–Saff number $a_n := a_n(w)$ is the (unique) positive root of the equation

$$n = \frac{2}{\pi} \int_0^1 \frac{a_n t Q'(a_n t)}{\sqrt{1 - t^2}} dt \qquad (n \in \mathbb{N}).$$

Research supported by the Hungarian National Scientific Research Foundation (OTKA) under Grant No. T032719, T32872 and by FKFP 0198/1999.

One of its property is that

$$\|r_n w\| := \max_{x\in I} |(r_n w)(x)| = \max_{|x|\le a_n} |(r_n w)(x)| \qquad (r_n \in \mathcal{P}_n,\ n \in \mathbb{N}),$$

where $\mathcal{P}_n$ denotes the polynomials of degree $\le n$.

Our classes of weights on $\mathbb{R}$ are the so-called *Freud-type weights*, where Q is even and polynomial growth at $+\infty$, and the *Erdős-type weights*, where Q is even and of faster than polynomial growth at $+\infty$. The exponential weights on $(-1, 1)$ that we discuss include $w_{k,\alpha} := e^{-Q_{k,\alpha}}$, where $Q_{k,\alpha}(x) := = \exp_k((1 - x^2)^{-\alpha})$. Here $\alpha > 0, k \ge 0$ and $\exp_k := \exp(\exp(\ldots\exp()))$ is the $k \ge 1$-th iterated exponential and $\exp_0(x) := x$. We shall denote these classes of weights by $\mathcal{F}(\mathbb{R})$, $\mathcal{E}(\mathbb{R})$ and $\mathcal{EXP}[-1, 1]$, respectively. For a formal definition of these classes, see [6, Definitions 1–3].

It is not difficult to see that the function $T(x) := 1+(xQ''(x))/Q'(x)$ $(x \in \in (0,+\infty)\cap I)$ guarantees the regular growth of Q. If T is bounded on I then Q is at most polynomial growth on I. This is true for Freud-type weights. In contrast, if $T(x) \to +\infty$ as x tends to the endpoints to I (it is true for $w \in \mathcal{E}(\mathbb{R})$ or $\mathcal{EXP}[-1, 1]$) then Q is faster than polynomial growth near the endpoint of I.

We shall derive the generalizations of results in [8][1]from the following statement.

THEOREM. *Let* $w \in \mathcal{F}(\mathbb{R})$, $\mathcal{E}(\mathbb{R})$ *or* $\mathcal{EXP}[-1, 1]$. *If* $r_n \in \mathcal{P}_n$ *satisfies*

$$\|r_n w\| < e^{c_w \frac{n}{\sqrt{T_n}}} \qquad (T_n := T(a_n)) \tag{2}$$

with a constant $c_w > 0$ *and for a point* $y \in I$ *we have* $(r_n w)(y) = 1$ *then*

$$|y| \le a_n \left(1 + d_w \left(\frac{\log \|r_n w\|}{n\, T_n}\right)^{2/3}\right)$$

with some constant $d_w > 0$ *depending only on* w.

REMARKS. 1. For our weights we have $n/\sqrt{T_n} \to +\infty$ if $n \to +\infty$. Indeed, if $w \in \mathcal{F}(\mathbb{R})$ then there exist $1 < A \le B$ such that $A \le T(x) \le B$ $(x \in \in (0,+\infty))$, see the definition of $\mathcal{F}(\mathbb{R})$ in [6, p. 153]. For $w \in \mathcal{E}(\mathbb{R})$ we know that for any $\varepsilon > 0$ there exists $c > 0$ independent of n such that $T_n \le cn^{\varepsilon}$

[1] Actually formula (2) in [8] holds only for $a_n \le |x| \le ra_n$ with some constant $r > 1$. That means the [8] Proposition 1 is true with some additional condition.

($n \in \mathbb{N}$) (see [2, (2.7)]). Finally, if $w \in \mathcal{EXP}[-1,1]$ then for some $\varepsilon > 0$ we have $T_n = O(n^{2-\varepsilon})$ ($n \to +\infty$), see [4, (3.8)].

2. Actually, the starting point of [8] was the problem of the behaviour of the point systems X_n ($n \in \mathbb{N}$) for which the weighted fundamental polynomials $q_{k,n}$ (see (1)) are uniformly bounded, i.e. $|q_{k,n}(x)| \le A$ uniformly in x, k and n. These point systems serve as a basis of constructing convergent weighted interpolation polynomials of degree at most $n(1+\varepsilon)$ (see [12] and [9]).

From the Theorem and Remark 1 we immediately obtain

COROLLARY 1. *Let $w \in \mathcal{F}(\mathbb{R})$, $\mathcal{E}(\mathbb{R})$ or $\mathcal{EXP}[-1,1]$ and suppose that the point system X_n ($n \in \mathbb{N}$) is such that the corresponding weighted fundamental polynomials of Lagrange interpolation $q_{k,n}$ ($k = 1,2,\dots,n, n \in \mathbb{N}$) are uniformly bounded. Then there exists $c > 0$ such that*

$$\max_{1\le k\le n} |x_{k,n}| \le a_n \left(1 + c\left(\frac{1}{n\,T_n}\right)^{2/3}\right) \qquad (n \in \mathbb{N}).$$

3. Now let us consider the Lebesgue constants $\Lambda_n(w, X_n)$ ($n \in \mathbb{N}$). Let $y_n \in \mathbb{R}$ be such that $\Lambda_n(w, X_n) = \lambda_n(w, X_n, y_n)$, and consider the weighted polynomial

$$(r_n w)(x) := \sum_{k=1}^{n} \left(\operatorname{sgn} q_{k,n}(y_n)\right) q_{k,n}(x).$$

Evidently

$$|(r_n w)(x)| \le \Lambda_n(w, X_n, x) \le \Lambda_n(w, X_n) = (r_n w)(y_n),$$

that is $\|r_n w\| = \Lambda_n(w, X_n)$. Since $|(r_n w)(x_{k,n})| = 1$ thus from the Theorem we obtain

COROLLARY 2. *Let $w \in \mathcal{F}(\mathbb{R})$, $\mathcal{E}(\mathbb{R})$ or $\mathcal{EXP}[-1,1]$. Suppose that the constant $c_w > 0$ satisfies* (6) *and the point system X_n ($n \in \mathbb{N}$) is such that*

$$\Lambda_n(w, X_n) < e^{c_w \frac{n}{\sqrt{T_n}}} \qquad (n \in \mathbb{N}).$$

Then there exists $c > 0$ such that

$$\max_{1\le k\le n} |x_{k,n}| \le a_n \left(1 + c\left(\frac{\log \Lambda_n(w, X_n)}{n\,T_n}\right)^{2/3}\right) \qquad (n \in \mathbb{N}).$$

It is known that for the above weights there are such point systems for which $\Lambda_n(w, X_n) \sim \log n$ $(n \in \mathbb{N})$ (see [7], [2], [1]), and the order is the best possible (see [10], [11]). Thus the "best" weighted Lagrange interpolation point systems satisfy

$$\max_{1\le k\le n} |x_{k,n}| \le a_n \left(1 + c\left(\frac{\log\log n}{n\,T_n}\right)^{2/3}\right) \qquad (n \in \mathbb{N})$$

with some constant $c > 0$ depending only on w.

Proof of Theorem. Fix a weight function w. Then for every $r_n \in \mathcal{P}_n$ $(n \in \mathbb{N})$ we have

$$|(r_n w)(x)| \le e^{n\,U_{n,a_n}\left(\frac{|x|}{a_n}\right)} \|r_n w\| \qquad (x \in \mathbb{R}), \tag{3}$$

where $U_{n,R}$ is the "majorizing function" (see [3, Lemma 7.1], [5, (4.11)], [4, (5.11)]).

Now we show that there exist the constants $c > 0$ and $D > 0$ (they depend only on w) such that

(4)

$$U_{n,a_n}(1+\varepsilon) \le -c \begin{cases} \varepsilon^{3/2} T_n, & \text{if } 0 \le \varepsilon \le \frac{D}{T_n} \\ \dfrac{\varepsilon^2}{1+\varepsilon} T_n^{3/2}, & \text{if } \frac{D}{T_n} \le \varepsilon < \begin{cases} +\infty, & \text{if } I = \mathbb{R} \\ \frac{1}{a_n} - 1, & \text{if } I = (-1,1) \end{cases} \end{cases}$$

For the interval $\varepsilon \in [0, D/T_n)$ this statement is Lemma 7.1(d) in [3] if $w \in \mathcal{F}(\mathbb{R})$; Theorem 4.3 of [5] for $w \in \mathcal{E}(\mathbb{R})$ and Theorem 5.3 in [4] if $w \in \mathcal{EXP}[-1,1]$.

Now let $\varepsilon \ge D/T_n$. Then there is a $\xi \in (a_n, a_n(1+\varepsilon)) \subset I$ such that

$$U_{n,a_n}(1+\varepsilon) \le -\frac{Q''(\xi)(a_n\varepsilon)^2}{2n} \tag{5}$$

(see [5, p. 228], [4, p. 53–56], [3, Lemma 7.1] with $R = a_n$ and use the inequality $\log(1+x) \le x$ if $x \ge 0$).

If $w \in \mathcal{F}(\mathbb{R})$ then there are $c_1, c_2 > 0$ such that $c_1 Q'(\xi) \le \xi Q''(\xi) \le c_2 Q'(\xi)$ (see [6, Definition 1]). Since $a_n Q'(a_n) \sim n$ $(n \in \mathbb{N})$ (see [3, (5.5)]) thus we have

$$\frac{a_n^2 Q''(\xi)}{n} \ge c_1 \frac{a_n Q'(a_n)}{n} \cdot \frac{1}{1+\varepsilon} \ge \frac{c_3}{1+\varepsilon}$$

which proves (4) for Freud type weights. If $w \in \mathcal{E}(\mathbb{R})$ or $\mathcal{EXP}[-1,1]$ then we obtain (4) in a similar way using that $x\,Q''(x)$ is increasing for $x \in (0,+\infty)\cap I$ (see [5, Lemma 2.1(ii)], [4, Lemma 3.1(ii)]) and

$$a_n^2 Q''(a_n) \sim n\,T_n^{3/2} \qquad (n \in \mathbb{N})$$

(see [5, Lemma 2.2(i)], [4, Lemma 3.2(i)]).

Since $\varepsilon^{3/2} T_n \sim \frac{\varepsilon^2}{1+\varepsilon} T_n^{3/2}$ $(n \in \mathbb{N})$ if $\varepsilon = D/T_n$ thus from (4) we obtain that there exists $c_w > 0$ such that

$$(6) \qquad |(r_n w)(x)| \le e^{-c_w \frac{n}{\sqrt{T_n}}} \|r_n w\| \qquad \left(\frac{|x|}{a_n} - 1 \ge \frac{D}{T_n}, r_n \in \mathcal{P}_n, n \in \mathbb{N}\right).$$

If $r_n \in \mathcal{P}_n$ satisfies (2) then $|(r_n w)(x)| < 1$ for $|x|/a_n - 1 \ge D/T_n$, i.e. $|y|/a_n - 1 < D/T_n$. From (3) and (4) we obtain that

$$1 = |(r_n w)(y)| \le e^{-c\left(\frac{|y|}{a_n}-1\right)^{3/2} n\,T_n} \|r_n w\|.$$

Hence, a simple rearrangement yields the statement of the Theorem. ∎

Acknowledgment

The author expresses his thanks to Professor Péter Vértesi for pointing out this topic and for his helpful remarks.

References

[1] S. DAMELIN, The weighted Lebesgue constant of Lagrange interpolation for exponential weights on $[-1,1]$, *Acta Math. Hungar.*, **81 (3)** (1998), 223–240.

[2] S. DAMELIN, The Lebesgue functions and Lebesgue constant of Lagrange interpolation for Erdős weights, *J. Approx. Theory*, **94** (1998), 235–262.

[3] A. L. LEVIN and D. S. LUBINSKY, Christoffel functions, orthogonal polynomials, and Nevai's conjecture for Freud weights, *Constr. Approx.*, **8** (1992), 463–535.

[4] A. L. LEVIN and D. S. LUBINSKY, *Christoffel Functions and Orthogonal Polynomials for Exponential Weights on* $[-1,1]$, Mem. Amer. Math. Soc. 535, Vol. 111 (1994).

[5] A. L. LEVIN, D. S. LUBINSKY and T. Z. MTHEMBU, Christoffel functions and orthogonal polynomials for Erdős weights on $(-\infty, \infty)$, *Rendiconti di Matematica, Serie VII*, Vol. 14, Roma (1994), 199–289.

[6] D. S. LUBINSKY and D. MACHE, $(C, 1)$ means of orthonormal expasion for exponential weights, *J. Approx. Theory*, **103** (2000), 151–182.

[7] J. SZABADOS, Weighted Lagrange and Hermite–Fejér interpolation on the real line, *J. of Inequal. and Appl.*, **1** (1997), 99–112.

[8] J. SZABADOS, Where are the nodes of "good" interpolation polynomials on the real line, *J. Approx. Theory*, **103** (2000), 357–359.

[9] L. SZILI and P. VÉRTESI, An Erdős type convergence process in weighted interpolation. II. (Exponential weights on $[-1, 1]$), *Acta Math. Hungar.* (submitted)

[10] P. VÉRTESI, On the Lebesgue function of weighted Lagrange interpolation. I. (Freud-type weights) *Constr. Approx.*, **15** (1999), 355–367.

[11] P. VÉRTESI, On the Lebesgue function of weighted Lagrange interpolation II., *J. Austral. Math. Soc. (Series A)*, **65** (1998), 145–162.

[12] P. VÉRTESI, An Erdős-type convergence process in weighted interpolation. I. (Freud-type weights), *Acta Math. Hungar.*, **91 (3)** (2001), 195–215.

ANNALES UNIV. SCI. BUDAPEST., **44** (2001), 39–52

MODIFIED WEIGHTED $(0;0,2)$-INTERPOLATION ON INFINITE INTERVAL $(-\infty,+\infty)$

By

P. MATHUR and S. DATTA

Department of Mathematics and Astronomy, Lucknow University, Lucknow, India

(Received October 15, 1999)

1. Introduction

In 1975, L. G. PÁL [8] introduced the following interpolation process.

Let

$$-\infty < x_{n,n} < \ldots < x_{1,n} < +\infty \tag{1.1}$$

be a system of distinct real points and put

$$W_n(x) = \prod_{i=1}^{n} (x - x_{i,n}). \tag{1.2}$$

The roots $y_{i,n}$ $(i = 1, \ldots, n-1)$ of $W_n'(x)$ are interscaled between the roots of $W_n(x)$, i.e.,

$$-\infty < x_{n,n} < y_{n-1,n} < \ldots < x_{2,n} < y_{1,n} < x_{1,n} < +\infty. \tag{1.3}$$

Pál proved that forgiven arbitrary numbers $\{\alpha_{i,n}\}_{i=1}^{n}$ and $\{\beta_{i,n}\}_{i=1}^{n-1}$ there exists a unique polynomial of degree $\leq 2n-1$ satisfying the conditions:

$$R_n(x_{i,n}) = \alpha_{i,n}, \quad i = 1, \ldots, n; \qquad R_n'(y_{i,n}) = \beta_{i,n}, \quad i = 1, \ldots, n-1 \tag{1.4}$$

and an initial condition $R_n(a) = 0$, where a is a given point, different from the nodal points $x_{i,n}$ $(i = 1, \ldots, n,\ n = 1, 2, \ldots)$.

SZILI [10], firstly applied this method in the case for $W_n(x) = H_n(x)$ (here H_n denotes the nth Hermite polynomial). Taking n even, he proved the existence, uniqueness, explicit representation and the convergence theorem for the polynomial $R_n(x)$ satisfying the conditions (1.4) together with an additional condition

$$R_n(0) = -\sum_{i=1}^{n} 2\alpha_{i,n}\left[\frac{H_n(0)}{H_n'(x_{i,n})}\right]^2. \tag{1.5}$$

If n is odd, the uniqueness fails to hold. Later I. JOÓ [5] improved his results by sharpening the estimates of the fundamental polynomials.

In a recent paper, Z. F. SEBESTYÉN [9] modified the the above methods by replacing the special condition (1.5) by an interpolatorial condition

$$R_n(0) = \alpha_{0,n} \qquad \text{for } n \text{ even} \tag{1.6}$$

or

$$R_n'(0) = \beta_{n,n} \qquad \text{for } n \text{ odd.} \tag{1.7}$$

Let

(1.8)

$$\overline{A}_{i,n}(x) = \frac{H_n'(x)}{H_n'(x_{i,n})} l_{i,n}(x) + 2n\frac{H_n(x)}{H_n'(x_{i,n})}\int_0^x l_{i,n}(t)dt - 2\left[\frac{H_n(x)}{H_n'(x_{i,n})}\right]^2 + \\ + \frac{2H_n(0)H_n(x)}{H_n'(x_{i,n})^2} + \frac{H_n'(0)H_n(x)}{x_{i,n}H_n'(x_{i,n})^2}, \qquad i = 1, \dots, n$$

$$\overline{B}_{i,n}(x)\frac{H_n(x)}{H_n(y_{i,n})}\int_0^x L_{i,n}(t)dt, \qquad i = 1, \dots, n-1. \tag{1.9}$$

For n even, taking

$$\overline{A}_{0,n}(x) = \frac{H_n(x)}{H_n(0)} \tag{1.10}$$

he showed that

$$\overline{R}_n(x) = \sum_{i=0}^{n} \alpha_{i,n}\overline{A}_{i,n}(x) + \sum_{i=1}^{n-1} \beta_{i,n}\overline{B}_{i,n}(x) \tag{1.11}$$

is the uniquely determined polynomial of degree $\leq 2n-1$ satisfying the conditions (1.4) and (1.6). For n odd, taking

$$\overline{B}_{n,n}(x) = \frac{H_n(x)}{H_n'(0)} \tag{1.12}$$

and

$$\overline{A}_{i,n}(x) = \frac{H_n'(x)}{H_n'(x_{i,n})} l_{i,n}(x) + 2n \frac{H_n(x)}{H_n'(x_{i,n})} \int_0^x l_{i,n}(t)dt - 2 \left[\frac{H_n(x)}{H_n'(x_{i,n})}\right]^2, \tag{1.13}$$

$$i = 1, \ldots, n$$

for $x_{i,n} = 0$, he showed that

$$\overline{R}_n(x) = \sum_{i=1}^{n} \alpha_{i,n} \overline{A}_{i,n}(x) + \sum_{i=1}^{n} \beta_{i,n} \overline{B}_{i,n}(x) \tag{1.14}$$

is the uniquely determined polynomial of degree $\leq 2n - 1$ satisfying the conditions (1.4) and (1.7). He also proved that for a function $f : \mathbf{R} \to \mathbf{R}$ which is continuously differentiable satisfying

$$\lim_{|x|\to\infty} x^{2r} f(x) e^{-x^2/2} = 0, \quad r = 0, 1, 2, \ldots, \qquad \lim_{|x|\to\infty} f'(x) e^{-x^2/2} = 0 \tag{1.15}$$

together with, for n even

$$\begin{aligned} \alpha_{i,n} &= f(x_{i,n}), \qquad i = 0, \ldots, n \\ \beta_{i,n} &= f'(x_{i,n}), \qquad i = 1, \ldots, n-1 \end{aligned} \tag{1.16}$$

and for n odd

$$\begin{aligned} \alpha_{i,n} &= f(x_{i,n}), \qquad i = 1, \ldots, n. \\ \beta_{i,n} &= f'(x_{i,n}), \qquad i = 1, \ldots, n \end{aligned} \tag{1.17}$$

following estimate holds

$$e^{-x^2} |f(x) - R_n(x)| = O(1)\omega\left(f', \frac{1}{\sqrt{n}}\right), \qquad x \in \mathbf{R}, \tag{1.18}$$

where O does not depend on n and x. $\omega(f', \cdot)$ is the Freud's modulus of continuity for f'.

J. Balázs [2] and L. Szili [11] have earlier studied analogous modified problems for weighted $(0, 2)$-interpolation and T. F. Xie [12], L. G. Pál [7] have investigated such modifications in Pál type interpolation.

Considering the nodes as the interscaled zeros of $H_n(x)$ and $H_n'(x)$, Srivastava and Mathur [6], taking n even, proved that there exists a unique polynomial of degree $\leq 3n - 2$ sates ing the conditions:

$$\begin{aligned} &G_n(x_{i,n}) = g_{i,n}, \qquad i = 1, \ldots, n, \\ &G_n(y_{i,n}) = g_{i,n}^*, \quad (wG_n)''(y_{i,n}) = g_{i,n}^*, \qquad i = 1, \ldots, n-1, \end{aligned} \tag{1.19}$$

where $w(x) = e^{-x^2/2}$ and

$$G_n'(0) = \sum_{i=1}^{n} g_{i,n} \frac{H_n''(0) l_{i,n}^2(0)}{H_n'(x_{i,n})}. \tag{1.20}$$

For n odd the uniqueness is not true. They also proved that for a function $f : \mathbf{R} \to \mathbf{R}$ satisfying the requirements (1.15), together with

$$\begin{aligned} &g_{i,n} = f(x_{i,n}), \qquad i = 1, \ldots, n, \\ &g_{i,n}^{*} = f(y_{i,n}) \quad \text{and} \quad g_{i,n}^{**} = O\left(\sqrt{n} e^{\delta y_{i,n}^2} \omega \left(f', \frac{1}{\sqrt{n}}\right)\right) \end{aligned} \tag{1.21}$$

the following estimate holds:

$$e^{-\gamma x^2} |f(x) - G_n(f, x)| = O(\log n) \omega \left(f', \frac{1}{\sqrt{n}}\right), \quad \gamma > \frac{3}{2}. \tag{1.22}$$

In [6] results have been obtained under a special condition (1.20), which looks to be artificial. Also it has been proved that for n odd, either the interpolatory polynomial does not exit or if it exists, they are infinitely many. In this connection we raise the following:

PROBLEM. For each positive integer n do there exist a weighted $(0; 0, 2)$-interpolatory polynomial G_n of degree $\leq 3n - 2$ satisfying the conditions (1.19) and

$$G_n(0) = g_0 \qquad \text{if } n \text{ is even} \tag{1.23}$$

or

$$G_n'(0) = g_0' \qquad \text{if } n \text{ is odd.} \tag{1.24}$$

If it exists what will be its explicit form and does it converse?

In this paper, we answer this problem in affirmative. In section 2, we give some preliminaries and state new results in section 3. The estimates of the fundamental polynomials and the convergence theorem have been proved in sections 4 and 5 respectively.

2. Preliminaries

Let $H_n(x)$ be the n^{th} Hermite polynomial with usual normalization

$$\int_{-\infty}^{+\infty} H_n(t)H_m(t)e^{-t^2}dt = \sqrt{\pi}2^n n!\delta_{n,m}, \qquad (n, m \in \mathbf{N}) \tag{2.1}$$

which satisfies the differential equation;

$$\begin{aligned} H_n''(x) - 2xH_n'(x) + 2nH_n(x) = 0, \\ H_n'(x) = 2nH_{n-1}(x). \end{aligned} \tag{2.2}$$

It is well known that $x_{i,n}$ roots of $H_n(x)$ satisfy the following relations:

$$\begin{aligned} &-\infty < x_{n,n} < \ldots < x_{\frac{n}{2}+1,n} < 0 < x_{\frac{n}{2},n} < \ldots < x_{1,n} < +\infty \ (n = 2m), \\ &-\infty < x_{n,n} < \ldots < x_{\frac{n+1}{2},n} = 0 < \ldots < x_{1,n} < +\infty \quad (n = 2m+1), \\ &x_{i,n} = -x_{n-i+1,n}, \qquad \left(i = 1, 2, \ldots, \frac{n}{2}\right). \end{aligned} \tag{2.3}$$

Let $l_{i,n}$ and $L_{i,n}$ denote the Lagrange fundamental polynomial corresponding to the nodal points $x_{i,n}$ and $y_{i,n}$ respectively, then

$$l_{i,n}(x) = \frac{H_n(x)}{H_n'(x_{i,n})(x - x_{i,n})}, \qquad i = 1, \ldots, n, \tag{2.4}$$

$$L_{i,n}(x) = \frac{H_n'(x)}{H_n''(y_{i,n})(x - y_{i,n})}, \qquad i = 1, \ldots, n-1, \tag{2.5}$$

For the roots of $H_n(x)$, we have

$$x_{i,n} \sim \frac{i^2}{n}, \tag{2.6}$$

$$H_n(x) = O(1)n^{-1/4}\sqrt{2^n n!}\left(1 + \sqrt[3]{|x|}\right)e^{x^2/2}, \quad x \in \mathbf{R}, \tag{2.7}$$

$$\sum_{i=1}^{n} e^{-\delta x_i^2} = O(\sqrt{n}) \text{ and } \sum_{i=1}^{n-1} e^{-\delta y_i^2} = O(\sqrt{n}), \quad \delta > 0, \tag{2.8}$$

$$\sum_{i=1}^{n} e^{x_{i,n}^2} l_{i,n}^2(x) = O(e^2) \text{ and } \sum_{i=1}^{n-1} e^{y_{i,n}^2} L_{i,n}^2(x) = O(e^{x^2}) \tag{2.9}$$

$$|H_n(0)| = \frac{n!}{\left(\frac{n}{2}\right)!} \quad \text{for } n \text{ even} \tag{2.10}$$

$$\frac{2^n\left(\left(\frac{n}{2}\right)!\right)^2}{(n+1)!} \sim n^{-1/2}. \tag{2.11}$$

The above results have been taken from [6], we shall also require the following estimates given by SEBESTYÉN [9]

$$\sum_{i=1}^{n} e^{x_{i,n}^2/2}|\overline{A}_{i,n}(x)| = O(\sqrt{n})e^{x^2} \tag{2.12}$$

and

$$\sum_{i=1}^{n} e^{x_{i,n}^2/2}|\overline{B}_{i,n}(x)| = O(1)e^{x^2}, \tag{2.13}$$

where $\overline{A}_{i,n}(x)$ and $\overline{B}_{i,n}(x)$ are given by (1.8) and (1.9) respectively. We shall use the following notations in the sequel $x_i = c_{i,n}$, $l_i = l_{i,n}$, $\overline{A}_i = \overline{A}_{i,n}$, $\overline{B}_i = \overline{B}_{i,n}$.

3. New Results

THEOREM 1. *Considering* (1.3) *as the roots of nodal points and the weight function* $w(x) = e^{-x^2/2}$, *there exists a unique polynomial* G_n *of degree* $\leq$ $\leq 3n-2$ *satisfying the conditions* (1.19) *and* (1.23) *or* (1.24) *according as n is even or odd.*

THEOREM 2. *Let*

$$A_i(x) = \frac{H_n'(x)}{H_n'(x_i)}l_i^2(x) - 2\frac{H_n(x)H_n'(x)}{H_n'(x_i)^3}\int_0^x \frac{H_n'(x_i)l_i'(t) - x_iH_n'(t)}{t-x_i}dt, \tag{3.1}$$
$$i = 1, \ldots, n,$$

$$\begin{aligned} B_i(n) = {} & \frac{H_n(x)}{H_n(y_i)}L_i^2(x) + \frac{H_n(x)H_n'(x)}{2nH_n(y_i)^2}\left[\int_0^x \frac{L_i'(t) - y_iL_i(t)}{t-y_i}dt - \right.\\ & \left. -\frac{2n+1-y_i^2}{2}\int_0^x L_i(t)dt - \frac{H_n'(0)}{2ny_i^2H_n(y_i)}\right], \quad i = 1\ldots, n-1 \end{aligned} \tag{3.2}$$

and

$$C_i(x) = -\frac{H_n(x)H_n'(x)}{4nw(y_i)H_n(y_i)^2}\int_0^x L_i(t)dt, \qquad i = 1, \dots, n-1. \tag{3.3}$$

For n even, let

$$A_0(x) = \frac{H_n(x)H_n'(x)^2}{4n^2x^2H_n(0)^3} + \frac{H_n(x)H_n'(x)}{4n^2H_n(0)^3}\int_0^x \frac{H_n'(t)+2ntH_n(t)}{t^3}dt \tag{3.4}$$

then

$$G_n(x) = \sum_{i=0}^{n} g_i A_i(x) + \sum_{i=1}^{n-1} g_i^* B_i(x) + \sum_{i=1}^{n-1} g_i^{**} C_i(x) \tag{3.5}$$

is the uniquely determined polynomial of degree $\leq 3n-2$ satisfying the conditions (1.19) *and* (1.23).

For n odd, let

$$B_n(x) = \frac{H_n(x)H_n'(x)}{H_n'(0)^2} \tag{3.6}$$

then

$$G_n^*(x) = \sum_{i=1}^{n} g_i A_i(x) + \sum_{i=1}^{n} g_i^* B_i(x) + \sum_{i=1}^{n-1} g_i^{**} C_i(x) \tag{3.7}$$

is the uniquely determined polynomial of degree $\leq 3n-2$ satisfying the conditions (1.19) *and* (1.24).

THEOREM 3. *If $f : \mathbf{R} \to \mathbf{R}$ is a continuously differentiable function satisfying the requirements* (1.15) *and* (1.21), *then*

$$e^{-3x^2/2}|f(x) - G_n(f,x)| = O\left(\omega\left(f', \frac{1}{\sqrt{n}}\right)\right), \quad \text{for even } n \tag{3.8}$$

and

$$e^{-3x^2/2}|f(x) - G_n^*(f,x)| = O\left(\omega\left(f', \frac{1}{\sqrt{n}}\right)\right), \quad \text{for } n \text{ odd.} \tag{3.9}$$

We will prove only our main Theorem 3 as the proof of other theorems is quite similar to that of theorems in [1].

4. Basic Estimates of Fundamental Polynomials (n even)

LEMMA 1. *For n even*

$$|A_0(x)| = O(1)e^{3x^2/2}, \tag{4.1}$$

$$\sum_{i=1}^{n} e^{x_i^2}|A_i(x)| = O(\sqrt{n})e^{3x^2/2}, \tag{4.2}$$

where $A_i(x)$ and $A_0(x)$ are given by (3.1) *and* (3.4) *respectively.*

PROOF. Integrating the last term of (3.4), by parts and using

$$\lim_{t\to 0}\frac{H_n'(t)+2ntH_n(t)}{t^2} = \lim_{t\to 0}\frac{H_n''(t)+2ntH_n'(t)+2nH_n(t)}{2t} =$$

$$= \lim_{t\to 0}\frac{(n+1)H_n'(t)}{t} = 0$$

together with (2.2) and (2.7), we get (4.1).

For n even, $A_i(x)$, $i = 1,\dots,n$ given by (3.1), can be written in a convenient form as:

$$A_i(x) = \frac{H_n'(x)}{H_n'(x_i)}\left[\frac{H_n'(x)}{H_n'(x_i)}l_i(x) + 2n\frac{H_n(x)}{H_n'(x_i)}\int_0^x l_i(t)dt\ -\right.$$

$$\left. -2\left[\frac{H_n(x)}{H_n'(x_i)}\right]^2 + \frac{2H_n(0)H_n(x)}{H_n'(x_i)^2}\right] \equiv \frac{H_n'(x)}{H_n'(x_i)}\overline{A}_i(x),$$

where $\overline{A}_i(x)$ is given by (1.8). Thus by (2.2) and (2.12), we have

$$\sum_{i=1}^{n} e^{x_i^2}|A_i(x)| = \sum_{i=1}^{n}\frac{|H_{n-1}(x)|}{H_{n-1}(x_i)|}e^{x_i^2/2}e^{x_i^2/2}|\overline{A}_i(x)| = O(\sqrt{n})e^{3x^2/2}, \tag{4.3}$$

which completes the proof of the lemma.

LEMMA 2. *For n even,*

$$\sum_{i=1}^{n} e^{y_i^2}|B_i(x)| = O(\sqrt{n})e^{3x^2/2}, \tag{4.4}$$

where $B_i(x)$ is given by (3.2).

PROOF. By (2.2) and (2.5), it follows that

$$\frac{L_i'(t) - y_i L_i(t)}{t - y_i} = -\frac{L_i''(t)}{2} + y_i L_i'(t) + \frac{H_n''(t)}{H_n''(y_i)} + (1-n)L_i(t). \tag{4.5}$$

Hence $B_i(x)$, given by (3.2), can be written in a convenient form as, for $y_i \neq 0$

$$\begin{aligned} B_i(x) &= \frac{H_n(x)}{H_n(y_i)} \left[L_i^2(x) + \frac{H_n'(x)}{2n H_n(y_i)} \left\{ -\frac{L_i'(x)}{2} + \frac{L_i'(0)}{2} + y_i L_i(x) - \right.\right. \\ &\quad \left.\left. -\frac{H_n'(x)}{2n H_n(y_i)} + (n+2-y_i^2) \int_0^x L_i(t)dt \right\} \right] = \\ &= \frac{H_n(x)}{H_n(y_i)} \left[\frac{L_i^2(x)}{2} + (n+2-y_i^2) \frac{H_n'(x)}{2n H_n(y_i)} \int_0^x L_i(x)dt + \right. \\ &\quad \left. + \frac{H_n(x)}{2H_n(y_i)} L_i(x) + \frac{H_n'(x) H_n(0)}{4n H_n(y_i)^2} \right]. \end{aligned} \tag{4.6}$$

For $y_i = 0$, by (3.2) and (4.5), we have

$$\begin{aligned} B_i(x) &= \frac{H_n(x)}{H_n(y_i)} \left[\frac{L_i^2(x)}{2} + \right. \\ &\quad \left. +(n+2-y_i^2) \frac{H_n'(x)}{2n H_n(y_i)} \int_0^x L_i(t)dt + \frac{H_n(x)}{2H_n(y_i)} L_i(x) \right]. \end{aligned} \tag{4.7}$$

Thus

$$\begin{aligned} \sum_{i=1}^{n} e^{y_i^2} |B_i(x)| &\le \sum_{i=1}^{n-1} \left| \frac{H_n(x)}{2H_n(y_i)} \right| e^{y_i^2} L_i^2(x) + \\ &\quad + \frac{(n+1)}{2n} \sum_{i=1}^{n-1} \frac{|H_n(x) H_n'(x)| e^{y_i^2}}{H_n(y_i)^2} \left| \int_0^x L_i(t)dt \right| + \\ &\quad + \sum_{i=1}^{n-1} |1 - y_i^2| \frac{|H_n(x) H_n'(x)| e^{y_i^2}}{2n H_n(y_i)^2} \left| \int_0^x L_i(t)dt \right| + \end{aligned}$$

$$+\sum_{i=1}^{n-1}\frac{e^{y_i^2}H_n^2(x)}{2H_n(y_i)^2}|L_i(t)|+\sum_{i=1}^{n-1}\frac{|H_n(x)H_n'(x)|e^{y_i^2}}{4ny_iH_n(y_i)^2}|H_n(0)|\equiv$$

$$\equiv I_1+I_2+I_3+I_4+I_5. \tag{4.8}$$

From (2.9), we have

$$I_1=O(1)e^{3x^2/2}\sum_{i=1}^{n-1}e^{-y_i^2/2}=O(\sqrt{n})e^{3x^2/2}. \tag{4.9}$$

By (2.13), we have

$$I_2=O(1)\sum_{i=1}^{n-1}\frac{2n|H_{n-1}(x)|}{|H_n(y_i)|}e^{y_i^2}|\overline{B}_i(x)|=O(\sqrt{n})e^{3x^2/2}, \tag{4.10}$$

where $\overline{B}_i(x)$ is given by (1.9). Since $|(1-y_i^2)|=O(e^{y_i^2/2})$, hence

$$I_3=O(1)\sum_{i=1}^{n-1}e^{y_i^2}\frac{|H_{n-1}(x)|}{|H_n(y_i)|}e^{y_i^2}|\overline{B}_i(x)|=O(\sqrt{n})e^{3x^2/2}, \tag{4.11}$$

$$I_4=O(\sqrt{n})e^{3x^2/2}. \tag{4.12}$$

By (2.10), we have

$$I_5=O(1)\sum_{i=1}^{n-1}e^{y_i^2}\frac{2|H_{n-1}(x)H_n(x)|}{y_iH_n(y_i)^2}\frac{n!}{\left(\frac{n}{2}\right)!}=O(1)e^{3x^2/2}. \tag{4.13}$$

Thus by using (4.9)–(4.13) in (4.8), the lemma follows.

LEMMA 3. *For n even,*

$$\sum_{i=1}^{n-1}e^{y_i^2/2}|C_i(x)|=O\left(\frac{1}{\sqrt{n}}\right)e^{3x^2/2}, \tag{4.14}$$

where $C_i(x)$ is given by (3.3).

PROOF. By (2.13), we have

$$\sum_{i=1}^{n-1}e^{y_i^2/2}|C_i(x)|=\sum_{i=1}^{n-1}\frac{e^{y_i^2/2}|H_{n-1}(x)|}{2|H_n(y_i)|}e^{y_i^2/2}|\overline{B}_i(x)|=O\left(\frac{1}{\sqrt{n}}\right)e^{3x^2/2}.$$

PROOF OF THEOREM 3. (n EVEN). By [3], Theorem 4 and [4] Theorem 1, there exists a polynomial $p_n(x)$ of degree $\leq n$, such that

$$e^{-x^2/2}|f(x)-p_n(x)| = O\left(\frac{1}{\sqrt{n}}\right)\omega\left(f', \frac{1}{\sqrt{n}}\right), \tag{4.15}$$

$$e^{-x^2/2}|f'(x)-p_n'(x)| = O(1)\omega\left(f', \frac{1}{\sqrt{n}}\right). \tag{4.16}$$

Further ([10], Lemma 4), we have for $x \in \mathbf{R}$

$$e^{-x^2/2}|p_n(x)| = O(1), \tag{4.17}$$

$$e^{-x^2/2}|p_n'(x)| = O(1) \tag{4.18}$$

and

$$e^{-x^2/2}|p_n''(x)| = O(1)\sqrt{n}\omega\left(f', \frac{1}{\sqrt{n}}\right) \quad \text{for } |x| < \sqrt{2n+1}. \tag{4.19}$$

From the uniqueness of $G_n(x)$ in (3.5), it follows that

$$p_n(x) = \sum_{i=0}^{n} p_n(x_i)A_i(x) + \sum_{i=1}^{n-1} p_n(y_i)B_i(n) + \sum_{i=1}^{n-1}(wp_n)''(y_i)C_i(x). \tag{4.20}$$

Using lemmas 1, 2, 3 and (4.15)–(4.20), we obtain

$$\begin{aligned} e^{-3x^2/2}|f(x)-G_n(f,x)| &\leq e^{-3x^2/2}|f(x)-p_n(x)| + \\ &+ e^{-3x^2/2}|p_n(x)-G_n(f,x)| \leq O(1)e^{-x^2}\omega\left(f', \frac{1}{\sqrt{n}}\right)\frac{1}{\sqrt{n}} + \\ &+ e^{-3x^2/2}\sum_{i=0}^{n}|p_n(x_i)-f(x_i)||A_i(x)| + \\ &+ e^{-3x^2/2}\sum_{i=1}^{n-1}|p_n(y_i)-f(y_i)||B_i(x)| + \\ &+ e^{-3x^2/2}\sum_{i=1}^{n-1}|(wp_n)''(y_i)-g_i^{**}||C_i(x)| \leq \end{aligned}$$

$$\leq O(1)\left[\omega\left(f', \frac{1}{\sqrt{n}}\right) + e^{-3x^2/2}\sum_{i=1}^{n-1}|g_i^{**}C_i(x)| + \right. \tag{4.21}$$

$$+ e^{-3x^2/2}\sum_{i=1}^{n-1}|w(y_i)p_n''(y_i)||C_i(x)| + e^{-3x^2/2}\sum_{i=1}^{n-1}|w'(y_i)p_n'(y_i)||C_i(x)|+$$

$$\left. + \; e^{-3x^2/2}\sum_{i=1}^{n-1}|w''(y_i)p_n(y_i)||C_i(x)|\right] .$$

By lemma 3 and (4.17)–(4.19), we have

$$e^{-x^2/2}\sum_{i=1}^{n-1}|w''(y_i)p_n(y_i)||C_i(x)| = O(1)\frac{1}{\sqrt{n}}, \tag{4.22}$$

$$e^{-x^2/2}\sum_{i=1}^{n-1}|w'(y_i)p_n'(y_i)||C_i(x)| = O(1)\frac{1}{\sqrt{n}} \tag{4.23}$$

and

$$e^{-x^2/2}\sum_{i=1}^{n-1}|w(y_i)p_n''(y_i)||C_i(x)| = O(1)\omega\left(f', \frac{1}{\sqrt{n}}\right). \tag{4.24}$$

Thus by using (4.22)–(4.24) in (4.21), the theorem follows.

5. Basic Estimates of Fundamental Polynomials (n odd)

LEMMA 4. *For n odd,*

$$\sum_{i=1}^{n} e^{x_i^2}|A_i(x)| = O(\sqrt{n})e^{3x^2/2}, \tag{5.1}$$

where $A_i(x)$ is given by (3.7).

PROOF. Since $x_{\frac{n+1}{2},n} = 0$, is a zero of $H_n(x)$ (n odd) thus for $x_i \neq 0$, we have

$$A_i(x) = \frac{H_n'(x)}{H_n'(x_i)}\overline{A}_i(x). \tag{5.2}$$

Thus by (2.12), we have

$$\sum_{i=1}^{n} e^{x_i^2}|A_i(x)| = \sum_{i=1}^{n}\frac{H_{n-1}(x)}{H_{n-1}(x_i)}e^{x_i^2}|\overline{A}_i(x)| = O(\sqrt{n})e^{3x^2/2}.$$

For $x_i = 0$, by (3.1), we have

$$A_i(x) = \frac{H_n'(x)}{H_n'(x_i)} l_i^2(x) - 2\frac{H_n(x)H_n'(x)}{H_n'(x_i)^2} \int_0^2 \frac{tH_n'(t) - H_n(t)}{t^3} dt. \tag{5.3}$$

Integrating the last term of (5.3), by parts, and using

$$\lim_{t\to 0} \frac{tH_n'(t) - H_n(t)}{t^2} = \lim_{t\to 0} \frac{H_n''(t)}{2} = 0$$

we have, by (1.13)

$$A_i(x) = \frac{H_n'(x)}{H_n'(x_i)} \overline{A}_i(x). \tag{5.4}$$

Thus

$$\sum_{i=1}^{n} e^{x_i^2} |A_i(x)| = O(\sqrt{n}) e^{3x^2/2},$$

hence the lemma is proved.

LEMMA 5. *For n odd,*

$$\sum_{i=1}^{n-1} e^{y_i^2} |B_i(x)| = O(\sqrt{n}) e^{3x^2/2}, \tag{5.5}$$

$$|B_0(x)| = O\left(\frac{1}{\sqrt{n}}\right) e^{3x^2/2}, \tag{5.6}$$

and

$$\sum_{i=1}^{n-1} e^{y_i^2} |C_i(x)| = O\left(\frac{1}{\sqrt{n}}\right) e^{3x^2/2}. \tag{5.7}$$

The proof of this lemma is quite similar to that of lemmas 2 and 3, so we omit details.

The proof of Theorem 3 (n odd) can be obtained following the same steps as in the case of n even. We omit details.

References

[1] BALÁZS, J., Súlyozott $(0,2)$-interpoláció ultraszférikus polinomok gyökein, *MTA III. Oszt. közl.*, **11** (1961), 305–338.

[2] BALÁZS, J., Modified weighted $(0,2)$-interpolation, *Approx. Theory,* Marcel Dekker Inc., New York, 1998, 61–73.

[3] FREUD, G., On two polynomial inequalities, *Acta Math. Acad. Sci. Hungar.,* **22** (1971), 109–116.

[4] FREUD, G., On polynomial approximation with weight $\exp\left(-\frac{1}{2}x^{2k}\right)$, *Acta Math. Acad. Sci. Hungar.,* **24** (1973), 363–371.

[5] JOÓ, I., On Pál interpolation, *Annales Univ. Sci. Budapest., Sect. Math.,* **37** (1994), 247–262.

[6] MATHUR, K. K. and SRIVASTAVA, S., Weighted $(0;0,2)$-interpolation on infinite interval, *Acta. Math. Hung.,* **70** 1996, 57–73.

[7] PÁL, L. G., A general lacunary $(0;0,1)$-interpolation process, *Annales Univ. Budapest., Sect. Comp.,* **16** (1996), 291–301.

[8] PÁL, L. G., A new modification to Hermite–Fejér interpolation, *Analysis Math.,* **1** (1975), 197–205.

[9] SEBESTYÉN, Z. F., Pál type interpolation on the roots of Hermite polynomials, *Pure Math. Appl.,* **9** (1998), 429–439.

[10] SZILI, L., A convergence theorem for the Pál-method of interpolation on the roots of Hermite polynomials, *Analysis Math.,* **11** (1985), 75–84.

[11] SZILI, L., Weighted $(0,2)$-interpolation on the roots of classical orthogonal polynomials, *Bull. of Allahabad Math. Soc.,* **8–9** (1993–94), 1–10.

[12] XIE, T. F., On Pál's problem, *Chinese Quart. J. Math.,* **7** (1992), 48–52.

ANNALES UNIV. SCI. BUDAPEST., **44** (2001), 53–64

ON MODIFIED WEIGHTED $(0, 1, \ldots, r-2, r)$-INTERPOLATION ON AN ARBITRARY SYSTEM OF NODES

By

P. MATHUR and S. DATTA

Department of Mathematics and Astronomy, Lucknow University, Lucknow, India

(Received October 15, 1999)

1. Introduction

Recently, J. BALÁZS [2] showed that there exists a modified weighted $(0,2)$-interpolation polynomial $R_n(x)$ of degree $\leq 2n-1$ satisfying the conditions:

$$\begin{aligned} &R_n(x_{i,n}) = y_{i,n}, \qquad i = 1, \ldots, n, \\ &R_n'(x_{n,n}) = y_{n,n} \\ &(wR_n)''(x_{i,n}) = y_{n,n}'', \quad i = 1, \ldots, n-1 \end{aligned}$$

where $y_{i,n}$, $y'_{n,n}$, $y''_{i,n}$ are arbitrary given real numbers, $x_{1,n}, \ldots, x_{n-1,n}$ are the zeros of the polynomial $W_{n-1}(x)$, i.e.,

$$W_{n-1}(x) = \prod_{i=1}^{n-1} (x - x_{i,n})$$

where $w(x) \in C^{(2)}(a,b)$ is such a weight function which satisfies the conditions:

$$(w\,W_{n-1})''(x_{i,n}) = 0, \qquad i = 1, \ldots, n-1$$

and

$$w(x_{i,n}) \neq 0, \qquad i = 1, \ldots, n-1.$$

Then $R_n(x)$ can be explicitly represented as:

$$R_n(x) = \sum_{i=1}^{n} y_{i,n} A_{i,n}(x) + \sum_{i=1}^{n-1} y''_{i,n} B_{i,n}(x) + y'_{n,n} \overline{A}_{n,n}(x)$$

where $A_{i,n}(x)$, $B_{i,n}(x)$ and $\overline{A}_{n,n}(x)$ are basic interpolation polynomials each of degree $\le 2n-1$.

This motivated us to consider the general problem of determining the modified weighted $(0,1,\ldots,r-2,r)$-interpolation $(r\ge 2)$ polynomial on an arbitrary system of nodes.

2. Definitions and New Results

Modified weighted $(0,1,\ldots,r-2,r)$-interpolation $(r \ge 2)$ means the solution of the following problem:

Let the system of knots

$$-\infty \le a \le x_{n,n} < \ldots < x_{1,n} \le b \le +\infty \qquad (n \in \mathbf{N},\ x_i := x_{i,n}) \tag{2.1}$$

be given in the finite or infinite open interval (a,b) and let $w(x) \in C^{(r)}(a,b)$ be a weight function. Find a polynomial $S_n(x)$ of minimal possible degree satisfying the conditions:

$$S_n^{(m)}(x_i) = y_i^{(m)}, \qquad i = 1,\ldots,n,\ m = 0,\ldots,r-2, \tag{2.2}$$

$$S_n^{(r-1)}(x_n) = y_n^{(r-1)}, \tag{2.3}$$

and

$$\left(w^{r-1}S_n(x)\right)^{(r)}(x_i) = y_i^{(r)}, \qquad i = 1,\ldots,n-1,\ n \in \mathbf{N}, \tag{2.4}$$

where $y_i^{(m)}$, $y_n^{(r-1)}$, $y_i^{(r)}$ are arbitrary given real numbers.

Let $W_{n-1}(x)$ denote a polynomial of degree $\le n-1$ having $x_{i,n}$'s as the zeros, i.e.,

$$W_{n-1}(x) = \prod_{i=1}^{n-1}(x - x_{i,n}). \tag{2.5}$$

If there exists a weight function $w(x) \in C^{(r)}(a,b)$ satisfying the conditions:

$$\{w^{r-1}W_{n-1}^{r-1}(x)\}^{(r)}(x_i) = 0, \qquad i = 1,\ldots,n-1, \tag{2.6}$$

and

$$w(x_i) \neq 0, \qquad i = 1,\ldots,n-1 \tag{2.7}$$

then the following holds.

THEOREM 1. *If there exists a weight function* $w(x) \in C^r(a,b)$ *satisfying the conditions* (2.6) *and* (2.7) *then there exists a unique modified weighted* $(0, 1, \dots, r-2, r)$*-interpolation polynomial* $S_n(x)$ *of degree* $\leq nr-1$ *corresponding to the system of knots* (2.1) *and satisfying the conditions* (2.2), (2.3) *and* (2.4).

We note that, if the zeros of the polynomial $W_{n-1}(x)$ are the zeros of the classical orthogonal polynomials, then the weight functions $w(x) \in C^r(a,b)$ satisfying the conditions (2.6) and (2.7) do always exist. We will show this in the sequel. To prove the above theorem we shall need the following

(2.8) $$W_{n-1}(x_n) \neq 0,$$

(2.9) $$\left(W_{n-1}^{r-1}\right)^{(q)}(x_i) = \begin{cases} 0, & q < r-1 \\ (r-1)!\, W_{n-1}'(x_i)^{r-1}, & q = r-1, \end{cases}$$

and

(2.10) $$\left(l_i^{r-1}\right)^{(q)}(x_i) = \begin{cases} 0, & q < r-1 \\ (r-1)!\, l_i'(x_i)^{r-1}, & q = r-1, \end{cases}$$

where

(2.11) $$l_i(x) = \frac{W_{n-1}(x)}{(x-x_i)W_n'(x_i)} \qquad i = 1, \dots, n-1.$$

3. Determination of fundamental polynomials

Suppose that for the basic system of knots (2.1) the polynomial $W_{n-1}(x)$ of degree $\leq n-1$ is given by (2.5) and there exists a weight function $w(x) \in C^r(a,b)$ satisfying the conditions (2.6) and (2.7).

Let $A_{tk}(x)$ $(k = 1, \dots, n,\ t = 0, \dots, r)$ denote polynomials of degree $\leq nr-1$ satisfying the conditions:

(3.1)
$$\begin{cases} A_{tk}^{(m)}(x_j) = \delta_{jk}\delta_{mt}, & j,k = 1, \dots, n;\ m,t = 0, \dots, r-2 \\ A_{rk}^{(r-1)}(x_n) = 0, & k = 1, \dots, n;\ t = 0, \dots, r-2 \\ (w^{r-1}A_{tk})^{(r)}(x_j) = 0, & k = 1, \dots, n;\ j = 1, \dots, n-1;\ t = 0, \dots, r-2, \end{cases}$$

(3.2) $$\begin{cases} A_{(r-1)n}^{(m)}(x_j) = 0, & j = 1, \dots, n;\ m = 0, \dots, r-2 \\ A_{(r-1)n}^{(r-1)}(x_n) = 1, & \\ (w^{r-1}A_{(r-1)n})^{(r)}(x_j) = 0, & j = 1, \dots, n-1, \end{cases}$$

$$(3.3)\quad \begin{cases} A_{rk}^{(m)}(x_j) = 0, & j=1,\dots,n-1, k=1,\dots,n-1, m=0,\dots,r-2, \\ A_{rk}^{(r-1)}(x_n) = 0, & k=1,\dots,n-1, \\ (w^{r-1}A_{rk})^{(r)}(x_j) = \delta_{jk}, & k=1,\dots,n-1;\ t=0,\dots,r-2. \end{cases}$$

We give.the explicit forms of $A_{tk}(x)$, $k=1,\dots,n$, $t=0,\dots r$ in the following:

LEMMA 1. *If for basic system of knots* (2.1) *there exists a weight function* $w(x) \in C^r(a,b)$ *satisfying the condition* (2.6) *and* (2.7), *then the polynomial* $A_{tk}(x)$, *satisfying the conditions* (3.1) *has the form:*

For $t=0,\dots,r-2$, $k=1,\dots,n-1$, *we have*

$$(3.4)\quad \begin{aligned} A_{tk}(x) &= a_{tk}(x-x_k)^{r-1}(x-x_n)^t l_k^r(x) + \\ &+ W_{n-1}^{r-1}(x)\left[\int_{x_n}^{x}(y-x_n)^{r-1}q_{tk}(y)dy + b_{tk}\int_{x_n}^{x} l_k(y)dy + \right. \\ &\left. + \int_{x_n}^{x}\left(\sum_{j=0}^{r-2} e_{jk}(y-x_n)^j\right) W_{n-1}(y)dy\right] + \sum_{j=t+1}^{r-2} d_{jk}A_{jk}(x), \end{aligned}$$

where, last summation is zero for $t=r-2$,

$$(3.5)\quad a_{tk} = \frac{1}{t!(x_k-x_n)^{r-1}},$$

$$(3.6)\quad \begin{aligned} q_{tk}(x) = \frac{a_{tk}}{W_{n-1}'(x_k)(x-x_k)^{r-t-1}} &\left[\left\{ l_k'(x_k) - \right.\right. \\ &\left.\left. - \sum_{j=1}^{r-t-2} c_{jk}(x-x_k)^{r-t-2}\right\} l_k(x) - l_k'(x)\right], \end{aligned}$$

where c_{ij} *'s are given by, for* $s=1,\dots,(r-t-2)$

$$(3.7)\quad l_k'(x_k)l_k^{(s)}(x_k) - l_k^{(s+1)}(x_k) - \sum_{u=1}^{s}\binom{s}{u} u!c_{uk}l_k^{(s-u)}(x_k) = 0,$$

$$b_{tk}=-\frac{a_{tk}}{(r-t)!w^{r-1}W'_{n-1}(x_k)^{r-1}}\left[\left\{w^{r-1}(x-x_n)^{r-1}l_k^r(x)\right\}_{x_k}^{(r-t)}+\right.$$

$$\text{(3.8)}\qquad +\omega^{r-1}(r-t)(x_k-x_n)^{r-1}\left\{l'_k(x_k)l_k^{(r-t-1)}(x_k)-\right.$$

$$\left.\left.-\sum_{v=1}^{r-t-1}\binom{r-t-1}{v}v!c_{vk}l_k^{(r-t-1-v)}(x_k)-l_k^{(r-t)}(x_k)\right\}\right],$$

$e'_{jk}s$, $j=0,\ldots,r-2$ are given by the equations:
(3.9)

$$\sum_{s=1}^{i}\binom{i}{s}\left(W_{n-1}^{r-1}\right)_{x_n}^{(i-s)}\left[b_{tk}l_k^{(s-1)}(x_n)+\right.$$

$$\left.+\sum_{m=0}^{s-1}\binom{s-1}{m}(W_{n-1}(x))_{x_n}^{(s-1-m)}m!e_{mk}\right]=0,\quad 1\le i\le r-2;$$

and for $i=r-1$,

$$\text{(3.10)}\qquad a_{tk}(r-1)!(x_n-x_k)^t l_k^r(x_n)+$$

$$+W_{n-1}^{r-1}(x_n)\left[b_{tk}l_k^{(r-2)}(x_n)+\sum_{m=0}^{r-2}\binom{r-2}{m}m!(W_{n-1}(x))_{x_n}^{(r-2-m)}e_{mk}\right]=0$$

and d_{jk}'s, $j=t+1,\ldots,r-2$, are given by equations, for $i=t+1,\ldots,r-2$

$$\text{(3.11)}\qquad a_{ik}\binom{i}{t}t!\{(x-x_n)^{r-1}l_k^r(x)\}_{x_k}^{(i-t)}+\sum_{j=t+1}^{i}d_{jk}=0.$$

For $k=n$, we have, when $t=0$

$$\text{(3.12)}\quad A_{0n}(x)=\frac{W_{n-1}^{r-1}(x)}{W_{n-1}^{r-1}(x_n)}+W_{n-1}^{r-1}(x)\int_{x_n}^{x}\left(\sum_{j=0}^{r-2}e_{jk}^{*}(y-x_n)^j\right)W_{n-1}(y)dy,$$

where e^*_{jk}'s, $j = 0, \dots, r-2$ are given by the relation

$$\begin{aligned} &\frac{(W^{r-1}_{n-1}(x))^{(m)}_{x_n}}{W^{r-1}_{n-1}(x_n)} + \\ &+\sum_{s=1}^{m} \binom{m}{s} (W^{r-1}_{n-1}(x))^{(m-s)}_{x_n} \left(\sum_{i=1}^{s} \binom{s}{i} (W_{n-1}(x))^{(s-i)}_{x_n} i!e^*_{ik} \right) = 0. \end{aligned} \tag{3.13}$$

Also for $t = 1, \dots, r-2$, we have

$$A_m(x) = W^{r-1}_{n-1}(x) \int_{x_n}^{x} \left(\sum_{j=t}^{r-2} e^{**}_{jk} (y - x_n)^j \right) W_{n-1}(y)dy \tag{3.14}$$

where r^{**}_{jk}'s, $j = t, \dots, r-2$ are given by the relation

$$\sum_{s=t}^{m} \binom{m}{s} (W^{r-1}_{n-1}(x))^{(m-s)}_{x_n} \left(\sum_{i=t}^{s} \binom{s}{i} (W_{n-1}(x))^{(s-i)}_{x_n} i!e^{**}_{ik} \right) = \delta_{mt} \tag{3.15}$$

for $m = t, \dots, r-2$.

PROOF. First, we determine $A_{r-2,k}(x)$, by (3.4), for which the last summation vanishes. Then A_{tk}, $t = 0, \dots, r-3$, $k = 1, \dots, n-1$, can be determined, in the terms of $A_{r-2,k}(x)$, from the recurrence relation (3.4).

Since q_{tk}, given by (3.6), is a polynomial of degree $\leq n-2$, then $A_{tk}(x)$, given by (3.4), is a polynomial of degree $\leq rn-1$. Obviously, by Leibnitz's theorem, for $x = x_j$, $j = 1, \dots, n-1$, by (2.9) and (2.10), we have for $k = 1, \dots, n-1$, $t, m = 0, \dots, r-2$,

$$A^{(m)}_{tk}(x_j) = \begin{cases} 0, & t > m \\ \delta_{jk} & t = m \\ 0 & t < m, \text{ due to (2.10).} \end{cases}$$

Now, for $x = x_n$, we have

$$\begin{aligned} A^{(m)}_{tk}(x_n) = a_{tk} &\sum_{s=1}^{m} \binom{m}{s} \{(x - x_n)^{r-1}\}^{(s)}_{x_n} \{(x - x_k)^t l^r_k(x)\}^{(m-s)}_{x_n} + \\ &+\sum_{s=1}^{m} \binom{m}{s} \left(W^{r-1}_{n-1}(x) \right)^{(m-s)}_{x_n} \Bigg[b_{kt} l^{(s-1)}_k(x_n) + \\ &+\sum_{i=1}^{s} \binom{s-1}{i} \left(W_{n-1}(x) \right)^{(s-1-i)}_{x_n} i!e_{ik} \Bigg]. \end{aligned}$$

For $1 \le m \le r-2$, by (3.9), we have $A_{tk}^{(m)}(x_n) = 0$. For $m = r-1$, we have

$$A_{tk}^{(r-1)}(x_n) = a_{tk}(r-1)!(x_n - x_k)^{r-2} l_k^r(x_n) +$$

$$+ W_{n-1}^{r-1}(x_n) \left[b_{tk} l_k^{(r-2)}(x_n) + \sum_{i=0}^{r-2} \binom{r-2}{i} \left(W_{n-1}(x) \right)_{x_n}^{(r-2-m)} i! e_i \right] = 0,$$

due to (3.10).

If the weight function $w(x) \in C^r(a,b)$ satisfies the condition (2.6) and (2.7), then for $j \ne k$, (3.6), (2.9), we have

$$\left(w^{r-1} A_{tk}(x) \right)_{x_j}^{(r)} = a_{tk} \left\{ w^{r-1}(x - x_n)^{r-1}(x - x_k)^t l_k^r(x) \right\}_{x_j}^{(r)} +$$

$$+ w^{r-1} r! W_{n-1}'(x_j)^{r-1}(x_j - x_n)^{r-1} q_{tk}(x) =$$

$$= r! a_{tk} w^{r-1}(x_j - x_n)^{r-1} \left[(x_j - x_k)^t l_k'(x_j)^r - \right.$$

$$\left. - \frac{W_{n-1}'(x_j)^{r-1} l_k'(x_j)}{W_{n-1}'(x_k)^{r-1}(x_j - x_k)^{r-t-1}} \right] = 0$$

due to

$$l_k'(x_j) = \frac{W_{n-1}'(x_j)}{(x_j - x_k) W_{n-1}'(x_k)}, \qquad j = 1, \ldots, n-1.$$

In the case $j = k$, we have

$$l_k(x_j) = 1, \;\; W_{n-1}(x_k) = 0 \text{ and } \left(w^{r-1} W_{r-1}^{n-1}(x) \right)_{x_k}^{(r)} = 0, \;\; k = 1, \ldots, n-1.$$

Also,

$$\lim_{x \to x_k} \frac{1}{(x - x_k)^{r-t-1}} \left[\left\{ l_k'(x_k) - \sum_{j=1}^{r-t-2} c_{jk}(x - x_k)^{r-t-2} \right\} l_k(x) - l_k'(x) \right] =$$

$$= \frac{1}{(r-t-1)!} \left[l_k'(x_k) l_k^{(r-t-1)}(x_k) - \right.$$

$$\left. - \sum_{v=1}^{r-t-1} \binom{r-t-1}{v} v! c_{vk} l_k^{(r-t-1-v)}(x_k) - l_k^{(r-t)}(x_k) \right].$$

Then by (2.9), we have

$$\left(w^{r-1}A_{tk}\right)^{(r)}(x_k)=\frac{a_{tk}r!}{(r-t)!}\{w^{r-1}(x-x_n)^{r-1}l_k^r(x)\}_{x_k}^{(r-t)}+$$

$$+r!w^{r-1}W'_{n-1}(x_k)^{r-1}\left[b_{tk}+\frac{a_{tk}(x_k-x_n)^{r-1}}{(r-t-1)!\,W'_{n-1}(x_k)^{r-1}}\cdot\right.$$

$$\cdot\left\{l'_k(x_k)l_k^{(r-t-1)}(x_k)-\right.$$

$$\left.\left.-\sum_{v=1}^{r-t-1}\binom{r-t-1}{v}v!c_{vk}l_k^{(r-t-1-v)}(x_k)-l_k^{(r-t)}(x_k)\right\}\right]=0.$$

This equality holds if we replace b_{tk} by the expression (3.8). Hence $A_{tk}(x)$, $t=0,\dots,r-2$, $k=1,\dots,n-1$, satisfies all the conditions given in (3.1).

If $k=n$, then obviously, for $t=0$, $A_{0n}(x_n)=1$. On differentiating (3.10), by Leibnitz's theorem, $m=1,\dots,r-2$ times, we have by (2.9), for $x=x_j$, $j=1,\dots,n-1$

$$A_{0n}^{(m)}(x_j)=0,\quad j=1,\dots,n-1,\quad m=1,\dots,r-2.$$

For $x=x_n$, by (3.11), we have

$$A_{0n}^{(m)}(x_n)=\frac{\left(W_{n-1}^{r-1}(x)\right)_{x_n}^{(m)}}{W_{n-1}^{r-1}(x_n)}+$$

$$+\sum_{s=1}^{m}\binom{m}{s}\left(W_{n-1}^{r-1}(x)\right)_{x_n}^{(m-s)}\left(\sum_{i=1}^{s}\left(W_{n-1}(x)\right)_{x_n}^{(s-i)}i!e_{ik}^{*}\right)=0.$$

Also,

$$\left(w^{r-1}A_{0n}(x)\right)_{x_j}^{(r)}=0,\qquad j=1,\dots,n-1.$$

Now, for $t=1,\dots,r-2$, $A_{rn}^{(m)}(x_j)=0, j=1,\dots,n-1$, $m=0,\dots,r-2$. For $x=x_n$ and $m=t,\dots,r-2$, we have

$$A_{tn}^{(m)}(x_n)=$$

$$=\sum_{s=t}^{m}\binom{m}{s}\left(W_{n-1}^{r-1}(x)\right)_{x_n}^{(m-s)}\left(\sum_{i=t}^{s}\binom{s}{i}\left(W_{n-1}(x)\right)_{x_n}^{(s-i)}i!e_{ik}^{**}\right)=\delta_{mt}$$

by (3.13). Also, $\left(w^{r-1} A_m(x)\right)^{(r)}_{x_j} = 0$, $j = 1, \ldots, n-1$.

Thus $A_m(x)$, $t = 0, \ldots, r-2$, as given in (3.10) and (3.12) satisfies all the conditions of (3.1).

LEMMA 2. $A_{rk}(x)$, $k = 1, \ldots, n-1$, *has the form*

$$A_{rk}(x) = W_{n-1}^{r-1}(x) \left[\alpha_{rk} \int_{x_n}^{x} (t - x_n)^{r-2} W_{n-1}(t) dt + \right. \tag{3.16}$$
$$\left. + \beta_{rk} \int_{x_n}^{x} (t - x_n)^{r-2} l_k(t) dt \right],$$

where

$$\beta_{rk} = \frac{1}{r! w^{r-1}(x_k)^{r-2} W'_{n-1}(x_k)^{r-1}} \tag{3.17}$$

and

$$\alpha_{rk} = -\beta_{rk} \frac{l_k(x_n)}{W_{n-1}(x_n)}. \tag{3.18}$$

PROOF. Obviously, $A_{rk}^{(m)}(x_j) = 0$, $j = 1, \ldots, n$, $m = 0, \ldots, r-2$. On differentiating (3.16), $(r-1)$ times, by Leibnitz's theorem, we have at $x = x_n$

$$A_{rk}^{(r-1)}(x_n) = (r-2)!\, W_{n-1}^{r-1}(x) \left[\alpha_{rk} W_{n-1}(x_n) + \beta_{rk} l_k(x_n)\right] = 0$$

due to (3.18).

If the weight function $w(x) \in C^r(a,b)$ satisfies, the conditions (2.6) and (2.7), then by (2.9) and (3.17), we have

$$\left(w^{r-1} A_{rk}(x)\right)^{(r)}_{x_j} = r w^{r-1} W'_{n-1}(x_j)^{r-1} \left[\beta_{rk} (x_j - x_n)^{r-2} l_k(x_j)\right] = \delta_{jk}.$$

Hence $A_{rk}(x)$, $k = 1, \ldots, n-1$, given by (3.16), satisfies all the conditions given in (3.3).

LEMMA 3. $A_{(r-1)n}(x)$ *is given by*

$$A_{(r-1)n}(x) = \frac{W_{n-1}^{r-1}(x)}{(r-1)!\, W_{n-1}^{r}(x_n)} \int_{x_n}^{x} (t - x_n)^{r-2} W_{n-1}(t) dt. \tag{3.19}$$

PROOF. Obviously, $A^{(m)}_{(r-1)n}(x_j) = 0$, $j = 1, \ldots, n$, $m = 0, \ldots, r-2$ and $A^{(r-1)}_{(r-1)n}(x_n) = 1$. Since the weight function $w(x) \in C^r(a,b)$ satisfies the conditions (2.6) and (2.7), hence $\left(w^{r-1}A_{(r-1)n}(x)\right)^{(r)}_{x_j} = 0$. Thus $A_{(r-1)n}(x)$ given by (3.19), satisfies the conditions (3.2).

4. Proof of the Main Theorem

As the polynomials $A_{tk}(x)$, $t = 0, \ldots, r$, of degree $\le rn-1$ are basic interpolation polynomials, hence the modified weighted $(0, 1, \ldots, r-2, r)$-interpolation polynomial $S_n(x)$ of degree $\le rn-1$ in Theorem 1, can be explicitly represented in the form

$$S_n(x) = \sum_{k=1}^{n}\left(\sum_{t=0}^{r-2} y_k^{(t)}A_{tk}(x)\right) + \sum_{k=1}^{n-1} y_k^{(r)}A_{rk}(x) + y_n^{(r-1)}A_{(r-1)n}(x).$$

Indeed, by Lemmas 1, 2 and 3, the polynomial $S_n(x)$ satisfies the conditions (2.2), (2.3) and (2.4), hence the Theorem is proved.

5. Remarks

1. Now, we show that if the zeros of the polynomial $W_{n-1}(x)$ are the zeros of the classical orthogonal polynomials of degree $\le rn-1$, then the weight function $w(x) \in C^r(a,b)$ satisfying the conditions (2.6) and (2.7) always exist. It is known that the zeros of the classical orthogonal polynomials are real and simple.

In [2] J. BALÁZS gave such weight functions w for which

$$\text{(5.1)} \qquad w(x_k) \neq 0 \quad \text{and} \quad \left(w\,W_{n-1}\right)^{n}(x_k) = 0, \quad k = 1, \ldots, n-1,$$

where W_{n-1} is the Jacobi, Laguerre of Hermite polinomial. So, we have only to show that in the cases

$$\text{(5.2)} \qquad \left(w^{r-1}W_{n-1}^{r-1}(x)\right)^{(r)}_{x_k} = 0, \qquad k = 1, \ldots, n-1.$$

We prove it by induction.

For $r=1$, (5.2) is true, obviously. It is also true for $r=2$, by (5.1). Let (5.2) be true for $r=i$, i.e.,

$$\left(w^{i-1}W_{n-1}^{i-1}(x)\right)^{(i)}_{x_k}=0,\qquad k=1,\dots,n-1 \tag{5.3}$$

then we have to show that (5.2) holds for $r=i+1$.

$$\begin{aligned}\left(w^{i}W_{n-1}^{i}(x)\right)^{(i+1)}_{x_k}&=\left\{\left(w^{i-1}W_{n-1}^{i-1}(x)\right)\left(w\,W_{n-1}(x)\right)\right\}^{(i+1)}_{x_k}=\\&=\left(w^{i-1}W_{n-1}^{i-1}(x)\right)^{(i+1)}_{x_k}\left(w\,W_{n-1}(x)\right)_{x_k}+\\&\quad+(i+1)\left(w^{i-1}W_{n-1}^{i-1}(x)\right)^{(i)}_{x_k}\left(w\,W_{n-1}(x)\right)'_{x_k}+\\&\quad+\binom{i+1}{2}\left(w^{i-1}W_{n-1}^{i-1}(x)\right)^{(i-1)}_{x_k}\left(w\,W_{n-1}(x)\right)''_{s_k}+\ldots=0,\end{aligned}$$

due to (2.5), (5.3), (5.1) and (2.9). Hence if the nodes are the zeros of classical orthogonal polynomials then by Theorem 1 there exists a modified weighted $(0,1,\dots,r-2,r)$-interpolation polynomial $S_n(x)$ of degree $\le rn-1$ satisfying the conditions (2.2), (2.3) and (2.4).

2. (i) BALÁZS's result [1] is a particular case of ours for $r=2$.

(ii) Taking the special weight functions $w(x)=\exp\left(\frac{-x^2}{2}\right)$ and $w(x)=$ $=\exp(-x^2)$ and the nodes as the zeros of $H_n(x)$, n^{th} Hermite polynomial, the problem reduces to DATTA and MATHUR's problems [2], [4] for $r=2$ and $r=3$ respectively.

3. The convergence problem will be dealt in another paper.

References

[1] BALÁZS, J., Súlyozott $(0,2)$-interpoláció ultraszférikus polinomok gyökein, *MTA III. Oszt. közl.*, **11** (1961), 305–338.

[2] BALÁZS, J., Modified weighted $(0,2)$-interpolation, *Approx. Theory,* Marcel Dekker Inc., New York, 1998, 61–73.

[3] S. DATTA and P. MATHUR, On Weighted $(0,2)$-Interpolation on Infinite Interval $(+\infty,-\infty)$, *Annales Univ. Sci. Budapest., Sect. Math.,* **42** (1999), 45–57.

[4] S. DATTA and P. MATHUR, On weighted $(0,1,3)$-interpolation on infinite interval $(-\infty,+\infty)$, communicated for publication.

[5] JOÓ, I., On weighted $(0, 2)$-interpolation, *Annales Univ. Sci. Budapest., Sect. Math.,* **38** (1995), 185–222.

[6] SZILI, L., Weighted $(0, 2)$-interpolation on the roots of Hermite polynomials, *Annales Univ. Sci. Budapest., Sect. Math.,* **27** (1984), 152–166.

ANNALES UNIV. SCI. BUDAPEST., **44** (2001), 65–81

A NOTE ON WEIGHTED $(0,1,3)$-INTERPOLATION ON INFINITE INTERVAL $(-\infty,+\infty)$

By

S. DATTA and P. MATHUR

Department of Mathematics and Astronomy, Lucknow University, Lucknow, India

(Received September 15, 2000)

1. Introduction

J. BALÁZS on the suggestion of P. TURÁN initiated the study of weighted $(0,2)$-interpolation, which means the determination of a polynomial $G_n(x)$ of degree $\leq 2n-1$ such that

$$(1.1)\qquad G_n(\xi_{i,n})=a_{i,n},\quad (wG_n)''(\xi_{i,n})=b_{i,n};\qquad i=1,2\dots,n$$

where $a_{i,n}$, $b_{i,n}$ are arbitrary given numbers and $\xi_{i,n}$ are the zeros of the n^{th}-ultraspherical polynomial $P_n^{(\alpha)}(x)$ $(\alpha>-1)$ with weight function $w(x)=$ $=(1-x^2)^{(1+\alpha)/2}$, $x\in[-1,1]$. He proved that generally there do not exist any polynomial of degree $\leq 2n-1$ satisfying the conditions (1.1). However, taking n even, he proved the existence, uniqueness, explicit representation and convergence theorem for the polynomial $G_n(x)$ of degree $\leq 2n$ satisfying (1.1) together with

$$(1.2)\qquad G_n(0)=\sum_{i=1}^{n}a_{i,n}l_{i,n}^2(0).$$

If n is odd, the uniqueness is not true. L. SZILI [9] studied an analogous problem on the nodes as the zeros of the n^{th} Hermite polynomial $H_n(x)$ with weight function $w(x)=e^{-x^2/2}$. Later, I. JOÓ [5] sharpened these results. In an earlier paper [3] authors have improved I. Joó's result by replacing the condition (1.2) with an interpolatory condition $G_n(0)=y_0$, where y_0 is an arbitrary given number in the case of n even and obtained that the necessary and sufficient condition for the existence of weighted $(0,2)$-interpolation in the case of n odd is $G_n(0)=y_0'$.

K. K. MATHUR and R. B. SAXENA [7] extended the study of weighted $(0,2)$-interpolation to the case of weighted $(0,1,3)$-interpolation, which means to determine a polynomial $T_n(x)$ of degree $\leq 3n-1$ such that

$$T_n(x_{i,n}) = y_{i,n}, \quad T_n'(x_{i,n}) = y_{i,n}', \quad (wT_n)'''(x_{i,n}) = y_{i,n}'''; \quad i = 1, \dots, n, \tag{1.3}$$

where $y_{i,n}$, $y_{i,n}'$, $y_{i,n}'''$ are arbitrary given numbers and weight function $w(x) = e^{-x^2}$ $(x \in \mathbf{R})$ and $x_{i,n}$'s are the zeros of Hermite polynomial $H_n(x)$ given by:

$$-\infty < x_{n,n} < \dots < x_{1,n} < \infty \qquad (n \in \mathbf{N}). \tag{1.4}$$

They proved that generally no polynomial of degree $\leq 3n-1$ satisfying the conditions (1.3) exists, as such taking an additional condition:

$$R_n(0) = \sum_{i=1}^{n} \left[\left(1 + 3x_{i,n}^2\right) y_{i,n} - x_{i,n} y_{i,n}'\right] l_{i,n}^3(0), \tag{1.5}$$

where 0 is not a nodal point belonging to (1.4). They showed that there do exists a unique polynomial of degree $\leq 3n$ (n even) and established its explicit representation and convergence. If n is odd, uniqueness fails to hold.

The object of this paper is to get an analogous result by replacing the artificial looking, condition (1.5) by an interpolatory condition $R_n(0) = y_0$, where y_0 is an arbitrary given number, in the case of n even. Further, what will be the necessary and sufficient condition for the existence and uniqueness of the $(0,1,3)$-interpolation in the case of n odd? If it exists what will be its explicit form and does it converge?

Here, we answer these questions in affirmative taking the nodes as the zeros of n^{th} Hermite polynomial $H_n(x)$.

In section 2, we have given preliminaries. New results have been stated in section 3. Sections 4 and 5 are devoted to the basic estimates of the fundamental polynomials and the proof of our main theorem for n odd and n even respectively.

2. Preliminaries

Let H_n be the n^{th} Hermite polynomial with usual normalisation

$$\int_{-\infty}^{+\infty} H_n(t) H_m(t) e^{-t^2} dt = \pi^{1/2} 2^n n! \delta_{n,m} \quad n, m \in \mathbf{N} \tag{2.1}$$

which satisfies the differential equation:

$$\begin{aligned} &H_n''(x) - 2x H_n'(x) + 2n H_n(x) = 0, \\ &H_n'(x) = 2n H_{n-1}(x). \end{aligned} \tag{2.2}$$

It is well known that $x_{i,n}$ (the roots of $H_n(x)$) satisfy the following relations:

$$\begin{cases} -\infty < x_{n,n} < \ldots < x_{\frac{n+1}{2}} < 0 < x_{\frac{n}{2},n} < \ldots x_{1,n} < \infty & (n = 2m), \\ -\infty < x_{n,n} < \ldots < x_{\frac{n+1}{2},n} = 0 < \ldots < x_{1,n} < \infty & (n = 2m+1), \\ x_{i,n} = -x_{n-i+1,n} \qquad \left(i = 1, 2, \ldots, \left[\frac{n}{2}\right]\right), \end{cases} \tag{2.3}$$

$$\begin{cases} H_n''(x_{i,n}) = 2x_{i,n} H_n'(x_{i,n}), \\ H_n'''(x_{i,n}) = 2[2x_{i,n}^2 - (n-1)] H_n'(x_{i,n}), \\ H_n^{(4)}(x_{i,n}) = 4x_{i,n}(3 - 2n + 2x_{i,n}^2) H_n'(x_{i,n}). \end{cases} \tag{2.4}$$

Let $l_{i,n}$ denote the fundamental polynomial of Lagrange interpolation corresponding to the nodal point $x_{i,n}$. Then

$$l_{i,n}(x) = \frac{H_n(x)}{(x - x_{i,n}) H_n'(x_{i,n})} \qquad (i = 1, \ldots, n), \tag{2.5}$$

$$l_{i,n}(x_{j,n}) = \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{for } i = j \end{cases}, \tag{2.6}$$

$$l_{i,n}'(x_{j,n}) = \begin{cases} \frac{H_n'(x_{j,n})}{(x_{j,n} - x_{i,n}) H_n'(x_{i,n})} & \text{for } j \neq i \\ x_{i,n} & \text{for } j = i \end{cases} \tag{2.7}$$

$$l_{i,n}''(x_{j,n}) = \begin{cases} \frac{2H_n'(x_{j,n}}{(x_{j,n} - x_{i,n}) H_n'(x_{i,n})} \left[x_{j,n} - \frac{1}{x_{j,n} - x_{i,n}}\right] & j \neq i \\ \frac{4x_{i,n}^2 - 2(n-1)}{3} & j = i \end{cases} \tag{2.8}$$

and

$$l_{i,n}'''(x_{j,n}) = \begin{cases} \frac{1}{x_{j,n} - x_{i,n}} \left[\frac{H_n'''(x_{j,n})}{H_n'(x_{i,n})} - 3l_{i,n}''(x_{j,n})\right] & j \neq i \\ x_{i,n}(2x_{i,n}^2 + 3 - 2n) & j = i. \end{cases} \tag{2.9}$$

For the roots of $H_n(x)$, we have

$$x_{i,n}^2 \sim \frac{i^2}{n}, \qquad i = 1, \ldots, n. \tag{2.10}$$

$$(2.11)\qquad H_n(x) = O(1)n^{-1/4}\sqrt{2^n n!}\left(1+3\sqrt{|x|}\right)e^{x^2/2}, \quad x \in \mathbf{R}.$$

$$(2.12)\qquad \sum_{i=1}^{n} \frac{e^{\delta x_{i,n}^2}}{H_n'(x_{i,n})^2} = O\left(\frac{1}{2^{n+1}n!}\right), \quad 0 < \delta < 1,$$

$$(2.13)\qquad \sum_{i=1}^{n} e^{x_{i,n}^2} l_{i,n}^2(x) = O(e^{x^2}),$$

$$(2.14)\qquad |H_n(0)| = \frac{n!}{\left(\frac{n}{2}\right)!} \quad \text{for } n \text{ even},$$

$$(2.15)\qquad \frac{2^n\left(\left(\frac{n}{2}\right)!\right)^2}{(n+1)!} \sim n^{-1/2}.$$

Also

$$(2.16)\qquad \int_0^x \frac{x_{i,n}l_{i,n}(t) - l_{i,n}'(t)}{t - x_{i,n}}dt = \frac{1}{2}(l_{i,n}'(x) - l_{i,n}'(0)) - x\,l_{i,n}(x) + n\int_0^x l_{i,n}(t)dt.$$

If $\mu_{i,n} = \frac{1}{3}(x_{i,n}^2 - 2(n-1))$, then

$$(2.17)\qquad \int_0^x \frac{(\mu_{i,n}(t - x_{i,n}) + x_{i,n})l_{i,n}(t) - l_{i,n}'(t)}{(t - x_{i,n})^2}dt =$$

$$= \frac{nx_{i,n}}{3}\int_0^x l_{i,n}(t)dt - \frac{1}{6}(l_{i,n}''(x) - l_{i,n}''(0)) +$$

$$+ \frac{x_{i,n}}{2}(l_{i,n}'(x) - l_{i,n}'(0)) - \frac{1}{3}(x_{i,n}^2 + n + 2)(l_{i,n}(x) - l_{i,n}(0)) +$$

$$+ \frac{1}{3H_n'(x_{i,n})}(H_n'(x) - H_n'(0)) - \frac{x_{i,n}}{3H_n'(x_{i,n})}(H_n(x) - H_n(0)),$$

3. New Results

Let n be odd in Theorems 1, 2, 3.

THEOREM 1. *Let the nodal points are the roots of n^{th} Hermite polynomial $H_n(x)$ and the weight function is $w(x) = e^{-x^2}$ $(x \in \mathbf{R})$.*

Then there exists a unique polynomial $R_n(x)$ *of degree* $\leq 3n$ *satisfying the conditions* (1.2) *and* $R_n''(0) = y_0''$, *where* y_0'' *is an arbitrary given number, and,* 0 *is a nodal point.*

THEOREM 2. *For* $k = 1, \ldots, n$,

(3.1) $A_k(x) = l_k^3(x) - 3x_k B_k(x) +$
$$\frac{H_n^2(x)}{H_n'(x_k)^2}\left[\int_0^x \frac{(x_k + \lambda_k(t - x_k) + \mu_k(t - x_k)^2)l_k(t) - l_k'(t)}{(t - x_k)^2}dt\right],$$

where

$$\lambda_k = \frac{x_k^2 - 2(n-1)}{3} \quad \text{and} \quad \mu_k = \frac{x_k(2x_k^2 - 3n)}{3}; \tag{3.2}$$

(3.3)
$$B_k(x) = (x - x_k)l_k^3(x) + \frac{H_n^2(x)}{H_n'(x_k)^2}\left[\int_0^x \frac{(x_k + S_k(t - x_k))l_k(t) - l_k'(t)}{(t - x_k)}\right]dt,$$

where

$$S_k = \frac{1}{3}(n + 2(1 - x_k^2)),$$

$$C_k(x) = \frac{e^{x_k^2}H_n^2(x)}{6H_n'(x_k)^2}\int_0^x l_k(t)dt; \tag{3.4}$$

and

$$D_0(x) = \frac{H_n^2(x)}{2H_n'(0)^2}. \tag{3.5}$$

Then

$$R_n(x) = \sum_{k=1}^n y_k A_k(x) + \sum_{k=1}^n y_k' B_k(x) + \sum_{k=1}^n y''' C_k(x) + y_0'' D_0(x) \tag{3.6}$$

is a uniquely determined polynomial of degree $\leq 3n$ *satisfying the conditions of Theorem* 1.

THEOREM 3. *Let $f : \mathbf{R} \to \mathbf{R}$ be a twice continuously differentiable function, such that*

$$(3.7)\qquad \begin{aligned} &\lim_{|x|\to\infty} x^{2r} f(x) e^{-x^2} = 0; \qquad r = 0, 1, \dots, \\ &\lim_{|x|\to\infty} e^{-x^2} f^{(r)}(x) = 0; \qquad r = 1, 2. \end{aligned}$$

Then the weighted $(0, 1, 3)$*-interpolatory polynomials* $R_n(x)$ *(*$n = 3, 5, 7, \dots$*) given by* (3.6) *together with*

$$(3.8)\qquad \begin{aligned} &y_k = f(x_k), \quad y_k' = f'(x_k) \\ &y_k''' = O\left(e^{x_k^2/2} n\omega\left(f', \frac{1}{\sqrt{n}}\right)\right) \\ &\text{and} \\ &y_0'' = f''(0) \end{aligned}$$

satisfy the estimate:

$$(3.9)\qquad e^{-x^2}|f(x) - R_n(f, x)| = O(1)\omega\left(f', \frac{1}{\sqrt{n}}\right),$$

where O does not depend on n and x. Here $\omega(f', \cdot)$ denotes the Freud's modulus of continuity of f'.

In the case of n even, analogous to Theorem 1, there do exist a weighted $(0, 1, 3)$-interpolatory polynomial $R_n^*(x)$ of degree $\leq 3n$ satisfying the conditions (1.2) and $R_n^*(0) = y_0^*$, where y_0^* is an arbitrary given number. Further, let

(3.10)

$$A_k^*(x) = \begin{cases} \dfrac{H_n^2(x)}{H_n^2(0)} & \text{for } k = 0 \ (x_0 = 0) \\ l_k^4(x) - 3x_k B_k^*(x) - \dfrac{H_n^2(x)}{x_k^2 H_n'(x_k)^2} l_k(0) + & \\ + \dfrac{H_n^2(x)}{H_n'(x_k)^2} \displaystyle\int_0^x \frac{(x_k + \lambda_k^*(t - x_k) + \mu_k^*(t - x_k)^2) l_k(t) - l_k^*(t)}{(t - x_k)^2} dt & \text{for } k = 1, \dots, n, \end{cases}$$

where

$$(3.11)\quad \lambda_k^* = \frac{x_k^2 - 2(n-1)}{3} \quad \text{and} \quad \mu_k^* = -x_k\left(n - \frac{2}{3}x_k^2\right), \quad k = 1, \dots, n,$$

$$B_k^*(x) = (x - x_k)l_k^3(x) + \frac{H_n^2(x)}{x_k H_n'(x_k)^2} l_k(0) +$$

$$(3.12) \qquad + \frac{H_n^2(x)}{H_n'(x_k)^2} \int_0^x \frac{(x_k + S_k^*(t - x_k))l_k(t) - l_k'(t)}{(t - x_k)} dt; \quad k = 1, \dots, n,$$

where

$$(3.13) \qquad S_k^* = \frac{1}{3}\left(n + 2(1 - x_k^2)\right)$$

and

$$(3.14) \qquad C_k^* = \frac{e^{x_k^2} H_n^2(x)}{H_n'(x_k)^2} \int_0^x l_k(t)dt, \quad k = 1, \dots, n.$$

Then

$$(3.15) \qquad R_n^*(x) = \sum_{k=0}^n y_k^* A_k^*(x) + \sum_{k=1}^n y'^* B_k^*(x) + \sum_{k=1}^n y'''^* C_k^*(x)$$

is a uniquely determined polynomial of degree $\le 3n$ satisfying the conditions (1.2) and $R_n(0) = y_0^*$.

THEOREM 4. *Let* $f : \mathbf{R} \to \mathbf{R}$ *be a twice continuously differentiable function satisfying the requirements* (3.7) *and the numbers* y_k^*, $y_k^{*\prime}$ *and* $y_k^{*\prime\prime\prime}$ *are such that*

$$(3.16) \qquad \begin{aligned} & y_k^* = f(x_k) \qquad k = 0, \dots, n \\ & y_k^{*\prime} = f'(x_k) \qquad k = 1, \dots, n \\ & \textit{and} \\ & y_k^{*\prime\prime\prime} = O\left(e^{x_k^2/2} n\omega\left(f', \frac{1}{\sqrt{n}}\right)\right). \end{aligned}$$

Then for interpolatory polynomial $R_n^*(x)$ *(*$n = 2, 4, 6, \dots$*) given by* (3.15)*, we have the estimate:*

$$(3.17) \qquad e^{-x^2}|f(x) - R_n^*(f, x)| = O(1)\omega\left(f', \frac{1}{\sqrt{n}}\right), \qquad x \in \mathbf{R},$$

where O *does not depend on* n *and* x*, and* $\omega\left(f', \frac{1}{\sqrt{n}}\right)$ *is the Freud's modulus of continuity.*

We shall prove only our main Theorem 3 and 4, as the proofs of the other theorems are quite similar to that of theorems in [1].

4. Basic Estimates with Respect to the Fundamental Polynomials (n odd)

LEMMA 1. ([6], Lemma 1.) *If λ_i $(i = 1,\dots,n)$ are the Christoffel-numbers on Hermite nodes, then*

$$\lambda_i \sim e^{-x_i^2} \frac{1}{n^{1/6} \cdot i^{1/3}} \sim e^{-x_i^2} \Phi_n(x_i), \quad i 1,\dots,\frac{n}{2},$$

where $\Phi(x_i) = x_i - x_{i+1}$ (certainly $x_1 > x_2 > \dots > x_n$), $\lambda_i = \lambda_{n-i+1}$.

LEMMA 2. *Let n be odd, then*

$$|D_0(x)| = O(e^{x^2}), \quad x \in \mathbf{R}, \tag{4.1}$$

$$\sup_{x \in \mathbf{R}} e^{-\frac{3x^2}{2}} \sum_{i=1}^{n} e^{x_i^2/2} |C_i(x)| \approx \frac{1}{n}. \tag{4.2}$$

PROOF. By using (2.2) and (2.11) in (3.5), (4,1) follows.

Without loss of generality, we can assume that $x \geq 0$. Using Lemma 1, we get

$$\begin{aligned} |H_n'(x_i)| &= 2n|H_{n-1}(x_i)| = 2n \cdot \pi^{1/4} \sqrt{2^{n-1}(n-1)!}|h_{n-1}(x_i)| = \\ &= \pi^{1.4} \sqrt{2} \sqrt{2^n n!} \lambda_i^{-1/2} \asymp c e^{x_i^2/2} \sqrt{2^n n!} \Phi_n^{-1/2}(x_i); \; i = 1,\dots,n. \end{aligned} \tag{4.3}$$

We have by (3.4)

$$|C_i(x)| \asymp e^{x_i^2/2} \frac{|H_n(x)|}{|H_n'(x_k)|} |\overline{B}_i(x)|, \tag{4.4}$$

where

$$\overline{B}_i(x) = \frac{e^{x_i^2/2} H_n(x)}{H_n'(x_i)} \int_0^x l_i(t)dt \tag{4.5}$$

and from [6]

$$\sum_{i=1}^{n} e^{x_i^2/2} |\overline{B}_i(x)| = O\left(\frac{e^{x^2/2}}{\sqrt{n}}\right). \tag{4.6}$$

We remark that (4.6) also holds for n odd. Thus

$$\sum_{i=1}^{n} e^{x_i^2/2}|C_i(x)| = \sum_{i=1}^{n} \frac{e^{x_i^2/2}|H_n(x)|}{|H_n'(x_i)} e^{x_i^2/2}|\overline{B}_i(x)| =$$

$$= \sum_{|x_i|\le\sqrt{n}} \ldots + \sum_{|x_i|>\sqrt{n}} \ldots =$$

$$= O(1)\frac{e^{x^2}}{\sqrt{n}} \sum_{|x_i|\le\sqrt{n}} e^{x_i^2/2}|\overline{B}_i(x)| + O(1)\frac{e^{x^2}}{n^{1/4}} \sum_{|x_i|>\sqrt{n}} \Phi_n^{\frac{1}{2}}(x_i)e^{\frac{x_i^2}{2}}|\overline{B}_i(x)| =$$

$$= O(1)\frac{e^{3x^2/2}}{n},$$

where we used $|H_n'(x_i)| \ge e^{x_i^2/2}\sqrt{2^n n!}n^{1/4}$, $|x_i| \le \sqrt{n}$, given in [5].

LEMMA 3. *For n odd*

$$\sum_{i=1}^{n} e^{x_i^2}|B_i(x)| = O\left(e^{3x^2/2}\right), \qquad x \in \mathbf{R}. \tag{4.7}$$

PROOF. From (3.3), due to (2.16), we have

$$\begin{aligned} B_i(x) = \frac{H_n(x)}{H_n'(x_i)} \Bigg[& \overline{A}_i(x) - \frac{7n}{3}\frac{H_n(x)}{H_n'(x_i)} \int_0^x l_i(t)dt - \\ & -\frac{5}{3}(1-x_i^2)\frac{H_n(x)}{H_n'(x_i)} \int_0^x l_i(t)dt \Bigg] \end{aligned} \tag{4.8}$$

where

$$\begin{aligned} \overline{A}_i(x) = {} & \frac{l_i^2(x)}{2} + (1-x_i^2)\frac{H_n(x)}{H_n'(x_i)} \int_0^x l_i(t)dt + \frac{nH_n(x)}{H_n'(x_i)} \int_0^x l_i(t)dt + \\ & + \frac{H_n'(x)}{2H_n'(x_i)} l_i(x) - x\frac{H_n(x)}{H_n'(x_i)} l_i(x) - \frac{H_n(x)}{2H_n'(x_i)} l_i'(0) \quad \text{for } x_i \ne 0 \end{aligned} \tag{4.9}$$

and

$$\overline{A}_i(x) = \frac{l_i^2(x)}{2} + (1 - x_i^2)\frac{H_n(x)}{H_n'(x_i)}\int_0^x l_i(t)dt + \frac{nH_n(x)}{H_n'(x_i)}\int_0^x l_i(t)dt + \tag{4.10}$$
$$+ \frac{H_n'(x)}{2H_n'(x_i)}l_i(x) - x\frac{H_n(x)}{H_n'(x_i)}l_i(x) \quad \text{for } x_i = 0.$$

By [[6], lemma 4], we have

$$\sum_{i=1}^{n} e^{x_i^2/2}|\overline{A}_i(x)| = O(1)e^{x^2}\sqrt{n}, \tag{4.11}$$

$$n\sum_{i=1}^{n} \frac{|H_n(x)|e^{x_i^2/2}}{|H_n'(x_i)|}\left|\int_0^x l_i(t)dt\right| = O(1)e^{x^2}\sqrt{n} \tag{4.12}$$

and

$$\sum_{i=1}^{n} \frac{|(1 - x_i^2)||H_n(x)|e^{x_i^2/2}}{|H_n'(x_i)|}\left|\int_0^x l_i(t)dt\right| = O(1)e^{x^2}\sqrt{n}. \tag{4.13}$$

(4.11)–(4.13) also hold for n odd. Thus by (4.8), (4.7) follows.

LEMMA 4. *For n odd*

$$\sum_{i=1}^{n} e^{x_i^2}|A_i(x)| = O(1)e^{3x^2/2}\sqrt{n}, \qquad x \in \mathbf{R}. \tag{4.14}$$

PROOF. $A_i(x)$, given by (3.2), can be represented alternatively as follows. For $x_i \neq 0$

$$A_i(x) = l_i^3(x) - 3x_iB_i(x) + \frac{H_n^2(x)}{H_n'(x_i)^2}I, \tag{4.15}$$

where

$$I = \frac{2x_i}{3}(2n - x_i^2)\int_0^x l_i(t)dt - \frac{1}{6}\left(l_i''(x) - l_i''(0) + \frac{x_i}{2}(l_i'(x) - l_i'(0))\right) - \tag{4.16}$$
$$- \frac{1}{3}(x_i^2 + n + 2)l_i(x) + \frac{1}{3H_n'(x_i)}(H_n'(x) - H_n'(0)) - \frac{x_i}{3H_n'(x_i)}H_n(x)$$

and for $x_i = 0$

$$A_i(x) = l_i^3(x) - 3x_i B_i(x) + \frac{H_n^2(x)}{H_n'(x_i)^2} J, \tag{4.17}$$

where

$$\begin{aligned} J = {} & \frac{2x_i}{3}(2n - x_i^2) \int_0^x l_i(t)dt - \frac{1}{6} l_i''(x) + \frac{x_i}{2} l_i'(x) - \\ & - \frac{1}{3}(x_i^2 + n + 2) l_i(x) + \frac{1}{3H_n'(x_i)}(H_n'(x) - H_n'(0)) - \frac{x_i H_n(x)}{3H_n'(x_i)}. \end{aligned} \tag{4.18}$$

Hence

$$\sum_{i=1}^{n} e^{\frac{x_i^2}{2}} |A_i(x)| = I_1 + I_2 + I_3 + I_4 + I_5 + I_6 + I_7 + I_8 + I_9. \tag{4.19}$$

where

$$I_1 = \sum_{i=1}^{n} e^{x_i^2} l_i^2(x) |l_i(x)| = O\left(\sqrt{n} e^{3x^2/2}\right). \tag{4.20}$$

$$I_2 = 3 \sum_{i=1}^{n} |x_i| + e^{x_i^2} |B_i(x)| = O\left(\sqrt{n} e^{3x^2/2}\right). \tag{4.21}$$

Using ($|x_i| = O(\sqrt{n})$ and Lemma 2

$$\begin{aligned} I_3 &= \frac{2}{3} \sum_{i=1}^{n} |x_i| |(2n - x_i^2)| e^{x_i^2} \frac{H_n^2(x)}{H_n'(x_i)^2} \left| \int_0^x l_i(t)dt \right| = \\ &= O\left(n^{3/2}\right) \sum_{i=1}^{n} |C_i(x)| = O\left(\sqrt{n} e^{3x^2/2}\right), \end{aligned} \tag{4.22}$$

$$I_4 = \frac{1}{6} \sum_{i=1}^{n} e^{x_i^2} \frac{H_n^2(x)}{H_n'(x_i)^2} |l_i''(x)|.$$

From (2.6), we have

(4.23)

$$I_4 = \frac{1}{6}\sum_{i=1}^{n} e^{x_i^2}\frac{H_n^2(x)}{H_n'(x_i)^2}\left[\left|\frac{(x(x-x_i)-1)H_n'(x)}{(x-x_i)^2 H_n'(x_i)} + \frac{l_i(x)}{(x-x_i)^2} - n l_i(x)\right|\right] \leq$$

$$\leq \frac{1}{6}\sum_{i=1}^{n} e^{x_i^2}\left[|x(x-x_i)-1|\left|\frac{H_n^2(x)}{H_n'(x_i)^2}\right| l_i^2(x) + (1+n(x-x_i)^2)|l_i^3(x)|\right] =$$

$$= \sum_{i=1}^{n}|x|e^{x_i^2}\frac{|H_n^2(x)H_n'(x)|}{H_n'(x_i)^2}|l_i(x)| + O(n)\sum_{i=1}^{n} e^{x_i^2}\frac{H_n^2(x)}{H_n'(x_i)^2}|l_i(x)| =$$

$$= |x|\frac{2n|H_{n-1}(x)H_n(x)|}{2^n n!}\sum_{i=1}^{n}\Phi(x_i)|l_i(x)| + O(n)\frac{H_n^2(x)}{2^n n!}\sum_{i=1}^{n}\Phi(x_i)|l_i(x)|.$$

If $|x| \geq 2\sqrt{n}$, then $\frac{|H_n(x)|e^{-x^2/2}}{\sqrt{n}n!} = O(1)e^{-cn}$, therefore, we can assume that $x \leq 2\sqrt{n}$. Hence

$$I_4 = O(1)e^{x^2}\sqrt{n}\sum_{i=1}^{n}\Phi_n(x_i)|l_i(x)| =$$

$$= O\left(\sqrt{n}e^{x^2}\right)\left[\sum_{|x_i|>2\sqrt{\log n}}\dots + \sum_{|x_i|\leq 2\sqrt{\log n}}\dots\right]. \tag{4.24}$$

Here

$$\sum_{|x_i|>2\sqrt{\log n}}\Phi_n(x_i)|l_i(x)| \leq c\frac{1}{n^{1/6}}\sum_{|x_i|>2\sqrt{\log n}}|l_i(x)| \leq c\frac{e^{x^2/2}}{n^{1/6}} \tag{4.25}$$

and

$$\sum_{|x_i|\leq 2\sqrt{\log n}}\Phi_n(x_i)|l_i(x)| \asymp \frac{1}{\sqrt{n}}\sum_{|x_i|\leq 2\sqrt{\log n}}|l_i(x)| \leq$$

$$\leq \frac{1}{\sqrt{n}}\sum_{i=1}^{n}|l_i(x)| \leq \frac{c}{\sqrt{n}}\left(\log n + e^{x^2/2}\right). \tag{4.26}$$

where we used [5], Hilfssatz 4, Satz, 11. Thus by (4.24), (4.25) and (4.26), we have

$$I_4 = O(1)\left(e^{3x^2/2} n^{-\frac{1}{6}}\right). \tag{4.27}$$

$$\begin{aligned} I_5 &= \frac{1}{2}\sum_{i=1}^{n} e^{x_i^2} \frac{H_n^2(x)}{H_n'(x_i)^2} |x_i| \left| \frac{H_n'(x)}{(x-x_i)H_n'(x_i)} - \frac{l_i(x)}{x-x_i} \right| \leq \\ &\leq \frac{1}{2}\sum_{i=1}^{n} e^{x_i^2} |x_i| \frac{|H_n(x)H_n'(x)|}{H_n'(x_i)^2} |l_i(x)| + \frac{1}{2}\sum_{i=1}^{n} |x_i| e^{x_i^2} \frac{|H_n(x)|}{|H_n'(x_i)|} l_i^2(x) = \\ &= O(1)\left(e^{3x^2/2}\sqrt{n}\right). \end{aligned} \tag{4.28}$$

$$\begin{aligned} I_6 &\leq \frac{1}{3}\sum_{i=1}^{n} e^{x_i^2} \frac{H_n^2(x)}{H_n'(x_i)^2} (n+2+x_i^2)|l_i(x)| = \\ &= O(n)\sum_{i=1}^{n} e^{x_i^2} \frac{H_n^2(x)}{H_n'(x_i)^2} |l_i(x)| = O\left(e^{3x^2/2}\sqrt{n}\right). \end{aligned} \tag{4.29}$$

(4.30)

$$\begin{aligned} I_7 &= \frac{1}{6}\sum_{i=1}^{n} e^{x_i^2} \frac{H_n^2(x)}{H_n'(x_i)^2} \left[\frac{2|H_n'(x)| + |H_n'(0)|}{|H_n'(x_i)|} \right] = \\ &= O(n)\sum_{i=1}^{n} e^{x_i^2} \frac{H_n^2(x)|H_{n-1}(x)|}{H_n'(x_i)^2 |H_n'(x_i)|} = O(n) H_n^2(x)|H_{n-1}(x)| \sum_{i=1}^{n} \frac{e^{x_i^2}}{|H_n'(x_i)^3|} = \\ &= O(n) H_n^2(x)|H_{n-1}(x) \left[\sum_{|x_i|\leq\sqrt{n}} + \sum_{|x_i|>\sqrt{n}} \right] = \\ &= O\left(\frac{e^{3x^2/2}}{n^{1/4}}\right) \frac{(2^n n!)^{3/2}}{(2^n n!)^{3/2}} \left[\sum_{|x_i|\leq\sqrt{n}} e^{\frac{n^{-3/4}}{x_i^2/2}} + e^{\frac{1}{n/2}} \sum_{|x_i|>\sqrt{n}} \Phi_n^{3/2}(x_i) \right] = \\ &= O\left(\frac{e^{3x^2/2}}{\sqrt{n}}\right). \end{aligned}$$

$$
\begin{aligned}
I_8 &= \frac{1}{3}\sum_{i=1}^{n} |x_i| e^{x_i^2} \frac{H_n^3(x)}{H_n'(x_i)^3} = \\
&= O(\sqrt{n})|H_n^3(x)\sum_{i=1}^{n} e^{x_i^2}\frac{1}{|H_n'(x_i)^3|} = O\left(\frac{e^{3x^2/2}}{\sqrt{n}}\right).
\end{aligned} \tag{4.31}
$$

(4.32)

$$
\begin{aligned}
I_9 &= \frac{1}{6}\sum_{i=1}^{n} e^{x_i^2}|2-3x_i^2|\frac{H_n^2(x)}{|x_i|H_n'(x_i)^2}|l_i'(0)| \le \\
&\le \frac{1}{3}\sum_{i=1}^{n} e^{x_i^2}|1-x_i^2|\frac{H_n^2(x)|H_n'(0)|}{x_i^2|H_n'(x_i)^3|} + \frac{1}{6}\sum_{i=1}^{n} e^{x_i^2}|x_i|\frac{H_n^2(x)|H_n'(0)|}{H_n'(x_i)^3|} = \\
&= O(n)H_n^2(x)|H_{n-1}(0)|\sum_{i=1}^{n}\frac{e^{3x_i^2/2}}{|H_n'(x_i)^3|} + \\
&\quad + O(n^{3/2})H_n^2(x)|H_{n-1}(0)|\sum_{i=1}^{n} e^{x_i^2}\frac{1}{|H_n'(x_i)^3|} = \\
&= O\left(\frac{1}{n^2}\right) + O\left(\sqrt{n}e^{x^2}\right) \le C\sqrt{n}e^{3x^2/2}.
\end{aligned}
$$

Using (4.19)–(4.32), we get the proof of (4.14).

LEMMA 5 ([6], Lemma 5]. *If $f \in C^1(\mathbf{R})$,*

$$\lim_{x\mapsto\pm\infty} x^{2r}f(x)w(x) = 0, \quad r = 0,1,\ldots$$

and

$$\lim_{x\to\pm\infty} f'(x)w(x) = 0,$$

then there exists a polynomial $p_n(x)$ of degree $\le n$, such that for $x \in \mathbf{R}$,

$$
\begin{aligned}
w(x)|f(x) - p_n(x)| &= O(1)\frac{1}{\sqrt{n}}\omega\left(f', \frac{1}{\sqrt{n}}\right) \\
w(x)|f'(x) - p_n'(x)| &= O(1)\frac{1}{\sqrt{n}}\omega\left(f', \frac{1}{\sqrt{n}}\right).
\end{aligned}
$$

Further ([9], Lemma 4*), we have for $x \in \mathbf{R}$*

$$
\begin{aligned}
w(x)|p_n(x)| &= O(1) \\
w(x)|p_n'(x)| &= O(1)
\end{aligned}
$$

and for $|x| < \sqrt{2n+1}$

$$w(x)|p_n''(x)| = O(1)\sqrt{n}\omega\left(f', \frac{1}{\sqrt{n}}\right).$$

Also ([7], Lemma 4*), we have for* $|x| < \sqrt{2n+1}$

$$w(x)|p_n'''(x)| = O(n)\omega\left(f', \frac{1}{\sqrt{n}}\right).$$

PROOF OF THEOREM 3. Let n be odd. From the uniqueness of polynomial $R_n(x)$ in (3.6), it follows that every polynomial $S_n(x)$ of degree $\leq 3n$ satisfies the relation:

$$S_n(x) = \sum_{i=1}^{n} S_n(x_i)A_i(x) + \sum_{i=1}^{n} S_n'(x_i)B_i(x) +$$
$$+ \sum_{i=1}^{n}(wS_n)'''(x_i)C_i(x) + S_n''(0)D_0(x).$$

Let $p_n(x)$ be a polynomial of degree $\leq 3n$ satisfying Lemma 5, then we have

$$e^{-3x^2/2}|f(x) - R_n(x)| < e^{-3x^2/2}|f(x) - p_n(x)| + e^{-3x^2/2}|p_n(x) - R_n(x)| =$$
$$= O(1)\left[e^{-x^2}e^{-x^2/2}|f(x) - p_n(x)| + e^{-3x^2/2}\left|\sum_{i=1}^{n}(f(x_i) - p_n(x_i))A_i(x)\right| + \right.$$
$$+e^{-3x^2/2}\left|\sum_{i=1}^{n}(f'(x_i) - p_n'(x_i))B_i(x)\right| +$$
$$\left. +e^{-3x^2/2}\left|\sum_{i=1}^{n}(y_i''' - (wp_n)'''(x_i))C_i(x)\right| + e^{-3x^2/2}|f''(0) - p_n''(0)||D_0(x)|\right].$$

Using lemmas 3, 4 and 5, we have

$$e^{-3x^2/2}|f(x) - R_n(x)| = O(1)\left[\omega\left(f', \frac{1}{\sqrt{n}}\right) + e^{-3x^2/2}\sum_{i=1}^{n}|y_i'''C_i(x)| + \right.$$
$$\left. +e^{-3x^2/2}\sum_{i=1}^{n}|(wp_n)'''(x_i)C_i(x)|\right].$$

By lemmas 2, 5 and (3.8), we have

$$e^{-3x^2/2}|f(x) - R_n(x)| = O(1)\omega\left(f', \frac{1}{\sqrt{n}}\right) +$$

$$+O(n)\omega\left(f', \frac{1}{\sqrt{n}}\right) \cdot e^{-3x^2/2}\left[\sum_{i=1}^{n}|C_i(x)|e^{x_i^2} + \sum_{i=1}^{n}|C_i(x)|\right] =$$

$$= O(1)\omega\left(f', \frac{1}{\sqrt{n}}\right).$$

Hence the theorem is proved.

5. Basic Estimates of Fundamental Polynomials (n even)

For n even, $A_i^*(x)$ and $B_i^*(x)$, given by (3.10) and (3.12), can be written in a convenient form as:

$$A_i^*(x) = l_i^3(x) - 3x_i B_i^*(x) + \frac{H_n^2(x)}{H_n'(x_i)^2} I - \frac{H_n^2(x)}{x_i^2 H_n'(x_i)^2} l_i(0), \tag{5.1}$$

where I is (4.16) and

$$B_i^*(x) = B_i(x) + \frac{H_n^2(x)}{x_i H_n'(x_i)^2} l_i(0), \tag{5.2}$$

where $B_i(x)$ is given by (4.8).

LEMMA 6. *For n even,*

$$\sum_{i=1}^{n} e^{x_i^2}|A_i^*(x)| = O\left(e^{3x^2/2}\sqrt{n}\right), \tag{5.3}$$

$$\sum_{i=1}^{n} e^{x_i^2}|B_i^*(x)| = O\left(e^{3x^2/2}\right), \tag{5.4}$$

and

$$\sum_{i=1}^{n} e^{x_i^2}|C_i^*(x)| = O\left(\frac{e^{3x^2/2}}{n}\right), \tag{5.5}$$

PROOF. The proof of this lemma is similar to that of Lemmas 2, 3 and 4, so we omit details.

PROOF OF THEOREM 4. Following the same steps as in the proof of Theorem 3, the theorem follows. We omit details.

References

[1] BALÁZS, J., Súlyozott $(0, 2)$-interpoláció ultraszférikus polinomok gyökein, *MTA III. Oszt. közl.*, **11** (1961), 305–338.

[2] BALÁZS, J. and TURÁN, P., Notes on the interpolation VIII, *Acta Math Acad. Sci. Hung.,* **12** (1961), 469–474.

[3] DATTA, S. and MATHUR P., On weighted $(0, 2)$-interpolation on infinite interval $(-\infty, +\infty)$, *Annales Univ. Sci. Budapest., Sect. Math.,* **42** (1999), 45–57.

[4] FREUD, G., On polynomial approximation with weight $\exp\left(-\frac{1}{2}x^{2k}\right)$, *Acta Math. Acad. Sci. Hungar.,* **24** (1973), 363–371.

[5] FREUD G., Lagrangesche interpolation über die Nullstellen der Hermiteschen Orthogonalpolynome, *Studia Sci. Math. Hung.,* **4** (1969), 179–190.

[6] JOÓ, I., On weighted $(0, 2)$-interpolation, *Annales Univ. Sci. Budapest., Sect. Math.,* **38** (1995), 185–222.

[7] MATHUR, K. K. and SAXENA, R. B., Weighted $(0, 1, 3)$-interpolation on the zeros of Hermite polynomial, *Acta. Math. Hung.,* **62** (1993), 31–47.

[8] SZEGŐ, G., *Orthogonal Polynomials,* Amer. Math. Soc. Coll. Publ. (New York, 1959).

[9] SZILI, L., Weighted $(0, 2)$-interpolation on the roots of Hermite polynomials, *Annales Univ. Sci. Budapest., Sect. Math.,* **27** (1984), 152–166.

ANNALES UNIV. SCI. BUDAPEST., **44** (2001), 83–93

INTERVAL OSCILLATION CRITERIA FOR SECOND ORDER FORCED NONLINEAR DIFFERENTIAL EQUATIONS WITH DAMPING*

By

W. LI and H-F. HUO

Department of Mathematics, Lanzhou University, Lanzhou, Gansu, Republic of China and Institute of Applied Mathematics, Gansu University of Technology Lanzhou, Gansu, People's Republic of China

(Received February 22, 2001)

1. Introduction

In this paper we shall be concerned with the second order forced nonlinear differential equation with damping

$$(r(t)y'(t))' + p(t)y'(t) + q(t)f(y(t))g(y'(t)) = e(t), \qquad t \geq t_0, \tag{1.1}$$

where r, q, f, g, e are to be specified in the following text.

We recall that a function $y : [t_0, t_1) \to (-\infty, +\infty)$, $t_i > t_0$ is called a solution of Eq. (1.1) if $y(t)$ satisfies Eq. (1.1) for all $t \in [t_0, t_1)$. In the sequel it will be always assumed that solutions of Eq. (1.1) exist for any $t_0 \geq 0$. A solution $x(t)$ of Eq. (1.1) is called oscillatory if it has arbitrary large zeros, otherwise it is called nonoscillatory.

When $r(t) - 1$, $p(t) \equiv 0$ and $e(t) \equiv 0$, Eq. (1.1) reduces to the equation

$$y''(t) + q(t)f(y(t))g(y'(t)) = 0, \tag{1.2}$$

which has been studied by GRACE and LALLI [7]. They mentioned that though stability, boundedness, and convergence to zero of all solutions of Eq. (1.2) have been investigated in the papers of BURTON and GRIMMER [1], GRACE and SPIKES [5, 6], LALLI [11], and WONG and BURTON [19], not much has been known regarding the oscillatory behavior of Eq. (1.2) except for the result by WONG and BURTON [19, Theorem 4] regarding oscillatory behavior of Eq. (1.2) in connection with that of the corresponding linear equation

$$y''(t) + q(t)y(t) = 0, \tag{1.3}$$

---

*Supported by the NNSF of China and the Foundation for University Key Teacher by the Ministry of Education of China.

Recently, ROGOVCHENKO [17] presented new sufficient conditions, which ensure oscillatory character of Eq. (1.2). They are different from those of GRACE and LALLI [7] and are applicable to other classes of equations which are not covered by the results of GRACE and LALLI [7]. However, all the mentioned above oscillation results involve the interval of q and hence require the information of q on the entire half-line $[t_0, +\infty)$. For related results refer to [2, 10, 13–16].

When $p(t) \equiv 0$ and $g(y) \equiv 1$, Eq. (1.1) reduces to the equation

$$(r(t)y'(t))' + q(t)f(y(t)) = e(t). \tag{1.4}$$

Numerous oscillation criteria have been obtained for Eq. (1.4); see KEENER [9], RAINKIN [16], SKIDMORE and BOWERS [20], SKIDMORE and LEIGHTON [21], and TEUFEL [22]. In these papers, the authors established oscillation criteria for a more general nonlinear equation by employing a technique introduced by KARTSATOS [8] where it is additionally assumed that $e(t)$ be the second derivative of an oscillatory function $h(t)$ and their oscillation results require the information of q on the entire half-line $[t_0, \infty)$.

However, from the Sturm Separation Theorem, we see that oscillation is only an interval property, i.e., if there exists a sequence of subintervals $[a_i, b_i]$ of $[t_0, \infty)$, as $a_i \to \infty$, such that for each i there exists a solution of equation

$$(r(t)y'(t))' + q(t)y(t) = 0, \qquad t \geq t_0 \tag{1.5}$$

that has at least two zeros in $[a_i, b_i]$, then every solution of Eq. (1.5) is oscillatory, no matter how "bad" Eq. (1.5) is (or r and q are) on the remaining parts of $[t_0, \infty)$.

EI-SAYED [4] applied this idea to oscillation and established an interval criterion for oscillation of a forced second order linear differential equation

$$(r(t)y'(t))' + q(t)y(t) = e(t), \qquad t \geq t_0. \tag{1.6}$$

THEOREM A. *Suppose that there exist two positive increasing divergent sequences $\{a_n\}$, $\{a_n\}$ and two sequences $\{c_n^+\}$, $\{c_n^-\}$ such that c_n^+, c_n^- are positive numbers and*

$$\begin{aligned} V_n^{\pm} = \int_{a_n^{\pm}}^{a_n^{\pm}+\pi/\sqrt{c_n^{\pm}}} &\left(c_n^{\pm}[1 - r(t)]\cos^2\left\{\sqrt{c_n^{\pm}}(t - a_n^{\pm}\right\} + \right. \\ &\left. + [q(t) - c_n^{\pm}]\sin^2\left\{\sqrt{c_n^{\pm}}(t - a_n^{\pm})\right\}\right) dt = 0, \end{aligned} \tag{1.7}$$

for all $n \geq n_0$*, where* n_0 *is a fixed positive integer. Suppose further that* $e(t)$ *satisfies*

$$e(t)\begin{cases} \geq 0, & t \in \left[a_n^+, a_n^+ + \frac{\pi}{\sqrt{c_n^+}}\right], \\ \leq 0, & t \in \left[a_n^-, a_n^- + \frac{\pi}{\sqrt{c_n^-}}\right], \end{cases} \tag{1.8}$$

for all $n \geq n_0$*. Then Eq.* (1.6) *is oscillatory.*

We note that the result is not very sharp, because it was proved with the aid of a comparison theorem of LEIGHTON [12] in the form given by COPPEL [3, Theorem 8, p11]. Recently, WONG [18] proved a more general oscillation result for Eq. (1.6).

THEOREM B. *Suppose that for any* $T \geq t_0$*, there exist* $T \leq s_1 < t_1 \leq \leq s_2 < t_2$ *such that*

$$e(t)\begin{cases} \leq 0, & t \in [s_1, t_1] \\ \geq 0, & t \in [s_2, t_2]. \end{cases} \tag{1.9}$$

Denote $D(s_i, t_i) = \{u \in C^1[s_i, t_i] : u(t) \not\equiv 0,\ u(s_i) = u(t_i) = 0\}$*,* $i = 1, 2$*. If there exists* $u \in D(s_i, t_i)$ *such that*

$$Q_i(u) = \int_{s_i}^{t_i} (qu^2 - ru')^2 \geq 0, \tag{1.10}$$

for $i = 1, 2$*, then Eq.* (1.6) *is oscillatory.*

Motivated by the ideas of EI-SAYED [4] and WONG [18], in this paper we obtain, by using a generalized Riccati technique which is introduced by LI [13] for the unforced equations and a new integral averaging technique, we obtain several new interval criteria for oscillation, that is, criteria given by the behavior of Eq. (1.1) (or of r, q, f, g and e) only on a sequence of subintervals of $[t_0, \infty)$. Finally, several examples that dwell upon the importance of our results are also included.

Hereinafter, we assume that

(H1) the functions $r : [t_0, \infty)$, $(0, \infty)$ and $p : [t_0, \infty) \to \mathbf{R}$ are continuous;

(H2) the function $q : [t_0, \infty) \to [0, \infty)$ is continuous and $q(t) \not\equiv 0$ on any ray $[T, \infty)$ for some $T \geq t_0$;

(H3) the function $f : \mathbf{R} \to \mathbf{R}$ is continuous and $yf(y) > 0$ for $y \neq 0$;

(H4) the function $g : \mathbf{R} \to \mathbf{R}$ is continuous and $g(y) \geq K > 0$ for $y \neq 0$.

In the sequel we say that a function $H := H(t)$ belongs to a function class $D(s_i, t_i) = \{H \in C^1[s_i, t_i] : H(t) \neq 0,\ H(s_i) = H(t_i) = 0\}$, $i = 1, 2$, denoted by $H \in D(s_i, t_i)$, $i = 1, 2$.

2. Main Results

THEOREM 1. *Let assumptions* (H1)–(H4) *hold. Suppose that*

$$f'(y) \geq \mu > 0 \qquad \text{for} \quad y \neq 0 \tag{2.1}$$

and that for any $T \geq t_0$, there exist $T \leq s_1 < t_1 \leq s_2 < t_2$ such that (1.9) *holds. If there exist $H \in D(s_i, t_i)$ and $g \in C^1([t_0, \infty), (0, \infty))$ such that*

$$\int_{s_1}^{t_1} H^2(t)\phi(t)dt \geq \frac{1}{4\mu}\int_{s_1}^{t_i} r(t)v(t)\left[-2H'(t) + \frac{p(t)}{r(t)}H(t)\right]^2 dt, \tag{2.2}$$

for $i = 1, 2$, where $v(t) = \exp(-2\mu \int^t g(s)ds)$ and

$$\phi(t) = v(t)[Kq(t) + \mu r(t)g^2(t) - p(t)g(t) - (r(t)g(t))'],$$

then every solution of Eq. (1.1) *is oscillatory.*

PROOF. Suppose that $y(t)$ is a nonoscillatory solution which is positive, say $y(t) > 0$ when $t \geq T_0$ for some T_0 depending on the solution $y(t)$. Denote

$$u(t)v(t)r(t)\left\{\frac{y'(t)}{f(y(t))} + g(t)\right\}, \qquad t \geq T_0. \tag{2.3}$$

It follows from (1.1) and (2.1) that $u(t)$ satisfies

$$\begin{aligned}
&u'(t) = \\
&= -2\mu g(t)u(t) + v(t)\left\{\frac{[r(t)y'(t)]'}{f(y(t))} - r(t)\frac{[y'(t)]^2 f'(y(t))}{f^2(y(t))} + [t(r)g(t)]'\right\} \leq \\
&\leq -2\mu g(t)r(t)v(t)\left\{\frac{y'(t)}{f(y(t))} + g(t)\right\} - v(t)Kq(t) + \frac{v(t)e(t)}{f(y(t))} - \\
&\quad - r(t)v(t)\frac{[y'(t)]^2\mu}{f^2(y(t))} + v(t)[r(t)g(t)]' - v(t)p(t)\frac{y'(t)}{f(y(t))} =
\end{aligned}$$

$$= -2\mu g(t)v(t)r(t)\frac{y'(t)}{f(y(t))} - 2\mu r(t)v(t)g^2(t) - v(t)Kq(t) + \frac{v(t)e(t)}{f(y(t))} -$$

$$- r(t)v(t)\frac{[y'(t)]^2\mu}{f^2(y(t))} + v(t)[r(t)g(t)]' - v(t)p(t)\frac{y'(t)}{f(y(t))} =$$

$$= \frac{\mu u^2(t)}{v(t)r(t)} - \frac{p(t)}{r(t)}u(t) - \phi(t) + \frac{v(t)e(t)}{f(y(t))}.$$

That is,

$$\phi(t) \leq -u'(t) - \frac{\mu u^2(t)}{v(t)r(t)} - \frac{p(t)}{r(t)}u(t) + \frac{v(t)e(t)}{f(y(t))}. \tag{2.4}$$

By assumption, we can choose s_1, $t_1 \geq T_0$ so that $e(t) \leq 0$ on the interval $I = [s_1, t_1]$ with $s_1 < t_1$. On the interval I, $u(t)$ satisfies by (2.4),

$$\phi(t) \leq -u'(t) - \frac{\mu u^2(t)}{v(t)r(t)} - \frac{p(t)}{r(t)}u(t). \tag{2.5}$$

Let $H \in D(s_1, t_1)$ be given as in hypothesis. Multiplying (2.5) by H^2 and integrating over I, we have

$$\int_{s_1}^{t_1} H^2(t)\phi(t)dt \leq -\int_{s_1}^{t_1} H^2(t)\left[u'(t) + \mu\frac{u^2(t)}{r(t)v(t)} + \frac{p(t)}{r(t)}\right]dt. \tag{2.6}$$

Integrating (2.6) by parts and using the fact that $H(s_1) = H(t_1) = 0$, we obtain

$$\int_{s_1}^{t_1} H^2(t)\phi(t)dt \leq -\int_{s_1}^{t_1} H^2(t)\left[u'(t) + \mu\frac{u^2(t)}{r(t)v(t)} + \frac{p(t)}{r(t)}\right]dt =$$

$$= -\int_{s_1}^{t_1}\left[\sqrt{\frac{\mu}{r(t)v(t)}}H(t)u(t) + \frac{1}{2}\sqrt{\frac{r(t)v(t)}{\mu}}\left(-2H'(t) + \frac{p(t)}{r(t)}H(t)\right)\right]^2 dt +$$

$$+ \int_{s_1}^{t_1}\frac{r(t)v(t)}{4\mu}\left[-2H'(t) + \frac{p(t)}{r(t)}H(t)\right]^2 dt <$$

$$< \int_{s_1}^{t_1}\frac{r(t)v(t)}{4\mu}\left[-2H'(t) + \frac{p(t)}{r(t)}H(t)\right]^2 dt.$$

which contradicts the condition (2.2). This contradiction proves that $y(t)$ is oscillatory.

When $y(t)$ is eventually negative, we see $H \in D(s_2, t_2)$ and $e(t) \geq 0$ on $[s_2, t_2]$ to reach a similar contradiction. The proof is complete.

When $p(t) \equiv 0$, by Theorem 1, we have the following corollary.

COROLLARY 1. *Let assumptions* (H1)–(H4) *and* (2.1) *hold. Suppose that for any* $T \geq t_0$, *there exist* $T \leq s_1 < t_1 \leq s_2 < t_2$ *such that* (1.9) *holds. If there exist* $H \in D(s_i, t_i)$ *and* $g \in C^1([t_0, \infty), (0, \infty))$ *such that*

$$\int_{s_i}^{t_i} H^2(t)\phi(t)dt \geq \frac{1}{\mu} \int_{s_i}^{t_i} r(t)v(t)[H'(t)]^2 dt, \tag{2.7}$$

for $i = 1, 2$, *where* $v(t) = \exp\left(-2\mu \int^t g(s)ds\right)$ *and*

$$\phi(t) = v(t)[Kq(t) + \mu r(t)g^2(t) - (r(t)v(t))'],$$

then every solution of Eq. (1.1) *is oscillatory.*

We remark that, if we take $g(t) = 0$, then $v(t) = 1$, $\phi(t) = q(t)$. Hence Corollary 1 also reduces to Theorem B of Wong if $f(y) = y$.

For the case when $f(y)$ is not monotonous or has no continuous derivative, the following result holds.

THEOREM 2. *Suppose that* (H1)–(H4) *hold. Let assumption* (2.1) *in Theorem be replaced by*

$$\frac{f(y)}{y} \geq c > 0 \qquad \text{for} \quad y \neq 0, \tag{2.8}$$

where c *is a constant. Suppose that* $q(t) \geq 0$ *and that for any* $T \geq t_0$, *there exist* $T \leq s_1 < t_1 \leq s_2 < t_2$ *such that* (1.9) *holds. If there exist* $H \in D(s_i, t_i)$ *and* $g \in C^1([t_0, \infty), (0, \infty))$ *such that*

$$\int_{s_i}^{t_i} H^2(t)\phi(t)dt \geq \frac{1}{4} \int_{s_i}^{t_u} r(t)v(t) \left[-2H'(t) + \frac{p(t)}{r(t)}H(t)\right]^2 dt, \tag{2.9}$$

for $i = 1, 2$, *where* $v(t) = \exp\left(-2 \int^t g(s)ds\right)$ *and*

$$\phi(t) = v(t)[Kcq(t) + r(t)g^2(t) - p(t)g(t) - (r(t)g(t))'],$$

then every solution of Eq. (1.1) *is oscillatory.*

PROOF. Suppose that $y(t)$ is a nonoscillatory solution which is positive, say $y(t) > 0$ when $t \geq T_0$ for some T_0 depending on the solution $y(t)$. Denote

$$u(t) = v(t)r(t)\left\{\frac{y'(t)}{y(t)} + g(t)\right\}, \qquad t \geq T_0. \tag{2.10}$$

It follows from (1.1) and (2.8) that $u(t)$ satisfies

$$\begin{aligned}
u'(t) &= -2g(t)u(t) + v(t)\left\{\frac{[r(t)y'(t)]'}{y(t)} - r(t)\frac{[y'(t)]^2}{y^2(t)} + [r(t)g(t)]'\right\} \leq \\
&\leq -2g(t)r(t)v(t)\left\{\frac{y'(t)}{y(t)} + g(t)\right\} - v(t)Kcq(t) + \frac{v(t)e(t)}{y(t)} - \\
&\quad - r(t)v(t)\frac{[y'(t)]^2}{y^2(t)} + v(t)[r(t)g(t)]' - v(t)p(t)\frac{y'(t)}{y(t)} = \\
&= -2g(t)v(t)r(t)\frac{y'(t)}{y(t)} - 2r(t)v(t)g^2(t) - v(t)Kcq(t) + \frac{v(t)e(t)}{y(t)} - \\
&\quad - r(t)v(t)\frac{[y'(t)]^2}{y^2(t)} + v(t)[r(t)g(t)]' - v(t)p(t)\frac{y'(t)}{y(t)} = \\
&= -\frac{u^2(t)}{v(t)r(t)} - \frac{p(t)}{r(t)}u(t) - \phi(t) + \frac{v(t)e(t)}{y(t)}.
\end{aligned}$$

That is,

$$\phi(t) \leq -u'(t) - \frac{u^2(t)}{v(t)r(t)} - \frac{p(t)}{r(t)}u(t) + \frac{v(t)e(t)}{y(t)}. \tag{2.11}$$

By assumption, we can choose s_1, $t_1 \geq T_0$ so that $e(t) \leq 0$ on the interval $I = [s_1, t_1]$ with $s_1 < t_1$. On the interval I, $u(t)$ satisfies by (2.11),

$$\phi(t) \leq -u'(t) - \frac{u^2(t)}{v(t)r(t)} - \frac{p(t)}{r(t)}u(t). \tag{2.12}$$

Similar to the proof of Theorem 1, we can obtain a contradiction. The proof is complete.

If $p(t) \equiv 0$, then, by Theorem 2, we have the following corollary.

COROLLARY 2. *Let the assumptions* (H1)–(H4) *and* (2.8) *hold. Suppose that* $q(t) \geq 0$ *and that for any* $T \geq t_0$, *there exist* $T \leq s_1 < t_1 \leq s_2 < t_2$ *such* *that* (1.9) *holds. If there exist* $H \in D(s_i, t_i)$ *and* $g \in C^1([t_0, \infty), (0, \infty))$ *such that*

$$\int_{s_i}^{t_i} H^2(t)\phi(t)dt \geq \int_{s_i}^{t_i} r(t)v(t)[H'(t)]^2 dt,$$

for $i = 1, 2$, *where* $v(t) = \exp\left(-2\int^t g(s)ds\right)$ *and*

$$\phi(t) = v(t)[Kcq(t) + r(t)g^2(t) - [r(t)g(t)]'],$$

then every solution of Eq. (1.1) *is oscillatory.*

3. Examples

In this section we will show the applications of our oscillation criteria by two examples. We will see that the equations in the examples are oscillatory based on the results in Section 2, though the oscillation cannot be demonstrated by the results of WONG [18].

EXAMPLE 1. Consider the following nonlinear differential equation

$$(\sqrt{t}y'(t))' - 2y'(t) + \frac{5}{4\sqrt{t}(1+\sin^4\sqrt{t})}y(t)(1+y^4(t)) =$$

$$= \frac{1}{\sqrt{t}}(\sin\sqrt{t} - \cos\sqrt{t}), \quad t \geq 1. \tag{3.1}$$

Here the zeros of the forcing term $\frac{1}{\sqrt{t}}(\sin\sqrt{t} - \cos\sqrt{t})$ are $\left(n\pi + \frac{\pi}{4}\right)^2$. Clearly,

$$f(y) = y(1+y^4) \qquad \text{and} \qquad f'(y) = 1 + 5y^4 \geq 1 = \mu.$$

Let $H(t) = \sin\sqrt{t}$. For any $T \geq 1$, choose n sufficiently large so that $\left(n\pi + \frac{\pi}{4}\right)^2 \geq T$ and set $s_1 = \left(n\pi + \frac{\pi}{4}\right)^2$ and $t_1 = \left[(n+1)\pi + \frac{\pi}{4}\right]^2$ in (2.2). Pick up $g(t) = 0$, then $v(t) = 1$. It is easy to verify that

$$\int_{\left[n\pi+\frac{\pi}{4}\right]^2}^{\left[(n+1)\pi+\frac{\pi}{4}\right]^2} H^2(t)\phi(t)dt = \int_{\left[n\pi+\frac{\pi}{4}\right]^2}^{\left[(n+1)\pi+\frac{\pi}{4}\right]^2} \sin^2\sqrt{t}\frac{5}{4\sqrt{t}(1+\sin^4\sqrt{t})}dt =$$

$$= \int_{n\pi+\frac{\pi}{4}}^{(n+1)\pi+\frac{\pi}{4}} \sin^2 s \frac{5}{4s(1+\sin^4 s)} 2s\, ds = \int_{n\pi+\frac{\pi}{4}}^{(n+1)\pi+\frac{\pi}{4}} \frac{5}{2} \sin^2 s \frac{1}{1+\sin^4 s} ds =$$

$$= \int_{n\pi+\frac{\pi}{4}}^{(n+1)\pi+\frac{\pi}{4}} \frac{5}{2} \sin^2 s \frac{1}{1+\sin^2 s(1-\cos^2 s)} ds \geq \int_{n\pi+\frac{\pi}{4}}^{(n+1)\pi+\frac{\pi}{4}} \frac{5}{2} \frac{\sin^2 s}{1+\sin^2 s} ds =$$

$$= \int_{n\pi+\frac{\pi}{4}}^{(n+1)\pi+\frac{\pi}{4}} \frac{5}{2} \frac{\sin^2 s + 1 - 1}{1+\sin^2 s} ds = \frac{5}{2}\pi - \int_{n\pi+\frac{\pi}{4}}^{(n+1)\pi+\frac{\pi}{4}} \frac{5}{2} \frac{1}{1+\sin^2 s} ds \geq$$

$$\leq \frac{5\pi}{2} - \frac{5}{2} \int_{n\pi+\frac{\pi}{4}}^{(n+1)\pi+\frac{\pi}{4}} \frac{1}{2|\sin s|} ds \geq \frac{5\pi}{2} - \frac{5\pi}{4} = \frac{5\pi}{4},$$

where we have used the inequality $1 + \sin^2 t \geq 2 \sin t$, and

$$\frac{1}{4\mu} \int_{s_1}^{t_1} r(t)v(t) \left[-2H'(t) + \frac{p(t)}{r(t)} H(t) \right]^2 dt =$$

$$= \frac{1}{4} \int_{\left(n\pi+\frac{\pi}{4}\right)}^{\left(n\pi+\pi+\frac{\pi}{4}\right)^2} \sqrt{t} \left[-2\frac{\cos\sqrt{t}}{2\sqrt{t}} - 2\sqrt{t}\sin\sqrt{t} \right]^2 dt =$$

$$= \frac{1}{4} \int_{\left(n\pi+\frac{\pi}{4}\right)}^{\left(n\pi+\pi+\frac{\pi}{4}\right)} s \left[-\frac{\cos s}{s} - \frac{2}{s} \sin s \right]^2 ds =$$

$$= \frac{1}{2} \int_{\left(n\pi+\frac{\pi}{4}\right)}^{\left(n\pi+\pi+\frac{\pi}{4}\right)} [\cos^2 s + 2\sin 2s + 4\sin^2 s] ds = \frac{5}{4}\pi,$$

which implies that (2.2) holds for $i = 1$.

Similarly, for $s_2 = \left[(n+1)\pi + \frac{\pi}{4}\right]^2$ and $t_2 = \left[(n+2)\pi + \frac{\pi}{4}\right]^2$, we can show that (2.2) holds. It follows from Theorem 1 that every solution of Eq. (3.1) is oscillatory. Observe that $y(t) = \sin\sqrt{t}$ is such a solution. However, the results of WONG [18] fail for the oscillation of Eq. (3.1).

EXAMPLE 2. Consider the following nonlinear differential equation

$$\begin{aligned}(\sqrt{t}y'(t))' - 2y'(t) + \frac{10(1+\sin^2\sqrt{t})}{\sqrt{t}(9+\sin^2\sqrt{t})}y(t)\left(\frac{1}{8} + \frac{1}{1+y^2(t)}\right) = \\ = \frac{1}{\sqrt{t}}(\sin\sqrt{t} - \cos\sqrt{t}), \qquad t \geq 1.\end{aligned} \tag{3.2}$$

Let $f(y) = y\left[\frac{1}{8} + \frac{1}{1+y^2}\right]$, then

$$f'(y) = \frac{(y^2-3)^2}{8(1+y^2)^2}.$$

Clearly, the condition, $f'(y) \geq \mu > 0$ for $y \neq 0$ does not hold. Hence, Theorem 1 is not valid for Eq. (3.2). However,

$$\frac{f(y)}{y} = \frac{1}{8} + \frac{1}{1+y^2} \geq \frac{1}{8} = K > 0.$$

Similar to the proof of Example 1, we see that the assumptions of Theorem 2 are satisfied. Hence, every solution of Eq. (3.2) is oscillatory. Observe that $y(t) = \sin\sqrt{t}$ is an oscillatory solution. Similarly, the results of WONG [18] are not valid for the oscillation of Eq. (3.2).

References

[1] T. A. BURTON and R. GRIMMER, Stability properties of $(ru')' + af(u)g(u') = 0$, *Monatsh Math.,* **74** (1970), 211–222.

[2] J. S. CASSELL, The asymptotic behavior of a class of linear oscillators, *Quart. J. Math. Oxford Ser.(3),* **32** (1981), 287–302.

[3] W. A. COPPEL, *Stability and Asymptotic Behavior of Differential Equations,* Heath, Boston, 1965.

[4] M. A. EI-SAYED, An oscillation criterion for a forced second order linear differential equations, *Proc. Amer. Math. Soc.,* **118** (1993), 813–817.

[5] J. R. GRAEF and P. W. SPIKES, Asymptotic behavior of solutions of a second order nonlinear differential equation, *J. Differential Equations,* **17** (1975), 451–476.

[6] J. R. GRAEF and P. W. SPIKES, Boundedness and convergence to zeros of solutions of a forced second order nonlinear differential equations, *J. Math. Anal. Appl.,* **62** (1978), 295–309.

[7] S. R. GRACE and B. S. LALLI, An oscillation criterion for second order strongly sublinear differential equations, *J. Math. Anal. Appl.,* **123** (1987), 584–588.

[8] A. G. KARTSATOS, Maintenance of oscillations under the effect of a periodic forcing term, *Proc. Amer. Math. Soc.,* **33** (1972), 377–383.

[9] M. S. KEENER, Solutions of a certain linear nonhomogeneous second order, differential equations, *Appl. Anal.,* **1** (1971), 57–63.

[10] C. C. HUANG, Oscillation and nonoscillation for second order linear differential equations, *J. Math. Anal. Appl.,* **210** (1997), 712–723.

[11] B. S. LALLI, On boundedness of solutions of certain second order differential equations, *J. Math. Anal. Appl.,* **25** (1969), 182–188.

[12] W. LEIGHTON, Comparion theorems for linear differential equations of second order, *Proc. Amer. Math. Soc.,* **13**, (1962), 603–610.

[13] W. T. LI, Oscillation of certain second-order nonlinear differential equations, *J. Math. Anal. Appl.,* **217** (1998), 1–14.

[14] W. T. LI and X. H. LI, Oscillation criteria for second order nonlinear differential equations with integrable coefficients, *Appl. Math. Lett.,* **8(3)** (2000), 1–6.

[15] W. T. LI, M. Y. ZHANG and X. L. FEI, Oscillation criteria for second order nonlinear differential equations with damping term, *Indian J. Pure Appl. Math.,* **30(10)** (1999), 1017–1029.

[16] S. M. RAINKIN, Oscillation theorems for second order nonhomogeneous linear differential equations, *J. Math. Anal. Appl.,* **153** (1976), 550–553.

[17] YURI V. ROGOVCHENKO, Oscillation criteria for certain nonlinear differential equations, *J. Math. Anal. Appl.,* **229** (1999), 399–416.

[18] J. S. WONG, Oscillation criteria for a forced second order linear differential equation, *J. Math. Anal. Appl.,* **231** (1999), 235–240.

[19] J. S. W. WONG and T. A. BURTON, Some properties of solution of $u'' + a(t)f(u)g(u') = 0$, *Monatsh. Math.,* **69** (1965), 364–374.

[20] A. SKIDMORE and J. J. BOWERS, Oscillatory behavior of solutions of $y''+p(x)y = f(x)$, *J. Math. Anal. Appl.,* **49** (1975), 317–323.

[21] A. SKIDMORE and W. LEIGHTON, On the equation $y'' + p(x)y = f(x)$, *J. Math. Anal. Appl.,* **43** (1973), 46–55.

[22] H. TEUFEL, Forced second order nonlinear oscillations, *J. Math. Anal. Appl.,* **40** (1972), 148–152.

ANNALES UNIV. SCI. BUDAPEST., **44** (2001), 95–105

AREA-INRADIUS AND DIAMETER-INRADIUS RELATIONS FOR COVERING PLANE SETS

By

MARÍA A. HERNÁNDEZ CIFRE

Departamento de Matemáticas, Universidad de Murcia, Murcia, Spain

(Received September 18, 2001)

1. Introduction and first results

Let $\mathcal{K}^2$ denote the family of compact convex sets K in the Euclidean plane E^2. For $K \in \mathcal{K}^2$, let $A(K)$, $r(K)$, $D(K)$ and $\omega(K)$ be the area, the inradius, the diameter and the minimal width (that is, the smallest distance between two parallel support lines) of K, respectively.

For an arbitrary lattice L and a set K, the *lattice point enumerator* is denoted by $G(K, L) = \mathrm{card}(\mathrm{int}(K) \cap L)$. A convex set K is called a *lattice-point-free convex set* with respect to L if $G(K, L) = 0$. Further, K is a *covering set* if

$$K + L = \{K + g,\ g \in L\} = E^2.$$

A great number of results concerning covering sets with respect to the integer lattice $\mathbb{Z}^2$ are known. However there are relatively few results on covering sets with respect to an arbitrary lattice L ([6], [7], [8], [9], [10]). The following elegant result was obtained by Awyong and Scott in [2]: an inequality concerning the inradius and the area of a planar lattice-point-free convex set, in the case where L is the integer lattice $\mathbb{Z}^2$.

Theorem 1. *Let K be a compact, planar, convex set with $G(K, \mathbb{Z}^2) = 0$. Then*

$$(2r - 1)A \leq 2(\sqrt{2} - 1), \tag{1}$$

This work is supported in part by Dirección General de Investigación (MCYT) BFM2001-2871.

with equality when and only when K is congruent to the diagonal square shown in Figure 1.

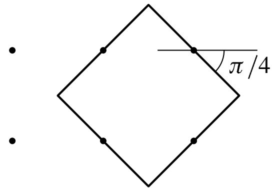

Figure 1

The d-dimensional analog was solved by AWYONG, HENK and SCOTT in [1]. In this paper, we generalize Theorem 1 to rectangular lattices, using then this result to obtain, in the last section, a more general inequality for arbitrary lattices, and some other related results.

Before stating the main theorem, let us introduce some notation. We denote by Γ_{uv} the rectangular lattice generated by the vectors $(u, 0)$ and $(0, v)$, with $0 < u \leq v$.

Now, we define the rombhus $\mathcal{Q}_r$ as follows:

Let $P = (u/2, v/2) \in E^2$, and $B_P(r)$ be the disc centered in P and with radius r. Following the notation of *Figure 2*, for each fixed r we denote by $\mathcal{Q}_r$ the rombhus with sides tangent to the disc $B_P(r)$, which pass through the points $P_1 = (u, v)$, $P_2 = (0, v)$, $P_3 = (0, 0)$ and $P_4 = (u, 0)$ respectively, and with angle $\alpha \leq \arctan(u/v)$.

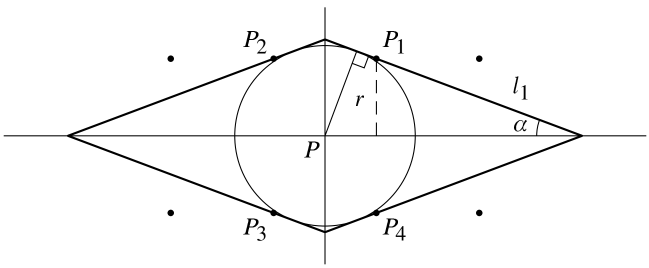

Figure 2. Optimal rombhus

This set will have an important role in our results. It is easy to check that

$$A(\mathcal{Q}_r) = \frac{2(u^2 + v^2)^2 r^2}{(2ur - v\sqrt{u^2 + v^2 - 4r^2})(2vr + u\sqrt{u^2 + v^2 - 4r^2})}.$$

Let us also denote by $b_{uv}(r)$ the function

$$(2) \qquad b_{uv}(r) = \frac{2(u^2+v^2)^2(2r-v)r^2}{(2ur - v\sqrt{u^2+v^2-4r^2})(2vr + u\sqrt{u^2+v^2-4r^2})}.$$

We prove the following result:

THEOREM 2. *Let r_* denote the unique solution of the equation*

$$(3) \quad uv(r+v)\sqrt{u^2+v^2-4r^2} + r[v(3u^2+v^2) - 8r^2(r+v) + 4u^2r] = 0.$$

For each $K \in \mathcal{K}^2$ with $G(K, \Gamma_{uv}) = 0$ it holds:

i) *If $v \geq 2u$ then $(2r(K) - v)A(K) \leq \frac{1}{2}uv^2$.*
ii) *If $v < 2u$ then $(2r(K) - v)A(K) \leq b_{uv}(r_*)$.*

These inequalities are tight in the following sense:

i) *$\frac{1}{2}uv^2$ cannot be replaced by $\frac{1}{2}uv^2 - \varepsilon$, because the equality would be attained for the case $r(K) = v/2$, when K is the infinite strip.*
ii) *Now, equality holds when $K = \mathcal{Q}_{r*}$ (up to congruence).*

2. Proof of the Theorem

The proof of the theorem will be established by proving two previous lemmas, the second of which is a result from elementary calculus.

LEMMA 1. *Let $K \in \mathcal{K}^2$ be a convex domain such that $G(K, \Gamma_{uv}) = 0$. Then there exists another convex domain $K^s \in \mathcal{K}^2$ containing no points of Γ_{uv}, such that:*

i) *$A(K^s) = A(K)$ and $r(K^s) \geq r(K)$*
ii) *K^s is symmetric about the lines $x = u/2$, $y = v/2$.*

PROOF. Let K' be the convex domain which is obtained from K by Steiner symmetrization with respect to the line $x = u/2$. It is well known that Steiner symmetrization preserves the convexity and the area, and does not decrease the inradius [3]. Therefore, K' is a convex domain with $A(K') = A(K)$, and $r(K') \geq r(K)$. Now, we have to see that $G(K', \Gamma_{uv}) = 0$.

Let us suppose that K' contains the lattice point of Γ_{uv}, $mu + nv$, with $m, n \in \mathbb{Z}$. Then, because of the symmetry of K' about $x = u/2$, the line $y = nv$ would intersect K' in a line segment of length greater than u. So, this line would also intersect K in a line segment with the same length, which implies that $G(K,\Gamma_{uv}) > 0$, contradicting the hypothesis. Hence, $G(K',\Gamma_{uv}) = 0$.

We can use an analogous argument, but now symmetrizing about the line $y = v/2$, to obtain a convex domain K^s with $A(K) = A(K^s)$, $r(K) \leq r(K^s)$ and $G(K^s,\Gamma_{uv}) = 0$. By construction, K^s is symmetric about the lines $x = u/2$ and $y = v/2$, and the lemma is proved. ∎

LEMMA 2. *Let*

$$h(r) = 2uv(r+v)\sqrt{u^2+v^2-4r^2} + 2r[v(3u^2+v^2) - 8r^2(r+v) + 4u^2r],$$

with $0 < u \leq v$ *and* $r \in (v/2, \sqrt{u^2+v^2}/2]$. *Then,*

i) *If* $v \geq 2u$, $h(r) < 0$.

ii) *If* $v < 2u$, $h(r)$ *vanishes exactly in an only point* $r^* \in (v/2, \sqrt{u^2+v^2}/2]$.

PROOF. Let us denote by

$$h_1(r) = (r+v)\sqrt{u^2+v^2-4r^2},$$
$$h_2(r) = r[v(3u^2+v^2) - 8r^2(r+v) + 4u^2r].$$

Then, $h(r) = 2uvh_1(r) + 2h_2(r)$.

It is easy to compute that when $r \in (v/2, \sqrt{u^2+v^2}/2]$, it holds

$$h_1''(r) = 4\frac{8r^3 - (u^2+v^2)(3r+v)}{(u^2+v^2-4r^2)^{3/2}} < 0,$$

so, $h_1(r)$ is a concave function, and moreover, $h_1'(r)$ is strictly decreasing. Analogously, we can check that

$$h_2''(r) = -96r^2 - 48vr + 8u^2 < 0,$$

and then, $h_2(r)$ is a concave function and $h_1'(r)$ is strictly decreasing.

Hence, we can deduce that the original function $h(r)$ is concave and its first derivative $h'(r)$ is strictly decreasing. So, $h'(r) < h'(v/2)$. Now,

$$h\left(\sqrt{u^2+v^2}/2\right) = -(v^2-u^2)\sqrt{u^2+v^2}\left(\sqrt{u^2+v^2}+v\right) \leq 0,$$
$$h(v/2) = 2v^2(4u^2-v^2) \qquad \text{and}$$
$$h'(v/2) = 8v(2u^2-3v^2).$$

Then we obtain that

i) If $v \geq 2u$, $h(v/2) \leq 0$ and also $h'(v/2) \leq 0$. So, $h'(r) < 0$ and therefore $h(r)$ is strictly decreasing and negative.

ii) If $v < 2u$, $h(v/2) > 0$. Then, there exists an $r^* \in (v/2, \sqrt{u^2+v^2}/2]$ such that $h(r^*) = 0$, and this point is unique because of the strict concavity of $h(r)$. ∎

Now we can prove our theorem.

Let us define the functional $f(K) = (2r(K) - v)A(K)$. Applying Lemma 1 to the set K, we may obtain a new convex set $K^s \in \mathcal{K}^2$ with $G(K^s, \Gamma_{uv}) = 0$, satisfying the conditions:

i) $A(K^s) = A(K)$ and $r(K^s) \geq r(K)$,

ii) K^s is symmetric about the lines $x = u/2$, $y = v/2$.

Then, it is clear that $f(K) \leq f(K^s)$. So, it suffices to prove the theorem for sets $K \in \mathcal{K}^2$ which are symmetric about the lines $x = u/2$ and $y = v/2$.

To fully utilize the symmetry of K about the lines $x = u/2$ and $y = v/2$, we move the origin to the point $P = (u/2, v/2)$.

Obviously, the area of a lattice-point-free convex set $K \in \mathcal{K}^2$ with respect to Γ_{uv} can be arbitrary large. However, the inradius of such a set is bounded above by $\sqrt{u^2+v^2}/2$. Besides, if $r(K) \leq v/2$, then $(2r(K) - - v)A(K) \leq 0$ and the result is trivially true. Hence, we may assume that $v/2 < r(K) \leq \sqrt{u^2+v^2}/2$.

Since int(K) does not contain the points

$$P_1 = \left(\frac{u}{2}, \frac{v}{2}\right), \qquad P_2 = \left(-\frac{u}{2}, \frac{v}{2}\right),$$
$$P_3 = \left(-\frac{u}{2}, -\frac{v}{2}\right), \quad P_4 = \left(\frac{u}{2}, -\frac{v}{2}\right),$$

it follows by the convexity of K that for each $i = 1, \ldots, 4$, K is bounded by a line l_i passing through P_i, with slopes $m(l_2) = m(l_4) = -m(l_1) = -m(l_3) \geq 0$ (by the symmetry of K). So, K lies within a rhombus Q determined by the lines l_i, $i = 1, \ldots, 4$. Since $K \subset Q$, clearly $A(K) \leq A(Q)$ and $r(K) \leq r(Q) = = r_Q$, and we have $f(K) \leq f(Q)$. It is therefore sufficient to maximize $f(K)$ over the set of all rhombi $Q \in \mathcal{K}^2$, determined by the lines l_i, $i = 1, \ldots, 4$. But moreover; if the acute angle α determined by the line l_1 and the OX-axis (see *Figure 2*) is not greater than $\arctan(u/v)$, then $Q = \mathcal{Q}_{r_Q}$.

But if $\alpha > \arctan(u/v)$, it is not difficult to compute that the area of such a rhombus Q in terms of its inradius r_Q takes the value

$$A(Q) = \frac{2(u^2+v^2)^2 r_Q^2}{\left(2ur_Q + v\sqrt{u^2+v^2-4r_Q^2}\right)\left(2vr_Q - u\sqrt{u^2+v^2-4r_Q^2}\right)}.$$

Since $u \le v$, we can obtain easily that

$$\left(2ur_Q - v\sqrt{u^2+v^2-4r_Q^2}\right)\left(2vr_Q + u\sqrt{u^2+v^2-4r_Q^2}\right) \le$$
$$\le \left(2ur_Q + v\sqrt{u^2+v^2-4r_Q^2}\right)\left(2vr_Q - u\sqrt{u^2+v^2-4r_Q^2}\right).$$

Then, the rhombus $\mathcal{Q}_{R_Q}$ (with the same inradius as Q) has area strictly greater than the area of Q. Therefore, again $f(Q) \le f(\mathcal{Q}_{r_Q})$, and so, we have just to maximize $f(K)$ over the set of all rhombi $\mathcal{Q}_r$.

We had gotten the area of $\mathcal{Q}_r$, as a function of r, so

$$f(\mathcal{Q}_r) = (2r(\mathcal{Q}_r) - v)A(\mathcal{Q}_r) = b_{uv}(r) =$$
$$= \frac{2(u^2+v^2)^2(2r-v)r^2}{\left(2ur - v\sqrt{u^2+v^2-4r^2}\right)\left(2vr + u\sqrt{u^2+v^2-4r^2}\right)}.$$

But $b_{uv}(r)$ can be written in the following way

$$b_{uv}(r) = 2(u^2+v^2)\frac{r^2\left(2ur + v\sqrt{u^2+v^2-4r^2}\right)}{(2r+v)\left(2vr + u\sqrt{u^2+v^2-4r^2}\right)},$$

and then, it is not difficult to see that

$$b'_{uv}(r) = \frac{2(u^2+v^2)^2 r}{(2r+v)\left(2vr + u\sqrt{u^2+v^2-4r^2}\right)^2\sqrt{u^2+v^2-4r^2}}h(r),$$

where $h(r)$ is the function defined in Lemma 2.

i) If $v \ge 2u$, Lemma 2 assures us that $h(r) < 0$; so $b_{uv}(r)$ is strictly monotonously decreasing, and then

$$b_{uv}(r) < \lim_{r \to v/2} b_{uv}(r) = \frac{1}{2}uv^2.$$

Clearly, this bound can not be replaced by a $\frac{1}{2}uv^2 - \varepsilon$, because the equality would be attained when $r = v/2$, i.e., when K is the infinite strip.

ii) If $v < 2u$, because of the proof of Lemma 2, we can assure that there exists a unique solution r^* of the corresponding equation (3); so $b_{uv}(r)$ attains its maximum value for $r = r^*$. Hence, in this case,

$$b_{uv}(r) \leq b_{uv}(r^*),$$

and the equality holds when $K = \mathcal{Q}_{r^*}$ (up to congruence).

This completes the proof of the theorem. ∎

3. Some results for arbitrary lattices

Let us denote by $\mathscr{L}^2$ the set of lattices $L \subset E^2$ with $\det L \neq 0$. Further, for $L \in \mathscr{L}^2$ let $\lambda_i = \lambda_i(L)$ be the successive minima of L, i.e., $\lambda_i(L) = \lambda_i(B^2, L) = \min\{\lambda > 0 \mid \dim \operatorname{aff}(\lambda B^2 \cap L) \geq i\}$, and let $\mu_i = \mu_i(L)$ be the covering minima of the lattice L, i.e., $\mu_i(L) = \mu_i(B^2, L) = \min\{\mu > 0 \mid \mu B^2 + g, g \in L$, meets every flat F of E^2 with $\dim(F) = 2 - i\}$.

We remember also that a basis $\{b_1, b_2\}$ of L is *reduced* (in the sense of Minkowski) if

i) $b_1 \in \{v \in L \setminus \{0\} : \|v\| \text{ is minimal}\}$

ii) $b_2 \in \{v \in L \setminus \{0\} : b_1, v \text{ are a basis of } L, \|v\| \text{ is minimal}\}$

iii) $b_1 \cdot b_2 \geq 0$.

For the sake of brevity we will represent by η the maximum $\eta = \max\{\lambda_1, 2\mu_1\}$ and by δ the minimum $\delta = \min\{\lambda_i, 2\mu_1\}$. Moreover, we denote by $b_{\delta\eta}(r)$ the corresponding function defined by (2) when $u = \delta$ and $v = \eta$. We will prove the following result.

THEOREM 3. *Let r_* denote the unique solution of the equation*

$$\delta\eta(r+\eta)\sqrt{\delta^2+\eta^2-4r^2} + r[\eta(3\delta^2+\eta^2) - 8r^2(r+\eta) + 4\delta^2 r] = 0 \tag{4}$$

For each $K \subset \mathcal{K}^2$ and $L \in \mathscr{L}^2$, with $G(K, L) = 0$ it holds:

i) *If $\eta \geq 2\delta$ then $(2r(K) - \eta)A(K) \leq \frac{1}{2}\delta\eta^2$.*

ii) *If $\eta < 2\delta$ then $(2r(K) - \eta)A(K) \leq b_{\delta\eta}(r_*)$.*

The inequalities are tight.

REMARK 1. Because of $2\mu_1 \geq \frac{\sqrt{3}}{2}\lambda_1$ (see [5]), it will never hold $\lambda_1 > 2(2\mu_1)$. So, when $\eta = \max\{\lambda_1, 2\mu_1\} = \lambda_1$, we will have just the upper bound $(2r-\eta)A \leq b_{2\mu_1\lambda_1}(r_*)$. Thus, for an arbitrary lattice, we will have the following three possible cases:

i) If $\mu_1 \geq \lambda_1$ then $(2r(K) - 2\mu_1)A(K) \leq 2\lambda_1\mu_1^2$.
ii) If $\mu_1 < \lambda_1 \leq 2\mu_1$ then $(2r(K) - 2\mu_1)A(K) \leq b_{\lambda_1 2\mu_1}(r_*)$.
iii) If $2\mu_1 \leq \lambda_1 < 4\mu_1$ then $(2r(K) - \lambda_1)A(K) \leq b_{2\mu_1\lambda_1}(r_*)$.

with r_* each solution of the corresponding equations we obtain from (4).

For instance, in the case of the integer lattice $\mathbb{Z}^2$, $\lambda_1 = 2\mu_1 = 1$, and we obtain $r_* = \sqrt{2}/2$. So, the upper bound for the relation $(2r-1)A$ takes the value $b_{11}(\sqrt{2}/2) = 2(\sqrt{2}-1)$, which was proved by AWYONG and SCOTT in [2].

From this theorem, we can obtain as an obvious consequence the following corollary.

COROLLARY 1. *Let $K \in \mathcal{K}^2$ and $L \in \mathcal{L}^2$ be given such that*

$$(2r(K) - \eta)A(K) > \max\left\{\frac{1}{2}\delta\eta^2, b_{\delta\eta}(r_*)\right\}.$$

Then K is a covering set.

We also prove an inequality relating the inradius and the diameter of a lattice-point-free convex set $K \in \mathcal{K}^2$.

PROPOSITION 1. *Let K be a convex set of $\mathcal{K}^2$ and $L \in \mathcal{L}^2$, with $G(K, L) = 0$. Then,*

$$(2r(K) - 2\mu_1)(D(K) - \lambda_1) \leq 2\mu_1\lambda_1.$$

The limiting infinite strip shows that the stated bound is best possible.

COROLLARY 2. *Let $K \in \mathcal{K}^2$ and $L \in \mathcal{L}^2$ be given such that*

$$(2r(K) - 2\mu_1)(D(K) - \lambda_1) > 2\mu_1\lambda_1.$$

Then K is a covering set.

We observe that this inequality can be rewritten as

$$\frac{\lambda_1}{D(K)} + \frac{\mu_1}{r(K)} < 1.$$

So, the following corollary generalizes the previous one:

COROLLARY 3. *Let $K \in \mathcal{K}^2$ and $L \in \mathcal{L}^2$ be given such that*

$$\frac{\lambda_1}{D(K)} + \frac{\mu_1}{r(K)} < k, \qquad k \in \mathbb{Z}.$$

Then, $G(K,L) \geq k^2$, i.e., $\{K + g \mid g \in L\}$ is, at least, a k^2-fold covering of E^2.

3.1. Proof of Theorem 3

Let $\{b_1, b_2\}$ be a reduced basis of L in the sense of Minkowski, with $\|b_1\| = \lambda_1(L)$, and let θ be the acute angle between b_1 and b_2 (so that $2\mu_1(L) = \|b_2\| \sin\theta$). Let $v_1 = b_1$, and let v_2 be a vector of length $2\mu_1$, which is perpendicular to v_1. Let now Γ denote the rectangular lattice determined by the basis $\{v_1, v_2\}$.

We reduce the problem to rectangular lattices and symmetric convex bodies. To this end, let K' be the Steiner symmetral of K with respect to the line $x = \lambda_1/2$. Then, K' is a convex domain with $A(K') = A(K)$, and $r(K') \geq r(K)$. We have to see that $G(K',\Gamma) = 0$. For, let us suppose that K' contains the Γ lattice point, $mv_1 + nv_2$, with m, $n \in \mathbb{Z}$. Then, the symmetry of K' about $x = \lambda_1/2$ assures that the line $y = 2\mu_1 n$ intersects K' in a line segment of length greater than λ_1. Thus, this line also intersects K in a line segment with the same length, which implies that $G(K,L) > 0$, contradicting the hypothesis. Hence, $G(K',\Gamma) = 0$. Now, we use an analogous argument but symmetrizing about the line $y = \mu_1$, and we obtain a convex set K^s with the same area, greater or equal inradius, and such that $G(K',\Gamma) = 0$.

Now, we may identify Γ with the rectangular lattice $\Gamma_{\delta\eta}$ generated by the vectors $(\delta, 0)$ and $(0,\eta)$ (note that either $\Gamma_{\delta\eta} = \Gamma$ or $\Gamma_{\delta\eta} = \rho_{\pi/2}(\Gamma)$, where $\rho_{\pi/2}$ is the rotation by $\pi/2$ at the origin). Then, applying Theorem 2 to K^s and $\Gamma_{\delta\eta} \equiv \Gamma$, we obtain finally

$$(2r(K) - \eta)A(K) \leq (2r(K^s) - \eta)A(K^s) \leq \begin{cases} \frac{1}{2}\delta\eta^2, & \text{if } \eta \geq 2\delta \\ b_{\delta\eta}(r_*) & \text{if } \eta < 2\delta. \end{cases}$$

This concludes the proof of the theorem. ∎

3.2. The inradius-diameter results

PROOF OF PROPOSITION 1. We will write $\omega(K) = \omega$, $r(K) = r$ and $D(K) = D$.

In [9] the following result is proved:

$$(\omega - 2\mu_1)(D - \lambda_1) \leq 2\mu_1\lambda_1 \tag{5}$$

with equality when and only when K is a triangle of diameter D and width $\omega = 2\mu_1 D/(D - \lambda_1)$.

Since $\omega \geq 2r$, we have

$$(2r - 2\mu_1)(D - \lambda_1) \leq (\omega - 2\mu_1)(D - \lambda_1) \leq 2\mu_1\lambda_1.$$

Taking the infinite strip to be the limit of a sequence of triangles which give the equality in (5), when ω tends to $2r$ we have

$$\lim_{\omega \to 2r} (2r - 2\mu_1)(D - \lambda_1) =$$

$$= \lim_{\omega \to 2r} (2r - 2\mu_1)\left(\frac{\omega\lambda_1}{\omega - 2\mu_1} - \lambda_1\right) = \lim_{2r \to 2\mu_1} 2r\lambda_1 = 2\mu_1\lambda_1.$$

So, the stated bound is best possible. ∎

PROOF OF COROLLARY 3. The proof of Corollary 3 follows an analogous proof by Hammer [4], and we repeat it here for opportunity.

If $k = 0$ the result is trivial. So, let us suppose that $k \geq 1$, and consider the similarity transformation $K \to K' = \frac{1}{k}K$. Obviously, $D(K') = \frac{1}{k}D(K)$ and $r(K') = \frac{1}{k}r(K)$. Now let $\{b_1, b_2\}$ be a basis of L with $\|b_i\| = \lambda_i$, $i = 1, 2$, and let $R = m_1 b_1 + m_2 b_2$ be a lattice point with $0 \leq m_i \leq (k-1)\lambda_i$, $i = 1, 2$.

Now, let us consider the translate K'' of K' given by $K'' = K' - \frac{1}{k}R$. We have:

$$\frac{\lambda_1}{D(K'')} + \frac{\mu_1}{r(K'')} = \frac{\lambda_1}{D(K')} + \frac{\mu_1}{r(K')} = \frac{k\lambda_1}{D(K)} + \frac{k\mu_1}{r(K)} =$$

$$= k\left(\frac{\lambda_1}{D(K)} + \frac{\mu_1}{r(K)}\right) < 1.$$

By Proposition 1, K'' contains a lattice point T. Hence, K' contains the point $T + \frac{1}{k}R$, and so, the original domain K contains the lattice point $U = k\left(T + \frac{1}{k}R\right) = kT + R$. But taking into account $0 \leq m_i \leq (k-1)\lambda_i$, we could have selected each m_i $i = 1, 2$ in k different ways. So, R might have been chosen in k^2 different ways. Therefore, K contains at least, k^2 distinct lattice points in its interior, i.e. $G(K, L) \geq k^2$.

Acknowledgements

The author wishes to thank Salvatore Vassallo for his valuable advices and helpful suggestions during the preparation of this paper. She also wishes to thank O. Pérez Blaya for enlightening discussions.

References

[1] W. Awyong, M. Henk and P. Scott, Note on lattice-point-free-convex bodies, *Mh. Math.,* **126** (1998), 7–12.

[2] W. Awyong and P. Scott, New inequalities for planar convex sets with lattice point constraints, *Bull. Austral. Math. Soc.,* **54** (1996), 391–396.

[3] T. Bonnesen and W. Fenchel, *Theorie der Konvexen Körper,* Springer, Berlin 1934, 1974; Chelsea, New York 1948.

[4] J. Hammer, Lattice Points and Area-Diameter Relation, *Math. Mag.,* **52** (1979), 25–25.

[5] R. Kannan and L. Lovász, Covering minima and lattice-points-free convex bodies, *Ann. Math.,* **128** (1988), 577–602.

[6] U. Schnell and J. M. Wills, Two Isoperimetric Inequalities with Lattice Constraints, *Mh. Math.,* **112** (1991), 227–233.

[7] U. Schnell and J. M. Wills, On Successive Minima and Intrinsic Volumes, *Mathematika,* **40** (1993), 144–147.

[8] S. Vassallo, A covering problem for plane lattices, *Geom. Ded.,* **43** (1992), 321–335.

[9] S. Vassallo, Area-diameter and area-width relations for covering plane sets, *Bull. Austral. Math. Soc.,* **51** (1995), 163–169.

[10] S. Vassallo and J. M. Wills, Covering sets for plane lattices, *Mh. Math.,* **121** (1996), 381–389.

ANNALES UNIV. SCI. BUDAPEST., **44** (2001), 107–113

A CHARACTERIZATION OF ADJOINT 1-SUMMING OPERATORS

By

QINGYING BU

Department of Mathematical Sciences, Kent State University, Kent, Ohio

(Received January 29, 2002)

For a Banach space X, let X^* denote its dual and B_X denote its closed unit ball. For $1 \leq p \leq \infty$, let p' denote its conjugate, i.e., $1/p + 1/p' = 1$. For $1 \leq p \leq \infty$, let $\ell_p(X)$ denote the space of absolutely p-summable sequences on a Banach space X, i.e.,

$$\ell_p(X) = \left\{ \overline{x} = (x_n)_n \in X^{\mathbb{N}} : \|\overline{x}\|_{(p)} = \left(\sum_{n=1}^{\infty} \|x_n\|^p \right)^{1/p} < \infty \right\},$$

where if $p = \infty$, let $\left(\sum_{n]1=}^{\infty} \|x_n\|^p \right)^{1/p} = \sup_n \|x_n\|$. Then $(\ell_p(X), \| \cdot \|_{(p)})$ is a Banach space (cf. [2, 7]). Let

$$c_0(X) = \left\{ \overline{x} = (x_n)_n \in X^{\mathbb{N}} : \lim_n x_n = 0 \right\}.$$

Then $c_0(X)$ is a closed subspace of $\ell_\infty(X)$. For $1 \leq p \leq \infty$, let $\ell_p[X]$ denote the space of weakly p-summable sequences on a Banach space X, i.e.,

$$\ell_p[X] = \left\{ \overline{x} = (x_n)_n \in X^{\mathbb{N}} : \sum_{n=1}^{\infty} |x^*(x_n)|^p < \infty \ \forall x^* \in X^* \right\}$$

and for $\forall \overline{x} \in \ell_p[X]$, let

$$\|\overline{x}\|_{[p]} = \sup \left\{ \left(\sum_{n=1}^{\infty} |x^*(x_n)|^p \right)^{1/p} : x^* \in B_{X^*} \right\}.$$

Then $(\ell_p[X], \|\cdot\|_{[p]})$ is a Banach space (cf. [2, 7, 8]). Here notice that if $p = \infty$, let $\left(\sum_{n=1}^{\infty} |x^*(x_n)|^p\right)^{1/p} = \sup_n |x^*(x_n)|$. For $1 \le p \le \infty$, let $\ell_p\langle X\rangle$ denote the space of strongly p-summable sequences on a Banach space X, i.e.,

$$\ell_p\langle X\rangle = \left\{\overline{x} = (x_n)_n \in X^{\mathbb{N}} : \sum_{n=1}^{\infty} |x_n^*(x_n)| < \infty \ \forall (x_n^*)_n \in \ell_{p'}[X^*]\right\}$$

and for $\forall \overline{x} \in \ell_p\langle X\rangle$, let

$$\|\overline{x}\|_{\langle p\rangle} = \sup\left\{\left|\sum_{n=1}^{\infty} x_n^*(x_n)\right| : (x_n^*)_n \in B_{\ell_{p'}[X^*]}\right\}.$$

Then $(\ell_p\langle X\rangle, \|\cdot\|_{\langle p\rangle})$ is a Banach space (cf. [2]). Let

$$c_0\langle X\rangle = \left\{\overline{x} = (x_n)_n \in \ell_\infty\langle X\rangle : \lim_n \|\overline{x}(i > n)\|_{\langle\infty\rangle} = 0\right\},$$

where $\overline{x}(i > n) = (0, \ldots, 0, x_{n+1}, x_{n+2}, \ldots)$. Then $c_0\langle X\rangle$ is a closed subspace of $\ell_p\langle X\rangle$.

DEFINITION 1. Let X, Y be Banach spaces. A Banach space operator $u : X \to Y$ is called strongly ∞-summing if there exists a constant $c > 0$ such that for any $x_1, x_2, \ldots, x_n \in X$ and $y_1^*, y_2^*, \ldots, y_n \in Y^*$,

$$\sum_{k=1}^{n} |\langle ux_k, y_k^*\rangle| \le c \cdot \sup_{1\le k\le n} \|x_k\| \cdot \sup_{y\in B_Y} \sum_{k=1}^{n} |y_k^*(y)|. \tag{1}$$

Let $D_\infty(u)$ denote the infimum taken over all possible c as above. Then $D_\infty(\cdot)$ is a norm (see [2]).

Recall that a Banach space operator $u : X \to Y$ is called absolutely 1-summing if there exists a constant $c > 0$ such that for any $x_1, x_2, \ldots, x_n \in X$,

$$\sum_{k=1}^{n} \|ux_k\| \le c \cdot \sup\left\{\sum_{k=1}^{n} |x^*(x_k)| : x^* \in B_X\right\}. \tag{2}$$

Let $\pi_1(\cdot)$ denote the absolutely 1-summing norm (see [4, p.31]). By Theorem 2.2.2 in [2], we have

THEOREM 2. (i) *A Banach space operator* $u : X \to Y$ *is absolutely* 1*-summing if and only if its adjoint operator* $u^* : Y^* \to X^*$ *is strongly* ∞*-summing. In this case,* $\pi_1(u) = D_\infty(u^*)$*;*

(ii) *A Banach space operator* $u : X \to Y$ *is strongly* ∞*-summing if and only if its adjoint operator* $u^* : Y^* \to X^*$ *is absolutely* 1*-summing. In this case,* $\pi_1(u^*) = D_\infty(u)$*.*

Let X, Y be Banach spaces. For a continuous linear operator $u : X \to \to Y$, define

$$\begin{array}{cccc} \hat{u} : & x^{\mathbb{N}} & \to & y^{\mathbb{N}}, \\ & (x_n)_n & \mapsto & (ux_n)_n. \end{array}$$

Then $\hat{u}$ is a linear operator. Define

$$\begin{array}{cccc} \psi : & x_0 \otimes X & \to & x^{\mathbb{N}}, \\ & \sum\limits_{k=1}^{n} s^k \otimes x_k & \mapsto & \left(\sum\limits_{k=1}^{n} s_i^{(k)} x_k \right)_i . \end{array}$$

Then ψ is well-defined linear map (see [1]). Let X, Y, Z, W be Banach spaces and let $u : X \to Z$ and $v : Y \to W$ be Banach space operators. Define

$$\begin{array}{cccc} u \otimes v : & X \otimes Y & \to & Z \otimes W, \\ & \sum\limits_{k=1}^{n} x_k \otimes y_k & \mapsto & \sum\limits_{k=1}^{n} (ux_k) \otimes (vy_k). \end{array}$$

Let $c_0 \overset{\wedge}{\otimes} X$ denote the completion of $c_0 \otimes X$ with respect to the injective tensor norm $\| \cdot \|_\vee$, and let $c_0 \overset{\vee}{\otimes} Y$ denote the completion of $c_0 \otimes Y$ with respect to the projective tensor norm $\| \cdot \|_\wedge$ (cf. [5, p. 223–227]).

THEOREM 3. *Let* $u : X \to Y$ *be a Banach space operator. Then the following are equivalent:*

(i) u *is strongly* ∞*-summing;*

(ii) $\mathrm{id}_{c_0}(\ell_\infty(X)) \subseteq \ell_\infty\langle Y \rangle$, *i.e.,* u *sends each bounded sequence in* X *to strongly* ∞*-summable sequence in* Y*;*

(iii) $\hat{u}(c_0(X)) \subseteq c_0\langle Y \rangle$*;*

(iv) $(\mathrm{id}_{c_0} \otimes u)(c_0 \overset{\vee}{\otimes} X) \subseteq c_0 \overset{\wedge}{\otimes} Y$, *where* id_{c_0} *is the identity operator on* c_0.

Furthermore, in this case, $\hat{u} : \ell_\infty(X) \to \ell_\infty\langle Y\rangle$, $\hat{u} : c_0(X) \to c_0\langle Y\rangle$, *and* $\mathrm{id}_{c_0} \otimes u : c_0 \overset{\vee}{\otimes} X \to c_0 \overset{\wedge}{\otimes} Y$ *are continuous with*

(3)
$$\| \hat{u} \|_{\ell_\infty(X)\to\ell_\infty\langle y\rangle} = \| \hat{u} \|_{c_0(X)\to c_0\langle Y\rangle} = \| \mathrm{id}_{c_0} \otimes u\|_{c_0 \overset{\vee}{\otimes} X\to c_0 \overset{\wedge}{\otimes} Y} = D_\infty(u).$$

PROOF. (i) $\Rightarrow$ (ii): It is easy to show that for $\forall\overline{x} = (x_n)_n \in \ell_\infty(X)$ and $\forall\overline{y^*} = (y_n^*)_n \in \ell_1[Y^*]$,

$$\sum_{n=1}^{\infty} |\langle ux_n, y_n^*\rangle| \le D_\infty(u)\cdot \|\overline{x}\|_{(\infty)} \cdot \|\overline{y^*}\|_{[1]}.$$

Since $\overline{y^*}$ is arbitrary in $\ell_1[Y^*]$, $\hat{u}(\overline{x}) = (ux_n)_n \in \ell_\infty\langle Y\rangle$ and

$$\| \hat{u} \|_{\ell_\infty(X)\to\ell_\infty\langle Y\rangle} \le D_\infty(u). \tag{4}$$

(ii) follows.

(ii) $\Rightarrow$ (iii): By Closed Graph Theorem, we can show that $\hat{u}$ is continuous. So

$$\| \hat{u}(\overline{x})\|_{\langle\infty\rangle} \le \| \hat{u} \|_{\ell_\infty(X)\to\ell_\infty\langle Y\rangle} \cdot \|\overline{x}\|_{(\infty)}, \quad \forall\overline{x} \in \ell_\infty(X). \tag{5}$$

Thus for $\forall n \in \mathbb{N}$,

$$\| \hat{u}(\overline{x})(i > n)\|_{\langle\infty\rangle} \le \| \hat{u} \|_{\ell_\infty(X)\to\ell_\infty\langle Y\rangle} \cdot \|\overline{x}(i > n)\|_{(\infty)},\ \forall\overline{x} \in \ell_\infty(X).$$

If $\overline{x} \in c_0(X)$, then $\lim_n \|\overline{x}(i > n)\|_{(\infty)} = 0$. So $\lim_n \| \hat{u}(\overline{x})(i > n)\|_{\langle\infty\rangle} = 0$, i.e., $\hat{u}(\overline{x}) \in c_0\langle X\rangle$. (iii) follows.

(iii) $\Rightarrow$ (i): By Closed Graph Theorem, we can show that $\hat{u}$ is continuous. So

$$\| \hat{u}(\overline{x})\|_{\langle\infty\rangle} \le \| \hat{u} \|_{c_0(X)\to c_0\langle Y\rangle} \cdot \|\overline{x}\|_{(\infty)}, \quad \forall\overline{x} \in c_0(X). \tag{5}$$

Now for any $x_1, \ldots, x_n \in X$, $y_1^*, \ldots, y_n^* \in Y^*$, let $\overline{x} = (x_1, \ldots, x_n, 0, 0, \ldots)$ and $\overline{y^*} = (s_1 y_1^*, \ldots, s_n, y_n^*, 0, 0, \ldots)$ where $s_k = \mathrm{sign}\langle ux_k, y_k^*\rangle$ for $k = 1, \ldots, n$. Then $\overline{x} \in c_0(X)$ and $y^* \in \ell_1[Y^*]$. So

$$\sum_{k=1}^{n} |\langle ux_k, y_k^*\rangle| = |\langle \hat{u}(\overline{x}), \overline{y^*}\rangle| \le \| \hat{u}(\overline{x})\|_{\langle\infty\rangle} \cdot \|\overline{y^*}\|_{[1]} \le$$

$$\leq \| \hat{u} \|_{c_0(X) \to c_0\langle Y \rangle} \cdot \|\overline{x}\|_{(\infty)} \cdot \|\overline{y^*}\|_{[1]} =$$

$$= \| \hat{u} \|_{c_0(X) \to c_0\langle Y \rangle} \cdot \sup_{1 \leq k \leq n} \|x_k\| \cdot \sup_{y \in B_Y} \sum_{k=1}^{n} |y_k^*(y)|.$$

Thus u is strongly ∞-summing and

$$D_\infty(y) \leq \| \hat{u} \|_{c_0(X) \to c_0\langle Y \rangle}. \tag{6}$$

(i) follows.

(iii) $\Rightarrow$ (iv): For $\forall z = \sum_{k=1}^{n} s^{(k)} \otimes x_k \in c_0 \otimes X$,

$$\begin{aligned} \psi((\mathrm{id}_{c_0} \otimes u)z) &= \psi\left(\sum_{k=1}^{n} s^{(k)} \otimes ux_k\right) = \left(\sum_{k=1}^{n} s^{(k)} ux_k\right)_i = \\ &= \left(u\left(\sum_{k=1}^{n} s_i^{(k)} x_k\right)\right)_i = \hat{u}\left(\psi\left(\sum_{k=1}^{n} s^{(k)} \otimes x_k\right)\right) = \hat{u}(\psi z). \end{aligned} \tag{7}$$

By [7] (also see [3]), $\psi(c_0 \overset{\vee}{\otimes} X) = c_0(X)$ with the isometry ψ. So by (5) and Theorem 9 in [1],

$$\|(\mathrm{id}_{c_0} \otimes u)z\|_\wedge = \|\psi((\mathrm{id}_{c_0} \otimes u)z)\|_{\langle\infty\rangle} = \| \hat{u}(\psi z)\|_{\langle\infty\rangle} \leq$$

$$\leq \| \hat{u} \|_{c_0(X) \to c_0\langle Y \rangle} \|\psi z\|_{(\infty)} = \| \hat{u} \|_{c_0(X) \to c_0\langle Y \rangle} \|z\|_\vee.$$

So $\mathrm{id}_{c_0} \otimes u$ is a continuous operator from $(c_0 \otimes X, \| \cdot \|_\vee)$ to $(x_0 \otimes Y, \| \cdot \|_\wedge)$. Thus $\mathrm{id}_{c_0} \otimes y$ can be norm-preserved extended to $c_0 \overset{\vee}{\otimes} X$. Therefore, $(\mathrm{id}_{c_0} \otimes u)(c_0 \overset{\vee}{\otimes} X) \subseteq c_0 \overset{\wedge}{\otimes} Y$ and

$$\| \mathrm{id}_{c_0} \otimes u \|_{c_0 \overset{\vee}{\otimes} X \to c_0 \overset{\wedge}{\otimes} Y} \leq \| \hat{u} \|_{c_0(X) \to c_0\langle Y \rangle}. \tag{8}$$

(iv) follows.

(iv) $\Rightarrow$ (iii): Let $\overline{x} \in c_0(X)$. Since $\psi(c_0 \overset{\vee}{\otimes} X) = c_0(X)$, $\exists z \in c_0 \overset{\vee}{\otimes} X$ such that $\psi(z)\overline{x}$. By (iv), $(\mathrm{id}_{c_0} \otimes u)_z \in c_0 \overset{\wedge}{\otimes} Y$. So by (7) and Theorem 9 in [1],

$$\hat{u}(\overline{x}) = \hat{u}(\psi z) = \psi((\mathrm{id}_{c_0} \otimes u)z) \in c_0\langle Y \rangle.$$

Thus (iii) follows. Furthermore,

$$\|\hat{u}(\overline{x})\|_{\langle\infty\rangle} = \|(\mathrm{id}_{c_0}\otimes u)z\|_{\vee} \leq$$
$$\leq \|\mathrm{id}_{c_0}\otimes u\|_{c_0\overset{\vee}{\otimes}X\to c_0\overset{\wedge}{\otimes}Y}\|z\|_{\vee} = \|\mathrm{id}_{c_0}\otimes u\|_{c_0\overset{\vee}{\otimes}X\to c_0\overset{\wedge}{\otimes}Y}\|\overline{x}\|_{(\infty)}.$$

Therefore,

$$\|\hat{u}\|_{c_0(X)\to c_0\langle Y\rangle} \leq \|\mathrm{id}_{x_0}\otimes u\|_{c_0\overset{\vee}{\otimes}X\to c_0\overset{\wedge}{\otimes}Y}. \tag{9}$$

Now combining (4), (6), (8), (9) and noticing that $\|\hat{u}\|_{c_0(X)\to c_0\langle Y\rangle} \leq \leq \|\hat{u}\|_{\ell_\infty(X)\to\ell_\infty\langle Y\rangle}$, (3) holds. The proof is completed. ∎

COROLLARY 4. (i) *A Banach space operator* $u : X \to Y$ *is absolutely* 1*-summing if and only if* $(\mathrm{id}_{c_0}\otimes u^*)(c_0\overset{\vee}{\otimes}Y^*) \subseteq c_0\overset{\wedge}{\otimes}X^*$. *In this case,* $\mathrm{id}_{c)}\otimes u^* : c_0\overset{\vee}{\otimes}Y^* \to c_0\overset{\wedge}{\otimes}X^*$ *is continuous and* $\|\mathrm{id}_{c_0}\otimes u^*\| = \pi_1(u)$*;*

(ii) *A Banach space operator* $u : X \to Y$ *satisfies that*

$$(\mathrm{id}_{c_0}\otimes u)(c_0\overset{\vee}{\otimes}X) \subseteq c_0\overset{\wedge}{\otimes}Y$$

if and only if its adjoint operator $u^* : Y^* \to X^*$ *is absolutely* 1*-summing. In this case,* $\mathrm{id}_{c_0}\otimes u : c_0\overset{\vee}{\otimes}X \to c_0\overset{\wedge}{\otimes}Y$ *is continuous and* $\|\mathrm{id}_{c_0}\otimes u\| = \pi_1(u^*)$.

ACKNOWLEDGEMENT. I am grateful to my advisor, Professor JOE DIESTEL, for his help.

References

[1] Q. BU and J. DIESTEL, Observations about the projective tensor product of Banach spaces, $I - \ell_p\overset{\wedge}{\otimes}X$, $1 < p <$, *Quaestiones Math.,* **24** (2001), 519–533.

[2] J. S. COHEN, Absolutely p-summing, p-nuclear operators, and their conjugates, *Math. Ann.,* **201** (1973), 177–200.

[3] J. DIESTEL, J. FOURIE and J. SWART, *The Metric Theory of Tensor Products,* to appear.

[4] J. DIESTEL, H. JARCHOW, and A. TONGE, *Absolutely Summing Operators,* Cambridge Univ. Press, Cambridge, 1995.

[5] J. DIESTEL and J. UHL, Vector Measures, Math. Surveys Vol. 15, *Amer. Math. Soc.,* Providence, 1977.

[6] A. GROTHENDIECK, Produits tensoriels topologiques et espaces nucléaires, *Memoires Amer. Math. Soc.,* **16** (1955).

[7] A. GROTHENDIECK, Résumé de la théorie métrique des produits tensoriels topologiques, *Bol. Soc. Mat. São Paulo,* **8** (1956), 1–79.

[8] M. GUPTA and Q. BU, On Banach-valued GAK-sequence spaces $\ell_p[X]$, *J. Analysis,* **2** (1994), 103–113.

ANNALES UNIV. SCI. BUDAPEST., **44** (2001), 115–121

SHORT PROOFS FOR THE PSEUDOINVERSE IN LINEAR ALGEBRA

By

CSABA J. HEGEDŰS

Department of Numerical Analysis, Eötvös Loránd University, Budapest

(Received February 25, 2002)

1. Introduction

We show how some theorems on projection matrices can be applied to give short proofs for the properties of the pseudoinverse and pseudosolution of under- or overdetermined linear system of equations. First, we recall some statements on projections and prove a maximum theorem for projections that map onto the same subspace. The symmetric projection is unique and the distance between a vector and a subspace can be given with the aid of a symmetric projection. As a by-product, one can give projections mapping onto the range space or null space of a matrix, and formulae for the distances from these subspaces.

On the basis of these results it is then possible to give short derivations for the pseudoinverse and its properties. With the suggested simplifications it will be easy to teach a short but full pseudoinverse theory in undergraduate or graduate courses. The given remarks may be used as exercises.

NOTATIONS. Matrices will be denoted by capital letters and vectors by lower case letters, a^{T} denotes the transpose of a. Real matrices and vectors are used throughout as it is straightforward to restate the results in complex variables.

2. Projections

We first recall some basic knowledge on projections and then prove our theorems. Let $S \subseteq \mathbf{R}^n$ be a subspace. $P \in \mathbf{R}^{n\times n}$ is a *projection* onto S if range$(P) = S$, and $P^2 = P$. Moreover, if $P^T = P$ then the projection is called *symmetric* or *orthogonal*. The only invertible projection is the identity I. To see this, multiply $P^2 = P$ with P^{-1}. Both P and $I - P$ are projections.

Only the first application of P may produce a new vector Px, all subsequent applications will leave it unchanged, $P^k x = Px$, $k > 1$. This property explains the phrase: P is *idempotent,* that is, any positive integer power of P is equal to itself.

Examples for projections:

$$P_1 = \frac{da^{\mathrm{T}}}{a^{\mathrm{T}}d}, \quad a^{\mathrm{T}}d \neq 0.$$

This is a rank one projection. The effect of $I - P_1$ on x is the following: vector x is projected along line d onto the plane $a^{\mathrm{T}}x = 0$: $(I - P_1)x = x - d\left(\frac{a^{\mathrm{T}}x}{a^{\mathrm{T}}d}\right)$ and $a^{\mathrm{T}}(I - P_1)x = 0$. Another example – generalizing the former – is

$$P_2 = I - A(D^{\mathrm{T}}A)^{-1}D^{\mathrm{T}}, \text{ where } D, A \in \mathbf{R}^{m\times n} \text{ and } D^{\mathrm{T}}A \text{ is invertible.}$$

Now P_2x is orthogonal to the column vectors of D. P_2 is *symmetric* or *orthogonal* if $D = A$ holds.

Observe that $P_1P_2 = P_2$ if P_1 and P_2 are mapping onto the same subspace.

THEOREM 1. *The symmetric projection is unique among projections mapping onto S.*

PROOF. Indirectly assume $P_1 \neq P_2$ are two symmetric projections onto S. Then

$$P_1P_2 = P_2 \;\Rightarrow\; P_2 = P_2^{\mathrm{T}} = P_2^{\mathrm{T}}P_1^{\mathrm{T}} = P_2P_1 = P_1,$$

a contradiction. (Cf. [4], Sect. 2.6.1.) ∎

THEOREM 2. *Let $H(S)$ be the set of projections that map onto the subspace S of $\mathbf{R}^m$. Moreover, let P_s be the unique symmetric projection onto S. Then for any vector $x \notin S$ and $Px \neq 0$*

$$\max_{P\in H(S)} \frac{x^TPx}{\|Px\|_2} = \|P_sx\|_2$$

holds. The maximum is also reached by any projections $\tilde{P}$ that satisfy $\tilde{P}x = = \lambda P_sx$, $\lambda > 0$.

PROOF. The theorem states that the angle between x and $Px \neq 0$ is smallest if $P = \tilde{P}$. If P_s, $P \in H(S)$ then $P_s P = P$, and applying Cauchy's inequality leads to

$$\frac{x^{\mathrm{T}} P x}{\|Px\|_2} = \frac{x^{\mathrm{T}} P_s P x}{\|Px\|_2} \leq \frac{\|P_s x\|_2 \, \|Px\|_2}{\|Px\|_2} = \|P_s x\|_2 \,,$$

where equality still holds for the projections $\tilde{P}$.

THEOREM 3. *The Euclidean distance of vector $x \notin S$ from subspace S is $\|(I - P_s)x\|_2$, where P_s is the symmetric projection onto S.*

PROOF. If P is a projection onto S then the distance of x from S is given by the minimum of $\|(I - \lambda P)x\|_2$ with respect to $\lambda > 0$ and P. We have seen in Theorem 2 that the smallest angle between x and λPx is reached for symmetric $P \in H(S)$ hence the minimum distance takes place between x and $\lambda P_s x$. But we have

$$\|x - \lambda P_s x\|_2^2 = \|x\|_2^2 - (2\lambda - \lambda^2)\, \|P_s x\|_2^2$$

from where it is seen by differentiation that the minimum with respect to $\lambda$is reached for $\lambda = 1$so that the distance vector $(I - P_s)x$ and $P_s x$ are mutually orthogonal to each other.

Now we recall two lemmas from standard linear algebra.

LEMMA 1. *Let L be a matrix of full column rank. If $LB = LC$ then $B = C$ follows.*

PROOF. We have $L(B - C) = 0$. Any linear combinations of the columns of L may result in zero only if the columns of $B - C$ are zero, thus $B =$ $= C$ follows.

LEMMA 2. *Assume $A \in R^{m\times n}$. Then $A^{\mathrm{T}}A$ is positive semidefinite. If A is of full column rank, then $A^{\mathrm{T}}A$ is positive definite.*

PROOF. Set $y = Ax$, then $x^{\mathrm{T}}A^{\mathrm{T}}Ax = y^{\mathrm{T}}y \geq 0$. For definiteness observe that$x = 0$follows from $y = 0$ if A is of full column rank.

EXAMPLES. (i) Let $m \geq n \geq k$, $A \in \mathbf{R}^{m\times n}$ and $A_1 \in \mathbf{R}^{m\times k}$ such that A_1 has the maximal number of linearly independent columns of A that is, $\mathrm{rank}(A) = k$ holds. Then for $B \in \mathbf{R}^{m\times k}$ and invertible $B^{\mathrm{T}}A_1$ define a projection by

$$P(A_1, B) = A_1(B^{\mathrm{T}}A_1)^{-1}B^{\mathrm{T}}.$$

Actually this is a mapping onto range(A_1)=range(A). It is orthogonal if $A_1 = B$, and then the inverse of $A^{\mathrm{T}} A_1$ exists by Lemma 2. The Euclidean distance of x from range(A) is given by $\|(I - P(A_1, A_1))x\|_2$.

(ii) Similarly, let $A_2 \in \mathbf{R}^{k\times n}$ have the maximal number of linearly independent rows of A. Then the projection onto range(A^{T}) is given by $P(A_2^{\mathrm{T}}, C^{\mathrm{T}})$ where CA_2^{T} is assumed to be invertible and $C \in \mathbf{R}^{k\times n}$. The projection onto nul(A) is $I - P(C^{\mathrm{T}}, A_2^{\mathrm{T}})$ and the Euclidean distance of y from nul(A) is $\left\|P(A_2^{\mathrm{T}}, A_2^{\mathrm{T}})y\right\|_2$.

3. Short theory of the pseudoinverse

There is a lot of books on the theoretical and computational aspects of the pseudoinverse. We mention here [1], [2], [4], [6]. These works, of course, go into much deeper details on generalized inverse theory than what will be addressed here. The computational aspects are handled e.g. in [4]. Reference 64 of [2] contains a bibliography of 1775 items on the theory, so the interested reader may consult them for details. Among Hungarian student texts, we mention [5] that gives a comprehensive theory of the pseudoinverse in Ch. 2. See also [3] and [7].

Now assume that we have a rank factorization of matrix $A \in \mathbf{R}^{m\times n}$, $A = LU$, where $L \in \mathbf{R}^{m\times r}$ and $U \in \mathbf{R}^{r\times n}$ so that rank(A) = r. We may think of an LU-factorization, but it is also possible to choose $L = Q$ and $U = R$ from the QR-factorization of A.

For the derivation of the pseudoinverse A^+ we start from the four defining Penrose equations:

$$\begin{array}{llll} 1. & AA^+A = A, & 2. & A^+AA^+ = A^+, \\ 3. & AA^+ = (AA^+)^{\mathrm{T}}, & 4. & A^+A = (A^+A)^{\mathrm{T}}. \end{array}$$

LEMMA 3. *AA^+ and A^+A are symmetric projections.*

PROOF. Multiply eq. 1 by A^+ from either side and observe equations 3 and 4. The same conclusions come from eq. 2 by multiplying with A.

THEOREM 4. *To every matrix A there exists uniquely the pseudoinverse $A^+ = U^+L^+$, where $L^+ = (L^{\mathrm{T}}L)^{-1}L^{\mathrm{T}}$, $U^+ = U^{\mathrm{T}}(UU^{\mathrm{T}})^{-1}$ and LU is an arbitrary rank factorization of A.*

PROOF. We begin with uniqueness. Assume indirectly that there are two different pseudoinverses, A_1^+ and A_2^+. By applying Lemma 3 and Theorem

1 on uniqueness, we have $A_1^+A = A_2^+A$ and $AA_1^+ = AA_2^+$ and that leads to contradiction:

$$A_1^+ = (A_1^+A)A_1^+ = A_2^+(AA_1^+) = A_2^+AA_2^+ = A_2^+.$$

The unique pseudoinverse will be given constructively. Observe that $\mathrm{range}(A) = \mathrm{range}(L)$ hence $AA^+ = LL^+ = L(L^\mathrm{T}L)^{-1}L^\mathrm{T}$ is the unique symmetric projection to there and $L^+ = (L^\mathrm{T}L)^{-1}L^\mathrm{T}$ follows by Lemma 1. Similarly, $\mathrm{range}(A^\mathrm{T}) = \mathrm{range}(U^\mathrm{T})$ holds so that the symmetric projection to there is $A^+A = A^\mathrm{T}(A^+)^\mathrm{T} = U^\mathrm{T}(U^+)^\mathrm{T} = U^+U \quad = U^\mathrm{T}(UU^\mathrm{T})^{-1}U$ from where one gets $U^+ = U^\mathrm{T}(UU^\mathrm{T})^{-1}$.

We have $L^+L = I_r$ and $UU^+ = I_r$ that is, L^+ is a left inverse and U^+ is a right inverse. With these

$$AA^+ = LL^+ = LUU^+L^+ \text{ and } A^+A = U^+U = U^+L^+LU,$$

from where one concludes that $A^+ = U^+L^+$.

REMARKS. If A has full column rank, then $L = A$ and $U = I_n$ is an appropriate choice, and $A^+ = (A^\mathrm{T}A)^{-1}A^\mathrm{T}$follows. With the QR-factorization of $A = QR$, one gets $A^+ = R^{-1}Q^\mathrm{T}$. If A has full row rank then $L = I_m$ and $U = A$ suffice, and then $A^+ = A^\mathrm{T}(AA^\mathrm{T})^{-1}$. If now $A^\mathrm{T} = QR$ then $A^+ = Q(R^\mathrm{T})^{-1}$. Finally, if A is rank deficient, then as a simple method, we take $A = Q_1BQ_2$ where Q_1, Q_2 are orthogonal matrices and B is an upper bidiagonal matrix. In that case $A^+ = Q_2^\mathrm{T}(B)^{-1}Q_1^\mathrm{T}$. Sometimes numerical rank determination is a delicate process, for details, see [4].

THEOREM 5. *Let P be a projection onto* $\mathrm{range}(A)$. *Then the linear system $Ax = b$ is consistent iff $Pb = b$.*

PROOF. Necessity. If the system is solvable then $b \in \mathrm{range}(A)$ and $Pb = b$ should hold. For sufficiency assume $P = UV^\mathrm{T}$ is a rank factorization of P. Then range(A)=range(U) because P is mapping onto range(A). Hence there exist a matrix Z such that $U = AZ$. Then

$$b = Pb = UV^\mathrm{T}b = AZV^\mathrm{T}b$$

such that $x = ZV^\mathrm{T}b$ is a solution.

REMARK. With the rank factorization of $P = UV^\mathrm{T}$, it is possible to give an explicit formula for Z. From $A = PA = UV^\mathrm{T}A$ it is seen that U and $B^\mathrm{T} = V^\mathrm{T}A$ gives a rank factorization of A, i.e. B^T is a full row rank matrix.

Now choose a matrix C such that $B^{\mathrm{T}}C$ is invertible (for example, $C^{\mathrm{T}} = B^{\mathrm{T}}$ suffices). Then

$$AC(B^{\mathrm{T}}C)^{-1} = UB^{\mathrm{T}}C(B^{\mathrm{T}}C)^{-1} = U$$

from where $Z = C(B^{\mathrm{T}}C)^{-1}$.

The pseudoinverse helps us to decide if a linear system is solvable. The consistency condition $b = AA^{+}b$ yields the solution $x^{+} = A^{+}b$ immediately.

THEOREM 6. *Assume the linear system $Ax = b$ is consistent. Then the general solution is given by*

$$x_g = x_p + (I - A^{+}A)t,\ t \in \mathbf{R}^n$$

where x_p is a particular solution and $(I - A^{+}A)t$ is the general solution of the homogeneous system $Ax = 0$.

PROOF. Assume x_1, x_2 are two solutions. Then $A(x_2 - x_1) = 0$ shows that the difference of two solutions is a solution of the homogeneous system and those solutions are in null(A). The pseudosolution x^{+} may serve as a particular solution.

If the linear system is inconsistent, then we can make it consistent by orthogonally projecting b onto range(A):

$$AA^{+}Ax = Ax = AA^{+}b.$$

The pseudosolution again is $x^{+} = A^{+}b$.

THEOREM 7. *The pseudosolution has the following properties: $\|b - Ax\|_2$ is minimal for $x = x^{+}$. $\|x^{+}\|_2$ is minimal among the possible solutions.*

PROOF. $\|b - Ax^{+}\|_2 = \|b - AA^{+}b\|_2$ is nothing else than the distance of vector b from range(A) by Theorem 3. The general solution for both cases (consistent or inconsistent systems) is expressible by

$$x_g = x^{+} + (I - A^{+}A)t = A^{+}AA^{+}b + (I - A^{+}A)t$$

so that it is the sum of two orthogonal vectors. Hence $\|x_g\|_2^2 = \|A^{+}b\|_2^2 + \|(I - A^{+}A)t\|_2^2$ which is minimal if $t = 0$.

REMARKS. Demanding either the first or the second Penrose condition is enough for AA^g or A^gA to be a projection, where A^g denotes a generalized inverse. If the first and third Penrose equation is fulfilled then A^gb is a least squares solution. In case of a consistent system the fulfilment of the second and fourth condition is enough to get a minimum norm solution.

ACKNOWLEDGEMENTS. The author thanks Dr. GISBERT STOYAN for his careful reading of the manuscript and his critical comments. Also, thanks are due to Dr. AURÉL GALÁNTAI and Dr. LAJOS LÁSZLÓ for their remarks.

References

[1] BEN ISRAEL, A. and GREVILLE, T. N. E., *Generalized Inverses: Theory and Applications,* John Wiley & Sons, New York, 1974.

[2] CAMPBELL, S. L. and MEYER JR., C. D., *Generalized Inverses of Linear Transformations,* Pitman, London, 1979.

[3] GALÁNTAI, A. and JENEY A., *Numerical Methods. Univ. Text.* (in Hungarian), Miskolc University Press, Miskolc, 1998.

[4] GOLUB, G. H. and VAN LOAN, C., *Matrix Computations,* 2^{nd} edition. John Hopkins Univ. P. Baltimore and London, 1989.

[5] MÓRICZ, F., *Numerical Methods in Algebra and Analysis* (in Hungarian), Polygon, Szeged, 1997.

[6] RAO, C. R. and MITRA S. K., *Generalized Inverse of Matrices and its Applications,* Wiley and Sons, Inc. New York, 1971.

[7] STOYAN, G. and TAKÓ, G., *Numerical Methods I.* 2^{nd} *Ed.* (in Hungarian), Sec. 2.3 and Sec. 2.6.3., Typotex, Budapest, 2002.

ANNALES UNIV. SCI. BUDAPEST., **44** (2001), 123–128

INFINITELY MANY BI-IDEALS WHICH ARE NOT QUASI-IDEALS IN SOME TRANSFORMATION SEMIGROUPS

By

YUPAPORN KEMPRASIT

Department of Mathematics, Faculty of Science, Chulalongkorn Univ., Bangkok, Thailand

(Received December 28, 2001)

1. Introduction and Preliminaries

A subsemigroup Q of a semigroup S is called a quasi-ideal of S if $SQ \cap QS \subseteq Q$, and by a *bi-ideal* of S we mean a subsemigroup B of S such that $BSB \subseteq B$. Quasi-ideals are a generalization of left ideals and right ideals and bi-ideals are a generalization of quasi-ideals. The notion of quasi-ideal was first introduced by O. STEINFELD in [7]. In fact, the notion of bi-ideal was given earlier. This can be seen in [3] and [2], page 84.

For a nonempty subset A of a semigroup S, $(A)_q$ and $(A)_b$ denote the quasi-ideal and the bi-ideal of S generated by A, respectively, that is, $(A)_q$ is the intersection of all quasi-ideals of S containing A and $(A)_b$ is the intersection of all bi-ideals of S containing A ([8], page 10 and 12).

PROPOSITION 1.1. ([2], page 84–85) *For any nonempty subset A of S,*

$$(A)_q = S^1 A \cap AS^1 \qquad \text{and} \qquad (A)_b = AS^1 A \cup A.$$

Let $\boldsymbol{BQ}$ denote the class of all semigroups whose sets of bi-ideals and quasi-ideals coincide. It is known that the following semigroups belong to $\boldsymbol{BQ}$: regular semigroups ([6]), left[right] simple semigroups ([4]) and left[right] 0-simple semigroups ([4]). Not only these semigroups are in $\boldsymbol{BQ}$. A nontrivial zero semigroup is an obvious example. In fact, J. CALAIS [1] has characterized the semigroups in $\boldsymbol{BQ}$ as follows: A semigroup S is in $\boldsymbol{BQ}$ if and only if $(x, y)_q = (x, y)_b$ for all x, $y \in S$. It is not easy to see from this characterization whether a given semigroup belongs to $\boldsymbol{BQ}$. Since every quasi-ideal of a semi-group S is a bi-ideal, it follows that $(x)_b \subseteq (x)_q$ for every $x \in S$. Therefore we have

PROPOSITION 1.2. *For an element x of a semigroup S, if $(x)_b$ is a quasi-ideal of S, then $(x)_b = (x)_q$.*

Let X be a nonempty set. It is well-known that the partial transformation semigroup on X, the full transformation semigroup on X and the one-to-one partial transformation semigroup on X (the symmetric inverse semigroup on X) are all regular, so they all belong to ***BQ*** ([6]). Let T_X denote the full transformation semigroup on X. The author has shown in [5] that transformation semigroup $\{a \in T_X \mid X \setminus X\alpha$ is infinite$\}$ where X is infinite, is not regular and neither left simple nor right simple but always belongs to ***BQ***. Let G_X, M_X and E_X denote respectively the symmetric group on X, the semigroup of all one-to-one transformations of X and the semigroup of all onto transformations of X. Then $G_X \in$ ***BQ***. For $a \in T_X$, α is said to be *one-to-one* at $x \in X$ if $(x\alpha)\alpha^{-1} = \{x\}$ and let $C(\alpha)$ be the set of all $x \in X$ such that α is not one-to-one at x. A transformation $\alpha \in T_X$ is said to be *almost one-to-one* if $C(\alpha)$ is finite. Hence if $\alpha \in T_X$ is almost one-to-one, then for every $x \in X$, $(x\alpha)\alpha^{-1}$ is finite. By an *almost onto transformation* of X we mean $\alpha \in T_X$ whose $X \setminus X\alpha$ is finite. Let AM_X and AE_X denote the set of all almost one-to-one transformations of X and the set of all almost onto transformations of X, respectively. Note that $G_X \subseteq M_X \subseteq AM_X$ and $G_X \subseteq E_X \subseteq AE_X$, and if X is finite, then $M_X = E_X = G_X$ and $AM_X = AE_X = T_X$. Also, we have

PROPOSITION 1.3. *AM_X and AE_X are subsemigroups of T_X.*

PROOF. Let $\alpha, \beta \in AM_X$ and $x \in X \setminus (C(\alpha) \cup C(\beta)\alpha^{-1})$. Then $x \in X \setminus C(\alpha)$ and $x\alpha \in X \setminus C(\beta)$. Thus we have $(x\alpha)\alpha^{-1} = \{x\}$ and $(x\alpha\beta)\beta^{-1} = = \{x\alpha\}$, and hence $(x\alpha\beta)(\alpha\beta)^{-1} = (x\alpha\beta)\beta^{-1}\alpha^{-1} = \{x\}$. Therefore $x \in \in X \setminus C(\alpha\beta)$. This proves that $X \setminus (C(\alpha) \cup C(\beta)\alpha^{-1}) \subseteq X \setminus C(\alpha\beta)$, so $C(\alpha\beta) \subseteq C(\alpha) \cup C(\beta)\alpha^{-1} = C(\alpha) \cup (C(\beta) \cap X\alpha)\alpha^{-1}$. Since α and β are almost one-to-one, $C(\alpha) \cup (C(\beta) \cap X\alpha)\alpha^{-1}$ is finite. Thus $\alpha\beta \in AM_X$.

Since for $\alpha, \beta \in T_X$,

$$X \setminus X\alpha\beta = (X \setminus X\beta) \cup (X\beta \setminus X\alpha\beta) \subseteq (X \setminus X\beta) \cup (X \setminus X\alpha),$$

it follows that AE_X is a subsemigroup of T_X. ∎

As was mentioned above, if X is finite, then M_X, E_X, AM_X and AE_X belong to ***BQ***. The aim of this paper is to show that if X is infinite, then all M_X, E_X, AM_X and AE_X contain at least k bi-ideals which are not quasi-ideals where $k = |X|$.

2. Main Results

We start by proving the following result for M_X and AM_X.

THEOREM 2.1. *If X is an infinite set, then the cardinality of the set of all bi-ideals in M_X which are not quasi-ideals and the cardinality of the set of all bi-ideals in AM_X which are not quasi-ideals are at least $|X|$.*

PROOF. Assume that S_X is M_X or AM_X. Since X is infinite, then $|X \times \mathbb{N}| = |X|$ where $\mathbb{N}$ denotes the set of all positive integers. Therefore there is a bijection from $X \times \mathbb{N}$ onto X and for clarity in what follows, we write its images as $s(t,n)$ for $t \in X$ and $n \in \mathbb{N}$. For each $t \in X$, let

$$A_t = s(t \times \mathbb{N})$$

and note that $\{A_t \mid t \in X\}$ forms a partition of X. For each $t \in X$, define $\alpha_t : X \to X$ by

$$x\alpha_t = \begin{cases} s(t,2n), & \text{if } x = s(t,n) \text{ for some } n \in \mathbb{N}, \\ \text{x} & \text{otherwise} \end{cases}$$

and note that $\alpha_t \in M_X$ and

$$X\alpha_t = X \setminus \{s(t,n) \mid n \text{ is odd}\}.$$

Clearly, $X\alpha_t \neq X\alpha_{t'}$ for distinct t, $t' \in X$. Now let $B_t = (\alpha_t)_b$, the bi-ideal of S_X generated by α_t. Then by Proposition 1.1, $B_t = \alpha_t S_X \alpha_t \cup \{\alpha_t\}$ and note that $X\lambda \subseteq X\alpha_t$ for all $\lambda \in B_t$. Thus if $B_t = B_{t'}$, then $X\alpha_t = X\alpha_{t'}$ and hence $t = t'$. Therefore $B_t \neq B_{t'}$, for distinct t, $t' \in X$. Thus $|\{B_t \mid t \in X\}| = |X|$. We assert that no B_t is a quasi-ideal of S_X. To show this by Proposition 1.2, that is, to show that $B_t \neq (\alpha_t)_q$, fix $t \in X$ and define $\beta, \gamma : X \to X$ by

$$x\beta = \begin{cases} s(t,n+1), & \text{if } x = s(t,n) \text{ for some } n \in \mathbb{N}, \\ x & \text{otherwise,} \end{cases}$$

$$x\gamma = \begin{cases} s(t,n+2) & \text{if } x = s(t,n) \text{ for some } n \in \mathbb{N}, \\ x & \text{otherwise.} \end{cases}$$

Then $\beta, \gamma \in M_X$ and

$$\begin{aligned} s(t,n)\beta\alpha_t &= s(t,2n+2) = s(t,n)\alpha_t y && \text{for all } n \in \mathbb{N}, \\ x\beta\alpha_t &= x = x\alpha_t\gamma && \text{for all } x \in X \setminus A_t. \end{aligned}$$

Therefore $\alpha_t \neq \beta\alpha_t = \alpha_t\gamma \in S_X\alpha_t \cap \alpha_t S_X = (\alpha_t)_q$. If $\beta\alpha_t \in B_t$, then $\beta\alpha_t = \alpha_t\eta\alpha_t$ for some $\eta \in S_X$ and hence $\beta = \alpha_t\eta$ since α_t is one-to-one.

Thus $C(\eta) = \emptyset$ if $S_X = M_X$ and $C(\eta)$ is finite if $S_X = AM_X$. By the definitions of α_t and β, we have

$$X \setminus \{s(t,1)\} = X\beta = X\alpha_t\eta = (X \setminus \{s(t,n) \mid n \text{ is odd}\})\eta. \tag{2.1.1}$$

Since $|s(t,1)\eta^{-1}| \leq 1$ if $S_X = M_X$ and $s(t,1)\eta^{-1}$ is finite if $S_X = AM_X$, it follows that $\{s(t,n) \mid n \text{ is odd}\} \setminus s(t,1)\eta^{-1}$ is an infinite set. From (2.1.1), we have

$$(\{s(t,n) \mid n \text{ is odd}\} \setminus s(t,1)\eta^{-1})\eta \subseteq (X \setminus \{s(t,n) \mid n \text{ is odd}\})\eta.$$

Consequently, $\{s(t,n) \mid n \text{ is odd}\} \setminus s(t,1)\eta^{-1} \subseteq C(\eta)$ which is a contradiction. Hence $\beta\alpha_t \in B_t$. That is, $\{B_t \mid t \in X\}$ is a family of bi-ideals in S_X as required by the theorem. ∎

From Theorem 2.1 and the fact that $M_X = G_X$ and $AM_X = T_X$ if X is finite, the following corollary is obtained

COROLLARY 2.2. *Let S_X be M_X or AM_X. Then $S_X \in$ **BQ** if and only if X is finite.*

THEOREM 2.3. *If X is an infinite set, then the cardinality of the set of all bi-ideals in E_X which are not quasi-ideals and the cardinality of the set of all bi-ideals in AE_X which are not quasi-ideals are at least $|X|$.*

PROOF. Assume that S_X be E_X or AE_X. For $t \in X$ and $n \in \mathbb{N}$, let A_t and $s(t,n)$ be defined as in the proof of Theorem 2.1. Next, for $t \in X$, define $\alpha_t : X \to X$ by

$$x\alpha_t = \begin{cases} s\left(t, \frac{n}{2}\right) & \text{if } x = s(t,n) \text{ for some even } n \in \mathbb{N}, \\ s(t,1) & \text{if } x = s(t,n) \text{ for some odd } n \in \mathbb{N}, \\ x & \text{otherwise.} \end{cases}$$

Then $\alpha_t \in E_X$ for every $t \in X$ and $\alpha_t \neq \alpha_{t'}$ for distinct t, $t' \in X$ (since s is one-to-one). Now, let $B_t = (\alpha_t)_b$, the bi-ideal of S_X generated by α_t. Then $B_t = \alpha_t S_X \alpha_t \cup \{\alpha_t\}$ by Proposition 1.1. Observe that $B_t \neq B_{t'}$ for distinct $t, t' \in X$. For, if not, then $\alpha_{t'} = \alpha_t\lambda\alpha_t$ for some $\lambda \in S_X$. Since $x\alpha_{t'} = x$ for all $x \in A_t$, we have that

$$\begin{aligned}(s(t,1)\lambda)\alpha_t &= (\{s(t,n) \mid n \text{ is odd}\})\alpha_t\lambda\alpha_t = \\ &= (\{s(t,n) \mid n \text{ is odd}\})\alpha_{t'} = \{s(t,n) \mid n \text{ is odd}\}\end{aligned}$$

which is impossible because λ and α_t are functions. Hence $|\{B_t \mid t \in X\}| = |X|$. We assert that no B_t is a quasi-ideal of S_X. To see this by Proposition 1.2, that is, $B_t \neq (\alpha_t)_q$, fix $t \in X$ and define $\beta, \gamma : X \to X$ by

$$x\beta = \begin{cases} s(t, n-2) & \text{if } x = s(t,n) \text{ and } n \in \mathbb{N} \setminus \{1,2\}, \\ s(t,1) & \text{if } x = s(t,2), \\ x & \text{otherwise,} \end{cases}$$

$$x\gamma = \begin{cases} s(t, n-1) & \text{if } x = s(t,n) \text{ and } n \in \mathbb{N} \setminus \{1\}, \\ x & \text{otherwise.} \end{cases}$$

Then $\beta, \gamma \in E_X$ and

$$\begin{aligned} x\beta\alpha_t &= x = x\alpha_t\gamma && \text{for all } x \in X \setminus A_t, \\ s(t,n)\beta\alpha_t &= s(t,1) = s(t,n)\alpha_t\gamma && \text{for } n \in \{1,2,3\}, \\ s(t,n)\beta\alpha_t &= s\left(t, \frac{n-2}{2}\right) = s(t,n)\alpha_t\gamma && \text{for even } n \in \mathbb{N}\setminus\{1,2,3\}, \\ s(t,n)\beta\alpha_t &= s(t,1) = s(t,n)\alpha_t\gamma && \text{for odd } n \in \mathbb{N}\setminus\{1,2,3\}, \end{aligned}$$

Consequently, $\alpha_t \neq \beta\alpha_t = \alpha_t\gamma$ and hence $\alpha_t\gamma \in S_X\alpha_t \cap \alpha_t S_X = (\alpha_t)_q$. If $\alpha_t\gamma \in B_t$, then $\alpha\gamma = \alpha_t\eta\alpha_t$ for some $\eta \in S_X$, and hence $\gamma = \eta\alpha_t$ since α_t is onto. It follows that η must fix $X \setminus A_t$ pointwise. In addition, since

$$(A_t \setminus \{s(t,1), s(t,2)\})\eta\alpha_t = (A_t \setminus \{s(t,1), s(t,2)\})\gamma = A_t \setminus \{s(t,1)\},$$

it follows from the definition of α_t that

$$(A_t \setminus \{s(t,1), s(t,2)\})\eta = \{s(t,n) \mid n \text{ is even and } n > 2\}.$$

Therefore we have

(2.3.1) $(X \setminus \{s(t,1), s(t,2)\})\eta = (X \setminus A_t) \cup \{s(t,n) \mid n \text{ is even and } n > 2\}.$

Since $\eta \in S_X$, we have that $X \setminus X\eta = \emptyset$ if $S_X = E_X$ and $X \setminus X\eta$ is finite if $S_X = AE_X$. But we obtain from (2.3.1) that

$$X \setminus X\eta = (\{s(n,t) \mid n \text{ is odd}\} \cup \{s(2,t)\}) \setminus \{s(t,1), s(t,2)\}\eta$$

which is an infinite set since η is a function, so we have a contradiction. Hence $\gamma\alpha_t \notin B_t$. That is, $\{B_t \mid t \in X\}$ is a family of bi-ideals in S_X as required by the theorem. ∎

From Theorem 2.3 and the fact that $E_X = G_X$ and $AE_X = T_X$ if X is finite, we have

COROLLARY 2.4. *Let S_X be E_X or AE_X. Then $S_X \in BQ$ if and only if X is finite.*

References

[1] J. Calais, Demi-groups quasi-inversif, *C.R. Acad. Paris,* **252** (1961), 1357–2359.

[2] A. H. Clifford and G. B. Preston, *The Algebraic Theory of Semigroups,* Vol. I., Amer. Math. Soc., (Providence, 1961).

[3] R. A. Good and D. R. Hughes, Associated groups for a semigroup, *Bull. Amer. Math. Soc.,* **58** (1952), 624–625 (Abstract).

[4] K. M. Kapp, On bi-ideals and quasi-ideals in semigroups, *Publ. Math. Debrecen,* **16** (1969), 179–185.

[5] Y. Kemprasit, Some transformation semigroups whose sets of bi-ideals and quasi-ideals coincide, *Comm. Algebra,* to appear.

[6] S. Lajos, Generalized ideals in semigroups, *Acta. Sci. Math.,* **22** (1961), 217–222.

[7] O. Steinfeld, Über die quasiideale von halbgruppen, *Publ. Math. Debrecen,* **4** 1956), 262–275.

[8] O. Steinfeld, Quasi-ideals in Rings and Semigroups, Akadémiai Kiadó Budapest, (1978).

ANNALES UNIV. SCI. BUDAPEST., **44** (2001), 129–150

L-COMMUTATIVITY OF THE OPERATORS IN SPLITTING METHODS FOR AIR POLLUTION MODELS

By

IVAN DIMOV, ISTVÁN FARAGÓ, ÁGNES HAVASI, ZAHARI ZLATEV

Central Laboratory of Parallel processing, Bulgarian Academy of Sciences, Sofia, Bulgaria
Eötvös Loránd University, Budapest
Eötvös Loránd University, Budapest
National Environmental Research Institute, Department of Atmospheric Environment,
Roskilde, Denmark

(Received March 13, 2002)

1. Introduction

Splitting procedures can lead to a substantial reduction of the computational work when large-scale problems are to be treated. Therefore, such procedures are often used in the numerical solution of many boundary value problems for differential equations describing real-life processes, see [3, 9, 5, 11, 16, 21, 13, 18]. A detailed theoretical study and analysis of the splitting procedure can be found in [8, 10].

An important example of real-life modelling is the problem of large-scale air pollution transport. Mathematical models of this kind are usually presented as a system of three-dimensional time-dependent partial differential equations which describe the processes of advection, diffusion, deposition, pollutant emission sources and chemical reactions. The environmental problems are becoming more and more important for the modern society, and their importance will certainly increase in the near future. High pollution levels (high concentrations and/or depositions of certain chemical species) may cause damages to plants, animals and humans. Moreover, some ecosystems can also be damaged (or even destroyed) when the pollution levels are very high. This is why the pollution levels must be carefully studied and controlled in the efforts to make it possible (i) to predict the appearance of high pollution levels and/or (ii) to decide what can be done to prevent the exceedance of prescribed critical levels. The control of the pollution levels in different highly developed and densely populated regions of Europe and North America is an important task for the modern society. Its importance has been steadily increasing during the last two decades. The necessity to establish reliable control strategies for air pollution levels will become even more important in the next two-three decades. Large scale air pollution models can be used to design reliable control strategies.

The numerical treatment of such mathematical models includes operator or time splitting. This procedure has several advantages: 1. the obtained sub-systems are easier to treat numerically than the original system; 2. we can exploit the special properties of the different sub-systems and apply the most suitable numerical method for each; 3. if each numerical method preserves the main qualitative properties then so does the global model. It is known that splitting procedures work well in the computer treatment of many air pollution models, [22, 12, 23]. At the same time, little attention has been devoted to the analysis of splitting procedures used in practice and to the question why splitting usually leads to good results.

As it was mentioned above, splitting procedures are used in order to facilitate the choice of efficient numerical methods in the treatment of the different operators involved in the model under consideration. Assume that the selected methods are not only efficient, but also sufficiently accurate. Then the success of splitting is determined by the splitting error. The recent paper [5] presents an analysis of operator splitting in air pollution models. By using the Lie operator formalism, a general expression is derived for a three-term splitting procedure in the pure initial value case. The procedure is called Strang splitting procedure in [5], however it has been introduced independently in 1968 by [6] and [16]. Therefore, it is reasonable to call it Marchuk–Strang splitting procedure. The splitting error for the advection–diffusion–reaction problem is analysed in the above mentioned reference [5]. Different conditions for reducing the errors, which are caused by the splitting procedures, are discussed there. Sufficient conditions, for which the splitting errors vanish, are also derived. These conditions are too strong and, thus, rather unrealistic when large models are used to treat practical problems. For example, it is obtained that if the velocity field $\mathbf{u}$ and diffusion matrix $\mathbf{k}$ are independent of the space coordinates $\mathbf{x}$, then there is no splitting error when *advection* and *diffusion* are splitted. The recent work [20] presents numerical methods, which are proposed for several splitted problems.

The *splitting errors* in the numerical treatment of the splitted problem are closely related to the requirement for L-commutativity of the operators involved in the splitting procedure. More precisely, the errors due to the application of splitting procedures disappear when the corresponding operators L-commute. This is why the L-commutativity of operators will be a major tool in the derivation of the results.

The goals of this work are:

- to analyse the L-commutativity of operators used in the mathematical model for studying large-scale air pollution transport,
- to formulate conditions under which the splitting errors vanish,
- to investigate the splitting errors of two widely used splitting procedures.

The paper is organized as follows. In Section 2 the definitions of the commutator operator and the L-commutativity are given. In Section 3 we define the operators associated with processes of *advection, diffusion, deposition, emission* and *chemistry.* The necessary and sufficient conditions for the L-commutativity of the operators are studied in Section 4. In Section 5 we introduce two different splitting procedures: the splitting procedure based on the separation of the physical processes involved in the Danish Eulerian Model (it will be called the DEM splitting in this paper, but this is done only in order to facilitate the references to it and, at the same time, to keep in mind that this procedure is used in a particular air pollution model) and the Marchuk–Strang splitting. It is also shown how the Lie operator formalism can be used to analyse the structure of the splitting error. Some concluding remarks are given in Section 6.

2. Background definitions

Throughout this paper we use the following notations. Let S denote some normed space of functions of type $\mathbb{R}^4 \to \mathbb{R}^m$ with the variables $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ and $t \in \mathbb{R}_0^+$. Clearly, any element $\mathbf{f}(\mathbf{x}, t) \in \mathbf{S}$ can be identified with the set of functions $f_l(\mathbf{x}, t) \in \mathbf{T}$, $l = 1, 2, \ldots, m$, where the notation $\mathbf{T}$ stands for the set of mappings of type $\mathbb{R}^4 \to \mathbb{R}$. The notation $\mathbf{S}_{\text{lin}}$ will be used for the linear functions in $\mathbf{S}$.

Assume that $\mathbf{A} : \mathbf{S} \to \mathbf{S}$ is a given operator. Such an operator can be identified with m operators of type $\mathbf{S} \to \mathbf{T}$, called components of the operator $\mathbf{A}$. We always assume that the operators $\mathbf{A} : \mathbf{S} \to \mathbf{S}$ are differentiable in Frechet sense [14] and the derivative operator is denoted by $\mathbf{A}'$. In the sequel $\mathbf{g} \in \mathbf{S}$, $\sigma_1(\mathbf{x}, t), \sigma_2(\mathbf{x}, t), \ldots, \sigma_m(\mathbf{x}, t) \in \mathbf{T}$, $k_1(\mathbf{x}, t), k_2(\mathbf{x}, t), k_3(\mathbf{x}, t) \in \mathbf{T}$, $\mathbf{k}(\mathbf{x}, t)$ is a mapping of type $\mathbb{R}^4 \to \mathbb{R}^{3\times 3}$ and has the form of a diagonal matrix

$$\mathbf{k}(\mathbf{x}, t) = \operatorname{diag}(k_1(\mathbf{x}, t), k_2(\mathbf{x}, t), k_3(\mathbf{x}, t)).$$

The functions $u_1(\mathbf{x}, t)$, $u_2(\mathbf{x}, t)$, $u_3(\mathbf{x}, t) \in \mathbf{T}$ define a vector field

$$\mathbf{u}(\mathbf{x}, t) = (u_1(\mathbf{x}, t), u_2(\mathbf{x}, t), u_3(\mathbf{x}, t))$$

of type $\mathbb{R}^4 \to \mathbb{R}^3$.

Let $\mathbf{p} \in \mathbb{R}^m$. The functions $R_l(\mathbf{x}, \mathbf{p})$, $l = 1, 2, \ldots, m$ are mappings of type $\mathbb{R}^{m+3} \to \mathbb{R}$, therefore the mapping

$$\mathbf{R}(\mathbf{x}, \mathbf{p}) = (R_1(\mathbf{x}, \mathbf{p}), R_2(\mathbf{x}, \mathbf{p}), \ldots, R_m(\mathbf{x}, \mathbf{p}))$$

is a mapping of type $\mathbb{R}^{m+3} \to \mathbb{R}^m$.

As usual, for the scalar valued function $f(\mathbf{x}, t) \in \mathbf{T}$ the notation $\partial_i f$ means the partial derivative w.r.t. the i-th component. The differential operator ∇ will be applied also in the usual way. It acts always with respect the space variables x_1, x_2 and x_3. That is, for a scalar valued function $f(\mathbf{x}, t) \in \mathbf{T}$ the symbol ∇f means the gradient operator w.r.t. $\mathbf{x}$ in the sense

$$\nabla f(\mathbf{x}, t) = (\partial_1 f(\mathbf{x}, t), \partial_2 f(\mathbf{x}, t), \partial_3 f(\mathbf{x}, t)).$$

For a three-dimensional vector field $\mathbf{f}(\mathbf{x}, t) = (f_1(\mathbf{x}, t), f_2(\mathbf{x}, t), f_3(\mathbf{x}, t))$ the symbol $\nabla \cdot f$ yields the divergence operator that is

$$\nabla \cdot \mathbf{f}(\mathbf{x}, t) = \partial_1 f_1(\mathbf{x}, t) + \partial_2 f_2(\mathbf{x}, t) + \partial_3 f_3(\mathbf{x}, t).$$

We remark that for an elements $\mathbf{f} \in \mathbf{S}$ the ∇ operator acts componentwise, that is $\nabla \mathbf{f} \in \mathbf{S}$ and $(\nabla \mathbf{f})_l = \nabla(f_l)$, $l = 1, 2, \ldots, m$. The same notation is used for the Laplace operator $\Delta = \nabla^2$. The use of the ∇ operator to the function of type $f(\mathbf{x}, p_1(\mathbf{x}), p_2(\mathbf{x}), \ldots, p_m(\mathbf{x}))$ may lead to some misunderstanding. To avoid this, we introduce the operator $\nabla_{\mathbf{x}}$ as

$$\nabla_{\mathbf{x}}(x_1, x_2, x_3, p_1(\mathbf{x}), p_2(\mathbf{x}), \ldots, p_m(\mathbf{x})) = (\partial_1 f, \partial_2 f, \partial_3 f), \tag{1}$$

that is it acts w.r.t. the first three variables of the function f, while ∇f means the gradient vector of the composite function

$$h(\mathbf{x}) = f(\mathbf{x}, p_1(\mathbf{x}), p_2(\mathbf{x}), (\mathbf{x}), \ldots, p_m(\mathbf{x})),$$

that is

$$\nabla f(x_1, x_2, x_3, p_1(\mathbf{x}), p_2(\mathbf{x}), \ldots, p_m(\mathbf{x})) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \right). \tag{2}$$

We would like to emphasize that the right-hand sides of expressions (1) and (2) are usually different because

$$\frac{\partial f}{\partial x_i} = \partial_i f + \sum_{k=1}^{m} \frac{\partial f}{\partial p_k} \frac{\partial p_k}{\partial x_i}, \qquad i = 1, 2, 3.$$

The multiplication of two elements of the space $\mathbb{R}^3$ means the standard scalar product. E.g. for $f, h \in \mathbf{T}$ the notation $\nabla f \nabla h$ yields $\partial_1 f \partial_1 h + \partial_2 f \partial_2 h + \partial_3 f \partial_3 h$ which will be applied in the sequel.

The following properties of the ∇ operator can be easily verified:

- For a scalar function $f \in \mathbf{T}$ and a vector field $\mathbf{g}$ the relation

$$\nabla\cdot(f\mathbf{g}) = (\nabla f)\mathbf{g} + f\,\nabla\cdot\mathbf{g} \tag{3}$$

  holds.

- Due to (3) we have

$$\nabla\cdot(f(\mathbf{M}\mathbf{g})) = (\nabla f)(\mathbf{M}\mathbf{g}) + f\,\nabla\cdot(\mathbf{M}\mathbf{g}), \tag{4}$$

  where $\mathbf{M}$ is any matrix.

- For a scalar function $f(\mathbf{x}, p_1(\mathbf{x}), p_2(\mathbf{x}), \ldots, p_m, (\mathbf{x}))$ and a vector function $\mathbf{g}$ the following relation holds:

$$\nabla\cdot(f\mathbf{g}) = f\,\nabla\cdot\mathbf{g} + \mathbf{g}\nabla_{\mathbf{x}}f + \sum_{j=1}^{m}(\partial_{j+3}f)(\nabla p_j)\mathbf{g}. \tag{5}$$

For the function $\mathbf{R}(\mathbf{x},\mathbf{p}) = (R_1(\mathbf{x},\mathbf{p}), R_2(\mathbf{x},\mathbf{p}), \ldots R_m(\mathbf{x},\mathbf{p}))$ we introduce two Jacobi matrices: the first is defined w.r.t. the variables x_1, x_2, x_3 and denoted by $\mathbf{R}_{\mathbf{x}}(\mathbf{x},\mathbf{p})$, the second one w.r.t. the variables $p_1, p_2, \ldots, p_m$, and denoted by $\mathbf{R}_{\mathbf{x}}(\mathbf{x},\mathbf{p})$. Consequently, they are matrices of type $\mathbb{R}^{m\times 3}$ and $\mathbb{R}^{m\times m}$, respectively, and are defined by the formulas

$$(\mathrm{R}_{\mathbf{x}}(\mathbf{x},\mathbf{p}))_{i,j} = \partial_j R_i(\mathbf{x},\mathbf{p}), \quad i = 1,2,\ldots,m \text{ and } j = 1,2,3,$$

and

$$(\mathbf{R}_{\mathbf{p}}(\mathbf{x},\mathbf{p}))_{i,j} = \partial_{3+j} R_i(\mathbf{x},\mathbf{p}), \quad i,j = 1,2,\ldots,m. \tag{6}$$

Here and further we assume the required smoothness of the functions in the definitions.

3. Operators used in air pollution models and their L-commutativity

First we define the operators $\mathbf{A}_i : \mathbf{S} \to \mathbf{S}$, $(i = 1,2,3,4,5)$, appearing in air pollution models as follows:

- $(\mathbf{A}_1(\mathbf{c}))_l := -\nabla\cdot(\mathbf{u}c_l)$, $l = 1,2,\ldots,m$, $\mathbf{c} \in \mathbf{S}$, which is associated with the process *advection;*
- $(\mathbf{A}_2(\mathbf{c}))_l := \nabla\cdot(\mathbf{k}\nabla c_l)$, $l = 1,2,\ldots,m$, $\mathbf{c} \in \mathbf{S}$, which is associated with the process *diffusion;*
- $(\mathbf{A}_3(\mathbf{c}))_l := \sigma_l c_l$, $l = 1,2,\ldots,m$, $\mathbf{c} \in \mathbf{S}$, which is associated with the process *deposition;*

- $(\mathbf{A}_4(\mathbf{c}))_l := g_l$, $l = 1, 2, \ldots, m$, $\mathbf{c} \in \mathbf{S}$, which is associated with the process *emission;*
- $(\mathbf{A}_5(\mathbf{c}))_j := R_l(\mathbf{x}, \mathbf{c})$, $l = 1, 2, \ldots, m$, $\mathbf{c} \in \mathbf{S}$, which is associated with the process *chemistry.*

In the following the *L-commutativity* of two differentiable operators plays a central role.

Assume that $\mathbf{A}, \mathbf{B} : \mathbf{S} \to \mathbf{S}$ are differentiable on $\mathbf{S}$. We define the operator $E_{\mathbf{A},\mathbf{B}} : \mathbf{S} \to \mathbf{S}$ as follows:

$$E_{\mathbf{A},\mathbf{B}}(\mathbf{s}) := (\mathbf{B}'(\mathbf{s}) \circ \mathbf{A})(\mathbf{s}) - (\mathbf{A}'(\mathbf{s}) \circ \mathbf{B})(\mathbf{s}), \quad \mathbf{s} \in \mathbf{S}. \tag{7}$$

DEFINITION 3.1. The operator $E_{\mathbf{A},\mathbf{B}}$ is called the commutator of the operators $\mathbf{A}$ and $\mathbf{B}$. The element $E_{\mathbf{A},\mathbf{B}}(\mathbf{s}) \in \mathbf{S}$ is called the commutator error of the operators $\mathbf{A}$ and $\mathbf{B}$ for the element $\mathbf{s} \in \mathbf{S}$.

Obviously, $E_{\mathbf{A},\mathbf{B}} = -E_{\mathbf{B},\mathbf{A}}$. Let $\Lambda_{\mathbf{A},\mathbf{B}}$ denote the subspace of those elements in $\mathbf{S}$ for which the commutator error turns into zero, that is $\Lambda_{\mathbf{A},\mathbf{B}} = \{\mathbf{s} \in \mathbf{S} : E_{\mathbf{A},\mathbf{B}}(\mathbf{s}) = 0\}$.

DEFINITION 3.2. We say that the operators $\mathbf{A}$ and $\mathbf{B}$ L-commute on Λ_0 if $\Lambda_0 = \Lambda_{\mathbf{A},\mathbf{B}}$. If $\Lambda_{\mathbf{A},\mathbf{B}} = \mathbf{S}$, then we say that the operators $\mathbf{A}$ and $\mathbf{B}$ L-commute, that is the operators $\mathbf{A}$ and $\mathbf{B}$ L-commute if the relation

$$E_{\mathbf{A},\mathbf{B}}(\mathbf{s}) = 0, \quad \forall \mathbf{s} \in \mathbf{S} \tag{8}$$

holds.

REMARK 3.1. If $\mathbf{A}$ and $\mathbf{B}$ are linear operators then $\mathbf{A}'(\mathbf{s}) = \mathbf{A}$ and $\mathbf{B}'(\mathbf{s}) = \mathbf{B}$ for any $\mathbf{s} \in \mathbf{S}$. In this case (8) turns into the formula $\mathbf{B} \circ \mathbf{A} = \mathbf{A} \circ \mathbf{B}$, hence the L-commutativity is equivalent to the usual commutativity.

Our goal is to analyse the L-commutativity of any pairs of the operators $\mathbf{A}_i$, $i = 1, 2, \ldots, 5$. To this aim we compute their derivatives.

The operators $\mathbf{A}_i$, $i = 1, 2, 3$ are linear. Therefore, the following relations hold for their derivatives:

$$\mathbf{A}_i'(\mathbf{c}) = \mathbf{A}_i, \qquad \forall \mathbf{c} \in \mathbf{S}, \quad i = 1, 2, 3. \tag{9}$$

Furthermore, the following relation follows from the fact that the operator $\mathbf{A}_4$ is constant:

$$\mathbf{A}_4'(\mathbf{c}) = 0, \qquad \forall \mathbf{c} \in \mathbf{S}. \tag{10}$$

The derivative of $\mathbf{A}_5$ at the point $\mathbf{c} \in \mathbf{S}$ is the Jacobi matrix $\mathbf{R}_{\mathbf{c}}(\mathbf{x}, \mathbf{c})$ and it acts as follows:

$$\mathbf{A}_5'(\mathbf{c})(\mathbf{c}) = \mathbf{R}_{\mathbf{c}}(\mathbf{x}, \mathbf{c})\mathbf{c}, \qquad \forall \mathbf{c} \in \mathbf{S}. \tag{11}$$

It is necessary to emphasize here the following fact: due to the special structures of the first four operators (the l-th component of these operators depend only on c_l), it is sufficient to study componentwise the L-commutativity properties of the pairs $(\mathbf{A}_i, \mathbf{A}_j)$, where $i, j = 1, 2, 3, 4$. This observation is used to facilitate the proofs of some of the following theorems.

Some *sufficient* conditions for the L-commutativity of some of the operators defined above can be found in the literature. For instance,

1. if $\mathbf{u}$ and $\mathbf{k}$ are independent of $\mathbf{x}$, then the operators $\mathbf{A}_1$ and $\mathbf{A}_2$;
2. if $\nabla \cdot \mathbf{u} = 0$ and $\mathbf{R}$ is independent of $\mathbf{x}$, then the operators $\mathbf{A}_1$ and $\mathbf{A}_5$;
3. if $\mathbf{R}$ is independent of $\mathbf{x}$ and linear in $\mathbf{c}$, then the operators $\mathbf{A}_2$ and $\mathbf{A}_5$

L-commute [5].

However, the *necessity* of these strong and unrealistic conditions is not clear. For example, the condition 1 (especially, the independence of $\mathbf{u}$ of $\mathbf{x}$) is very unrealistic because the velocity field $\mathbf{u}$ can strongly depend on both $\mathbf{x}$ and t. Therefore, it is worthwhile to examine the possibility to relax these assumptions by replacing them by some weaker, more realistic conditions. E.g. the condition $\nabla \cdot \mathbf{u} = 0$ is much more realistic because it describes the continuity principle in the lower layers of the atmosphere. We shall analyse the commutativity in the next section under this natural condition, too. We shall also give some exact (necessary and sufficient) conditions for the L-commutativity of the operators $\mathbf{A}_i$ and $\mathbf{A}_j$ $(i, j = 1, \ldots, 5)$.

4. Condition of L-commutativity of the operators in air pollution models

In this section, we shall derive conditions for the L-commutativity of different pairs of the operators $\mathbf{A}_i$, $\mathbf{A}_j$, $i, j = 1, 2, 3, 4, 5$. For the sake of brevity, we shall use the notations $E_{i,j} := E_{\mathbf{A}_i, \mathbf{A}_j}$ and $\Lambda_{i,j} := \Lambda_{\mathbf{A}_i, \mathbf{A}_j}$.

4.1. L-commutativity of the advection and diffusion operators

As was stated in the previous section, it is sufficient to treat these operators componentwise. This means that if an arbitrary component of the advection operator L-commutes with the corresponding component of the diffusion operator, then the operators will be L-commutative. By use of this fact, we obtain that the commutator operator reads

$$(E_{1,2}(\mathbf{c}))_l := \nabla \cdot [\mathbf{k} \nabla (-\nabla \cdot (\mathbf{u} c_l))] + \nabla \cdot [\mathbf{u} (\nabla \cdot (\mathbf{k} \nabla c_l))], \tag{12}$$

for all $l = 1, 2, \ldots, m$.

In the following we examine the commutator under the condition

$$\mathbf{k} \text{ is independent of } \mathbf{x} \text{ and } \nabla\cdot\mathbf{u} = 0. \tag{13}$$

Then, by using the properties of the ∇ operator, after straightforward, but tedious calculations we obtain that (12) yields the relation

$$(E_{1,2}(\mathbf{c}))_l = -2\sum_{s=1}^{3}\sum_{r=1}^{3} k_s(\partial_s\partial_r c_l)(\partial_s u_r) - \sum_{s=1}^{3}\sum_{r=1}^{3} k_s(\partial_r c_l)(\partial_s^2 u_r). \tag{14}$$

The equations $(E_{1,2}(\mathbf{c}))_l = 0$, $l = 1, 2, \ldots, m$ define a system of second order PDE's and the set of its solution is $\Lambda_{1,2} \subset \mathbf{S}$.

As one can easily check in case $\mathbf{u}$ is independent of $\mathbf{x}$, the velocity field is divergencefree (continuity assumption) and the relation $\Lambda_{1,2} = \mathbf{S}$ (that is the L-commutativity of the operators $\mathbf{A}_1$ and $\mathbf{A}_2$) holds. On the other hand, if one of the following conditions are satisfied:

- $\mathbf{u}$ is linear,
- $k_1 = k_2 = k_3 =$ const. and the functions u_1, u_2 and u_3 are harmonic functions w.r.t. $\mathbf{x}$ [4], i.e., $\Delta\mathbf{u} = 0$,

then $\mathbf{S}_{\text{lin}} \subset \Lambda_{1,2}$ that is the operators $\mathbf{A}_1$ and $\mathbf{A}_2$ L-commute on the linear elements. (The latter choices can be interpreted as an approximation to the general case.)

4.2. *L*-commutativity of the advection and deposition operators

Due to property (3), in a similar way as in Subsection 4.1, we obtain

$$(E_{1,3}(\mathbf{c}))_l = \sigma_l[-\nabla\cdot(c_l\mathbf{u})] + \nabla\cdot[\mathbf{u}(\sigma_l c_l)] = c_l(\nabla\sigma_l)\mathbf{u}. \tag{15}$$

By using the relation (15) we get: the operators $\mathbf{A}_1$ and $\mathbf{A}_3$ are L-commuting if and only if the gradient of each deposition function is orthogonal to the velocity field, that is the condition

$$(\nabla\sigma_l)\mathbf{u} = 0 \tag{16}$$

holds for $l = 1, 2, \ldots, m$.

4.3. *L*-commutativity of the advection and emission operators

By using the formula (3), we obtain

$$(E_{1,4}(\mathbf{c}))_l = \nabla\cdot(g_l\mathbf{u}) = (\nabla g_l)\mathbf{u} + g_l\nabla\cdot\mathbf{u}, \tag{17}$$

which implies the following: The operators $\mathbf{A}_1$ and $\mathbf{A}_4$ are *L*-commuting if and only if the condition

$$\nabla\cdot(g_l\mathbf{u}) = 0 \tag{18}$$

holds for $l = 1, 2, \ldots, m$. If the continuity condition $\nabla\cdot\mathbf{u} = 0$ is assumed, then the commutativity holds if and only if the gradients of each emission functions are orthogonal to the velocity field, that is the condition

$$(\nabla g_l)\mathbf{u} = 0 \tag{19}$$

holds for $l = 1, 2, \ldots, m$.

4.4. *L*-commutativity of the advection and chemistry operators

For the commutator of the advection and chemistry operators we can write:

$$(E_{1,5}(\mathbf{c}))_l = -\sum_{j=1}^{m}(\mathbf{R_c}(\mathbf{x},\mathbf{c}))_{l,j}\nabla\cdot(c_j\mathbf{u}) + \nabla\cdot(R_l(\mathbf{x},\mathbf{c})\mathbf{u}). \tag{20}$$

Using (3) and the notation (6), we obtain

$$\sum_{j=1}^{m}(\mathbf{R_c}(\mathbf{x},\mathbf{c}))_{l,j}\nabla\cdot(c_j\mathbf{u}) = \sum_{j=1}^{m}\frac{\partial R_l(\mathbf{x},\mathbf{c})}{\partial c_j}((\nabla c_j)\mathbf{u} + c_j\cdot\nabla\mathbf{u}). \tag{21}$$

Further, applying the formula (5) we get

$$\nabla\cdot(R_l(\mathbf{x},\mathbf{c})\mathbf{u}) = R_l(\mathbf{x},\mathbf{c})\nabla\cdot\mathbf{u} + \sum_{j=1}^{m}\frac{\partial R_l(\mathbf{x},\mathbf{c})}{\partial c_j}(\nabla c_j)\mathbf{u} + \mathbf{u}\nabla_{\mathbf{x}}R_l(\mathbf{x},\mathbf{c}). \tag{22}$$

Combining (21) and (22) with (20), for the l-th component of the commutator we obtain

$$(E_{1,5}(\mathbf{c}))_l = -\sum_{j=1}^{m}\frac{\partial R_l(\mathbf{x},\mathbf{c})}{\partial c_j}c_j\nabla\cdot\mathbf{u} + R_l(\mathbf{x},\mathbf{c})\nabla\cdot\mathbf{u} + \mathbf{u}\nabla_{\mathbf{x}}R_l(\mathbf{x},\mathbf{c}). \tag{23}$$

Consequently, the operators $\mathbf{A}_1$ and $\mathbf{A}_5$ L-commute if the relations

$$\nabla\cdot\mathbf{u} = 0 \qquad \text{and} \qquad \mathbf{u}\nabla_{\mathbf{x}} R_l(\mathbf{x}, \mathbf{c}) = 0 \tag{24}$$

hold for all $l = 1, 2, \ldots, m$ and $\mathbf{c} \in \mathbf{S}$. Obviously, under the continuity assumption $\nabla\cdot\mathbf{u} = 0$, a fixed element $\mathbf{c} \in \mathbf{S}$ belongs to $\Lambda_{1,5}$ (that is $\mathbf{A}_1$ and $\mathbf{A}_5$ are L-commuting on this element) if and only if the conditions

$$\sum_{i=1}^{3} u_i(\mathbf{x}, t)\partial_i R_l(\mathbf{x}, \mathbf{c}(\mathbf{x}, t)) = 0 \qquad \forall l = 1, 2, \ldots, m \tag{25}$$

are satisfied. Therefore, in case of explicit independence of the functions R_l of the variable $\mathbf{x}$, that is in the case $R_l(\mathbf{x}, \mathbf{c}) = R_l(\mathbf{c})$, the conditions in (25) are fulfilled for all $\mathbf{c} \in \mathbf{S}$, so, under these assumptions the operators $\mathbf{A}_1$ and $\mathbf{A}_5$ are L-commuting.

4.5. L-commutativity of the diffusion and deposition operators

By use of the relation (4) we obtain

$$\begin{aligned}(E_{2,3}(\mathbf{c}))_l &= \sigma_l[\nabla\cdot(\mathbf{k}\nabla c_l)] - \nabla\cdot[\mathbf{k}\nabla(\sigma_l c_l)] = \\ &= -(\nabla\sigma_l)(\mathbf{k}\nabla c_l) - c_l\nabla\cdot[\mathbf{k}(\nabla\sigma_l)] - (\nabla c_l)\mathbf{k}\nabla\sigma_l.\end{aligned} \tag{26}$$

This means that the operators $\mathbf{A}_2$ and $\mathbf{A}_3$ L-commute if the condition

$$\nabla\sigma_l = 0 \tag{27}$$

is satisfied for all $l = 1, 2, \ldots, m$. Therefore in case $\sigma_l = \text{const}$ the operators $\mathbf{A}_2$ and $\mathbf{A}_3$ L-commute on any element of $\mathbf{S}$.

4.6. L-commutativity of the diffusion and emission operators

For this case we get the relation

$$(E_{2,4}(\mathbf{c}))_l = -\nabla\cdot[\mathbf{k}\nabla g_l], \tag{28}$$

which means the following: The operators $\mathbf{A}_2$ and $\mathbf{A}_4$ are L-commuting if and only if the condition

$$\frac{\partial}{\partial x_1}\left(k_1\frac{\partial g_l}{\partial x_1}\right) + \frac{\partial}{\partial x_2}\left(k_2\frac{\partial g_l}{\partial x_2}\right) + \frac{\partial}{\partial x_3}\left(k_3\frac{\partial g_l}{\partial x_3}\right) = 0 \tag{29}$$

is satisfied for all $l = 1, 2, \ldots, m$. Clearly, if $\nabla g_l = 0$ for all $l = 1, 2, \ldots, m$, then the operators $\mathbf{A}_2$ and $\mathbf{A}_4$ are L-commuting.

On the base of (29) we can formulate an important for the practice corollary, which gives the necessary and sufficient condition of L-commutativity of the diffusion and emission operators: if $k_1(\mathbf{x},t) = k_2(\mathbf{x},t) = k_3(\mathbf{x},t) = \text{const}$ then the operators $\mathbf{A}_2$ and $\mathbf{A}_4$ are L-commuting if and only if $\nabla\mathbf{g} = 0$ that is the components of the given $\mathbf{g} \in \mathbf{S}$ are harmonic functions.

4.7. L-commutativity of the diffusion and chemistry operators

For the commutator operator we have

$$(E_{2,5}(\mathbf{c}))_l = \sum_{j=1}^{m} \partial_{j+3} R_l(\mathbf{x},\mathbf{c}) \nabla\cdot(\mathbf{k}\nabla c_j) - \nabla\cdot(\mathbf{k}\nabla_{\mathbf{x}} R_l(\mathbf{x},\mathbf{c})). \tag{30}$$

A cumbersome calculation gives the following result:

$$\begin{aligned}(E_{2,5}(\mathbf{c}))_l = -\nabla\cdot(\mathbf{k}\nabla_{\mathbf{x}} R_l(\mathbf{x},\mathbf{c})) - \sum_{j=1}^{m}(\nabla_x \partial_{j+3} R_l(\mathbf{x},\mathbf{c}))(\mathbf{k}\nabla c_j) - \\ - \sum_{j=1}^{m}\sum_{r=1}^{m}((\partial_{r+3}\partial_{j+3} R_l(\mathbf{x},\mathbf{c}))\nabla c_r)(\mathbf{k}\nabla c_j).\end{aligned} \tag{31}$$

By use of (31) clearly we have: under the conditions $\partial_1 R_1(\mathbf{x},\mathbf{c})) = \partial_2 R_l(\mathbf{x},\mathbf{c})) = \partial_3 R_l(\mathbf{x},\mathbf{c})) = 0$ and $\partial_{r+3}\partial_{j+3} R_l(\mathbf{x},\mathbf{c}) = 0$ for all $r, j, l = 1, 2, \ldots, m$ the operators $\mathbf{A}_2$ and $\mathbf{A}_5$ L-commute. Consequently, in case $\mathbf{R}(\mathbf{x},\mathbf{c}) = \mathbf{R}(\mathbf{c})$ and $\mathbf{R}(\mathbf{c}) \in \mathbf{S}_{\text{lin}}$ the operators L-commute.

4.8. L-commutativity of the deposition and emission operators

Obviously, the following relationships are valid for all $l = 1, 2, \ldots, m$:

$$(E_{3,4}(\mathbf{c}))_l = -\sigma_l g_l. \tag{32}$$

Consequently, the operators $\mathbf{A}_3$ and $\mathbf{A}_4$ are not L-commuting in any realistic case.

4.9. *L*-commutativity of the deposition and chemistry operators

Clearly, by definition

$$(E_{3,5}(\mathbf{c}))_l = \sum_{j=1}^{m}[\partial_{j+3}R_l(\mathbf{x},\mathbf{c})]\sigma_j c_j - \sigma_l R_l(\mathbf{x},\mathbf{c}). \tag{33}$$

Assume that $\sigma_1 = \sigma_2 = \ldots = \sigma_m = \sigma$. Then we have

$$(E_{3,5}(\mathbf{c}))_l = \sigma \left[\sum_{j=1}^{m}(\partial_{j+3}R_l(\mathbf{x},\mathbf{c}))c_j - R_l(\mathbf{x},\mathbf{c})\right]. \tag{34}$$

Obviously, the splitting error turns into zero for all $\mathbf{c} \in \mathbf{S}$ if and only if the relation

$$\sum_{j=1}^{m} \frac{\partial R_l(\mathbf{x},\mathbf{c})}{\partial c_j} c_j = R_l(\mathbf{x},\mathbf{c}) \tag{35}$$

is satisfied for all $l = 1, 2, \ldots, m$. Let us examine the case $R_l(\mathbf{x},\mathbf{c}) = R_l(\mathbf{c})$. Then for all fixed l, (35) yields a partial differential equation of first order. This equation has the general solution

$$R_l(c_1, c_2, \ldots, c_m) = c_m \varphi_l \left(\frac{c_1}{c_m}, \frac{c_2}{c_m}, \ldots, \frac{c_{m-1}}{c_m}\right), \tag{36}$$

where $\varphi_l : \mathbb{R}^{m-1} \to \mathbb{R}$ is any continuously differentiable function for all $l = 1, \ldots, m$. Therefore, we obtained: under the conditions $\sigma_1 = \sigma_2 = \ldots = \sigma_m = \sigma$ and $\mathbf{R}(\mathbf{x},\mathbf{c}) = \mathbf{R}(\mathbf{c})$ the operators $\mathbf{A}_3$ and $\mathbf{A}_5$ are L-commuting if and only the functions $R_l(c_1, \ldots, c_m)$ have the form (36) for all $l = 1, \ldots, m$.

4.10. *L*-commutativity of the emission and chemistry operators

By definition

$$(E_{4,5}(\mathbf{c}))_l = \sum_{j=1}^{m}[\partial_{j+3}R_l(\mathbf{x},\mathbf{c})]g_j. \tag{37}$$

Consequently, the operators $\mathbf{A}_4$ and $\mathbf{A}_5$ L-commute if and only if $\mathbf{g}$ lies in the null space of the Jacobian $\mathbf{R_c}(\mathbf{x},\mathbf{c})$, that is $\mathbf{g} \in \ker \mathbf{R_c}(\mathbf{x},\mathbf{c})$.

4.11. Summarizing the *L*-commutativity of the operators

Here we give a short summary of the results obtained in the previous sections.

- The commutator error $E_{3,4}$ does not vanish in any realistic case.
- Under the assumptions

  (38) $$\nabla\cdot\mathbf{u} = 0, \quad \nabla\sigma_l = 0, \quad \nabla\mathbf{g} = 0, \quad \mathbf{R}(\mathbf{x}, \mathbf{c}) = \mathbf{R}(\mathbf{c})$$

  the commutator errors $E_{1,3}$, $E_{1,4}$, $E_{1,5}$, $E_{2,3}$ and $E_{2,4}$ vanish.
- Under the further conditions

  (39) $$\mathbf{k}(\mathbf{x}) = \text{const}, \quad \mathbf{u},\ \mathbf{R},\ \mathbf{c} \quad \text{are linear}, \quad \sigma_1 = \ldots = \sigma_m = \text{const}.$$

  the commutator errors $E_{1,2}$, $E_{2,5}$ and $E_{3,5}$ are also zero.
- If in addition $\mathbf{g} \in \ker \mathbf{J}_\mathbf{R}$, then even the operators $\mathbf{A}_4$ and $\mathbf{A}_5$ *L*-commute.

Clearly, these results cover those of [5], however, for certain pairs of operators they give more general conditions for the *L*-commutativity than the requirements formulated there. For instance, if $\mathbf{u}$ is linear (not necessarily constant), then for concentration functions of a special from (linear functions) the operators $\mathbf{A}_1$ and $\mathbf{A}_2$ *L*-commute.

Finally we remark that the assumptions of the linearity can be interpreted in the following way, too. The operators $\mathbf{A}_i$, $i = 1, 2, \ldots, 5$, are defined on the linear finite elements and the derivation is understood in generalized sense. We define the operators $\mathbf{A}_i$, $1, 2, \ldots, 5$, as the mappings which are obtained after the semidiscretization of the weak form of the original fully continuous PDE's, in the linear finite element spaces. Then the functions $\mathbf{g}$ and $\mathbf{u}$ in the definition of the operators can be considered as the projections of the original functions into the linear finite element subspace.

5. Application of the splitting error analysis in air pollution models

In this section we present two examples of air pollution models in which the above results can be applied: the Danish Eulerian Model (DEM) [22] and the advection – diffusion – reaction model as defined in [5].

For the first model a splitting procedure based on a separation of the physical processes is used. This way of implementing a splitting procedure in air pollution modelling was first proposed in [8]. We shall use the abbreviation *DEM splitting* in this paper in order (i) to facilitate the references to this splitting and (ii) to emphasize the fact that it has been used in the Danish Eulerian Model.

For the second model we shall use a symmetrical splitting proposed simultaneously in [6, 7] and [16]. We refer to this splitting procedure as the Marchuk–Strang splitting procedure in this paper. A very good description of this way of splitting, which is particularly oriented to air pollution models, is given in [12].

5.1. Air pollution models and mathematical formulation of the long-range transport of air pollutants

The air pollution models must satisfy several important requirements [22, 12, 23]:

1. The mathematical models must be defined on large space domains, because the long range transport of air pollution is an important environmental phenomenon, and high pollution levels are not limited to the areas where the high emission sources are located.
2. All relevant physical and chemical processes must be adequately described in the models used.
3. Enormous files of input data (both meteorological data and emission data) are needed.
4. The output files are also very big, and fast visualization tools must be used in order to represent the trends and tendencies, hidden behind many megabytes (or even many gigabytes) of digital information, so that even non-specialists can easily understand them.

All important physical and chemical processes must be taken into account when an air pollution model is to be developed. Systems of partial differential equations (PDE's) are often used to describe mathematically an air pollution model. Consider a three-dimensional space domain Ω and assume that $\mathbf{x} \equiv$ $\equiv (x_1, x_2, x_3) \in \Omega$. Then the PDE systems are of the following type:

(40)
$$\frac{\partial c_l(\mathbf{x},t)}{\partial t} = \mathbf{A}(\mathbf{x},t)c_l(\mathbf{x},t) + f(\mathbf{x},t), \ t \in [0,T], \ c_l(\mathbf{x},0) = c_{l0}(\mathbf{x}), \ l = 1,\dots,m$$

where

$$\mathbf{A}(\mathbf{x},t)c_l \equiv -\nabla\cdot(\mathbf{u}c_l) + \nabla\cdot(\mathbf{k}\nabla c_l) - (\sigma_1 + \sigma_2)c_l, \quad l = 1,\dots,m, \tag{41}$$

$$\nabla\cdot(\mathbf{u}c_l) = \frac{\partial(u_1 c_l)}{\partial x_1} + \frac{\partial(u_2 c_l)}{\partial x_2} + \frac{\partial(u_3 c_l)}{\partial x_3}, \quad l = 1,\dots,m. \tag{42}$$

$$\nabla(k\nabla c_l) = \frac{\partial}{\partial x_1}\left(k_1\frac{\partial c_l}{\partial x_1}\right) + \frac{\partial}{\partial x_2}\left(k_2\frac{\partial c_l}{\partial x_2}\right) + \frac{\partial}{\partial x_3}\left(k_3\frac{\partial c_l}{\partial x_3}\right), \quad l = 1,\dots,m. \tag{43}$$

The vector-function $f(\mathbf{x},t)$ is defined as a sum

$$f(\mathbf{x},t) = g + R, \tag{44}$$

where

$$g \equiv (g_1,\dots,g_m)^T,$$
$$R \equiv (R_1,\dots,R_m)^T,$$

and

$$R_l = R_l(\mathbf{x}, c_1, c_2, \dots, c_m), \qquad l = 1,2,\dots,m. \tag{45}$$

The different quantities that are involved in the mathematical model have the following meaning:

- c_l denotes the concentration of the l-th species;
- u_1, u_2 and u_3 are velocities;
- k_1, k_2 and k_3 are diffusion coefficients;
- the functions g_l describe the emission sources in the space domain;
- α_{1l} and α_{2l} are the dry and wet deposition coefficients, respectively;
- the nonlinear functions $R_l(c_1, c_2, \dots, c_m)$ describe the chemical reactions.

The functions R_l, representing the chemical reactions in which the l-th pollutant is involved are of the form

$$R_l(c_1, c_2, \dots, c_m) = -\sum_{j=1}^{m} \alpha_{lj} c_j + \sum_{j=1}^{m}\sum_{k=1}^{m} \beta_{ljk} c_j c_k, \qquad l = 1,2,\dots,m.$$

This is a special kind of nonlinearity (it is seen that the chemical terms are described by quadratic functions), but it is not clear if this property can efficiently be exploited. To the authors' knowledge, it is not exploited in the existing large scale air pollution models.

The models defined by (40)–(45) are traditionally used to calculate some concentration fields by using both meteorological and emission data as input [23]. This gives an answer to the question: what are the concentration levels and/or the deposition levels caused by the existing emissions under the particular meteorological conditions that take place in the time-period under consideration? However, it is much more important to study the question: how can the concentrations be kept under certain critical levels?

5.2. Splitting and its role

It is very difficult to treat directly the system (40). Therefore, some kind of *splitting* is to be used. *Splitting*, or problem decomposition, is commonly used during the first step of the numerical treatment of large air pollution models (as in many other large-scale scientific and engineering problems). The big problem, the model described by the system of equations (40), is divided into several smaller problems through some splitting procedure. These smaller problems might have special properties that can be exploited in the numerical solutions. For example, the systems of linear algebraic equations that arise, after splitting, from the diffusion part of the model normally have banded, symmetric, and positive definite matrices. On the other hand, it is not easy to evaluate the error that arises from the splitting techniques used.

Splitting according to the major physical processes is very popular; see, for example, [9], [12] and [22]. Such splitting procedures lead often to a number of sub-models which are to be treated cyclicly at every time-step [22]. In the DEM [22] these sub-models are describing the horizontal advection (46), the horizontal diffusion (47), the chemical reactions including the emissions (48), the deposition (49) and the vertical exchange (50), so there are five splitted systems of the form

$$\frac{\partial c_l^{(1)}}{\partial t} = -\frac{\partial (u_1 c_l^{(1)})}{\partial x_1} - \frac{\partial (u_2 c_l^{(1)})}{\partial x_2} \tag{46}$$

$$\frac{\partial c_l^{(2)}}{\partial t} = \frac{\partial}{\partial x_1}\left(k_1 \frac{\partial c_l^{(2)}}{\partial x_1}\right) + \frac{\partial}{\partial x_2}\left(k_2 \frac{\partial c_l^{(2)}}{\partial x_2}\right) \tag{47}$$

$$\frac{\partial c_l^{(3)}}{\partial t} = g_l + R_l(c_1^{(3)}, c_2^{(3)}, \ldots, c_m^{(3)}) \tag{48}$$

$$\frac{\partial c_l^{(4)}}{\partial t} = -(\sigma_{1l} + \sigma_{2l})c_l^{(4)} \tag{49}$$

$$\frac{\partial c_l^{(5)}}{\partial t} = -\frac{\partial (u_3 c_l^{(5)})}{\partial x_3} + \frac{\partial}{\partial x_3}\left(k_3 \frac{\partial c_l^{(5)}}{\partial x_3}\right). \tag{50}$$

The values $c_l^{(j)}$, $j = 1, \ldots, 5$ are connected through the initial conditions, that is $c_l^{(j)}$ is used as an initial condition for $c_l^{(j+1)}$, $j = 1, \ldots, 4$, and for the next time-step the process continues cyclicly. We shall call the splitting procedure (46)–(50) *DEM splitting procedure.*

An alternative of the above splitting procedure can be the already mentioned *symmetrical Marchuk–Strang splitting scheme* [6, 7, 16]. Usually, this splitting scheme is applied to a model of air pollution transport, where the deposition and emission parts are included into the reaction part of the problem operator. So, instead of (40) we consider an advection-diffusion-reaction problem

$$\frac{\partial \mathbf{c}(\mathbf{x},t)}{\partial t} = A(\mathbf{c}(\mathbf{x},t)), \quad t \in (0,T], \quad \mathbf{c}(\mathbf{x},0) = \mathbf{c}_0(\mathbf{x}), \tag{51}$$

where

$$A(\mathbf{c}(\mathbf{x},t)) = A_1(\mathbf{c}(\mathbf{x},t)) + A_2(\mathbf{c}(\mathbf{x},t)) + A_5(\mathbf{c}(\mathbf{x},t)),$$

and it is assumed that into $\mathbf{A}_5(c(\mathbf{x},t))$ the deposition and emissions are included. In the Marchuk–Strang splitting the problem (51) is split by ordering the operators A_1, A_2 and A_5 symmetrically in the following way:

$$\frac{\partial \mathbf{c}^{(1)}(\mathbf{x},t)}{\partial t} = A_1(\mathbf{c}^{(1)}(\mathbf{x},t)), \quad t \in \left(0, \frac{\tau}{2}\right], \quad \mathbf{c}^{(1)}(\mathbf{x},0) = \mathbf{c}_0(\mathbf{x}), \tag{52}$$

$$\frac{\partial \mathbf{c}^{(2)}(\mathbf{x},t)}{\partial t} = A_2(\mathbf{c}^{(2)}(\mathbf{x},t)), \quad t \in \left(0, \frac{\tau}{2}\right], \quad \mathbf{c}^{(2)}(\mathbf{x},0) = \mathbf{c}^{(1)}\left(\mathbf{x}, \frac{\tau}{2}\right), \tag{53}$$

$$\frac{\partial \mathbf{c}^{(3)}(\mathbf{x},t)}{\partial t} = A_5(\mathbf{c}^{(3)}(\mathbf{x},t)), \quad t \in (0,\tau], \quad \mathbf{c}^{(3)}(\mathbf{x},0) = \mathbf{c}^{(2)}\left(\mathbf{x}, \frac{\tau}{2}\right), \tag{54}$$

$$\frac{\partial \mathbf{c}^{(4)}(\mathbf{x},t)}{\partial t} = A_2(\mathbf{c}^{(4)}(\mathbf{x},t)), \quad t \in \left(0, \frac{\tau}{2}, \tau\right], \quad \mathbf{c}^{(4)}\left(\mathbf{x}, \frac{\tau}{2}\right) = \mathbf{c}^{(3)}\left(\mathbf{x}, \frac{\tau}{2}\right), \tag{55}$$

$$\frac{\partial \mathbf{c}^{(5)}(\mathbf{x},t)}{\partial t} = A_1(\mathbf{c}^{(5)}(\mathbf{x},t)), \quad t \in \left(\frac{\tau}{2}, \tau\right], \quad \mathbf{c}^{(5)}\left(\mathbf{x}, \frac{\tau}{2}\right) = \mathbf{c}^{(4)}\left(\mathbf{x}, \frac{\tau}{2}\right). \tag{56}$$

Let us now suppose that one can solve both the original problem and the splitted subproblems exactly. In this case it is possible to express the splitting error with the help of the so-called Lie operator formalism, as will be shown in the next chapter.

5.3. Lie operator formalism and splitting error

In this chapter, following the technique of [15] and [5], we shall derive the local splitting error of the Marchuk-Strang splitting procedure and give the results of a similar error analysis for the DEM splitting. We will see that in terms of the local error the order of the Marchuk–Strang splitting scheme is higher than that of the DEM splitting.

First we need to introduce the concept of Lie operator, since it plays an important role in the derivation of the error formula.

Let A be a generally non-linear operator of type $\mathbf{S} \to \mathbf{S}$. With this given operator we associate a new operator, which we will denote by $\mathscr{A}$ and call it the Lie operator associated to A. This operator acts on the space of differentiable operators $\mathbf{S} \to \mathbf{S}$ and maps each operator F into the new operator $\mathscr{A}F$, such that for any element $c \in \mathbf{S}$,

$$(\mathscr{A}F)(c) = (F'(c) \circ A)(c). \tag{57}$$

It is easy to see that the Lie operator is linear.

Let us consider the initial value problem

$$\frac{\partial c}{\partial t}(\mathbf{x}, t) = A(c(\mathbf{x}, t)), \quad \text{on } (0, T], \quad c(\mathbf{x}, 0) = c_0(\mathbf{x}), \tag{58}$$

and denote by $\mathscr{A}$ the Lie operator associated to the particular operator A of problem (58). Let F be any differentiable operator $\mathbf{S} \to \mathbf{S}$. Applying the operator $\mathscr{A}F$ to the solution $c(\mathbf{x}, t)$ of (51) and using the chain-rule of differentiating, one obtains the relation

$$(\mathscr{A}F)(c(\mathbf{x}, t)) = \frac{\partial}{\partial t} F(c(\mathbf{x}, t)), \tag{59}$$

from which by induction follows also that

$$\frac{\partial^i}{\partial t^i} F(c(\mathbf{x}, t)) = (\mathscr{A}' F)(c(\mathbf{x}, t)), \quad i = 2, 3\ldots \tag{60}$$

Assume that the solution $\mathbf{c}(\mathbf{x}, t)$ of (51) is an analytic function. Then, using its Taylor series expansion, one can easily show that

$$c(\mathbf{x}, \tau) = (e^{\tau\mathscr{A}}, I)(c_0(\mathbf{x})), \tag{61}$$

where I is the identity operator $\mathbf{S} \to \mathbf{S}$.

Applying now to each of the subproblems the corresponding exponentiated Lie operator and composing them in the order defining the Marchuk–Strang splitting procedure (52)–(56), for the solution $\hat{c}$ of the splitted problem at time τ one can get

$$\hat{c}(\mathbf{x}, \tau) = \left(e^{\frac{1}{2}\tau\mathscr{A}_1} e^{\frac{1}{2}\tau\mathscr{A}_2} e^{\frac{1}{2}\tau\mathscr{A}_5} e^{\frac{1}{2}\tau\mathscr{A}_2} e^{\frac{1}{2}\tau\mathscr{A}_1} I \right) (c_0(\mathbf{x})).$$

In order to compute the product of exponentials on the right-hand side, we can use the well-known Baker–Campbell–Hausdorff (BCH) theorem [19]. This claims that for any pair of linear operators X, Y the product $e^X e^Y$ can locally be written as the exponential e^Z of the Lie operator

$$Z = X+Y+\frac{1}{2}[X,Y]+\frac{1}{12}([X,X,Y]+[Y,Y,X])+\frac{1}{24}[X,Y,Y,X]+\ldots, \tag{62}$$

where $[X, Y]$ is the commutator $[X, Y] = XY - YX$ and $[X, X, Y]$ is recursively defined by $[X, X, Y] = [X, [X, Y]]$ etc. Substituting $X = \frac{1}{2}\tau A_1$ etc. and applying (62) four times, we obtain that the Marchuk–Strang solution $\hat{c}$ can be expressed as

$$\hat{c}(\mathbf{x}, \tau) = \left(e^{\tau\hat{\mathcal{A}}} I\right)(c_0(\mathbf{x})),$$

where the new Lie operator $\hat{\mathcal{A}}$ has the form

$$\hat{\mathcal{A}} = \mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_5 + \tau^2 \mathcal{E}_A + O(\tau^4) \tag{63}$$

with

$$\begin{aligned} \mathcal{E}_A = &-\frac{1}{24}[\mathcal{A}_1,\mathcal{A}_1,\mathcal{A}_2] - \frac{1}{24}[\mathcal{A}_1,\mathcal{A}_1,\mathcal{A}_5]+ \\ &+\frac{1}{12}[\mathcal{A}_2,\mathcal{A}_2,\mathcal{A}_1] - \frac{1}{24}[\mathcal{A}_2,\mathcal{A}_2,\mathcal{A}_5] + \frac{1}{12}[\mathcal{A}_5,\mathcal{A}_5,\mathcal{A}_1]+ \\ &+\frac{1}{12}[\mathcal{A}_5,\mathcal{A}_5,\mathcal{A}_2] + \frac{1}{12}[\mathcal{A}_2,\mathcal{A}_5,\mathcal{A}_1] + \frac{1}{12}[\mathcal{A}_5,\mathcal{A}_2,\mathcal{A}_1]. \end{aligned} \tag{64}$$

REMARK 5.1. If $\mathcal{A}_1$ and $\mathcal{A}_2$ are Lie operators, then $[\mathcal{A}_1,\mathcal{A}_2] = 0$ is equivalent to $E_{A_1,A_2} = 0$, where A_1 and A_2 are the operators belonging to the Lie operators $\mathcal{A}_1$ and $\mathcal{A}_2$, respectively.

In order to characterize the error at time τ that arises if we apply operator splitting on the interval $[0,\tau]$, we can use the notion of the local splitting error, defined as the difference between the exact solution of the splitted problem and the exact solution of the original problem [17, 20]. According to the above considerations, for the Marchuk–Strang splitting scheme this local error can be given as

$$\mathrm{Err}_{S_p}(\tau) := \left(e^{\tau\hat{\mathcal{A}}} I - e^{\tau\mathcal{A}} I\right)(c_0(\mathbf{x})). \tag{65}$$

Applying now (63) and the definition of the exponential, we obtain that the behaviour of the error function as $\tau \to 0$ is as follows:

$$\mathrm{Err}_{S_p}(\tau) = \tau^3(\mathcal{E}_A I)(c_0(\mathbf{x})) + o(\tau^4),$$

i.e. the local splitting error of the Marchuk–Strang scheme is $o(\tau^3)$. Therefore, we say that the Marchuk–Strang splitting is a second order splitting scheme [17].

A similar analysis shows that the local error of the DEM splitting is only $o(\tau^2)$, which means that in the above sense the Marchuk–Strang splitting scheme provides a one order higher approximation to the original problem than the DEM splitting.

We remark that if we apply operator splitting on the interval $[k\tau, (k+1)\tau]$, $k = 1, 2, \dots$, then (65) becomes

$$\mathrm{Err}_{S_p}(\tau) := \left(e^{\tau\hat{\mathscr{A}}}I - e^{\tau\mathscr{A}}I\right)(\hat{c}(\mathbf{x}, k\tau)), \tag{66}$$

where clearly $\hat{c}(\mathbf{x}, k\tau)$ contains some error due to applying splitting in the first k steps.

6. Concluding remarks

Analyzing the splitting error both for the DEM and the Marchuk–Strang splitting procedure, one can conclude:

- If all the pairs (A_i, A_j), where $i,j = 1,2,3,4,5$ and $i \neq j$, in the DEM splitting procedure L-commute, then no splitting error occurs.
- If all the pairs (A_i, A_j), where $i,j = 1,2,5$ and $i \neq j$, in the Marchuk–Strang splitting procedure applied to the advection – diffusion – reaction problem commute, then no splitting error occurs.
- The splitting error in the DEM splitting procedure is of first order if at least one pair (A_i, A_j), where $i,j = 1,2,5$ and $i \neq j$, does not commute.
- The splitting error in the Marchuk–Strang splitting procedure is of second order if at least one pair (A_i, A_j), where $i,j = 1,2,5$ and $i \neq j$, does not commute.
- As we proved in Section 4, for the realistic situations the splitting errors for the operators $\mathbf{A}_3$ and $\mathbf{A}_4$ do not vanish. On the other hand, for the other cases under the assumptions

  $$\nabla\cdot\mathbf{u} = 0, \qquad \nabla\sigma_l = 0, \qquad \nabla\mathbf{g} = 0, \qquad \mathbf{R}(\mathbf{x}, \mathbf{c}) = \mathbf{R}(\mathbf{c})$$

  the splitting errors are equal to zero for each pair of operators in the air pollution modeling with the exception of $(\mathbf{A}_1, \mathbf{A}_2)$, $(\mathbf{A}_2, \mathbf{A}_5)$, $(\mathbf{A}_3, \mathbf{A}_5)$ and $(\mathbf{A}_4, \mathbf{A}_5)$. If additionally we assume the linearity of $\mathbf{u}$, $\mathbf{R}$ and the solution $\mathbf{c}$, $\sigma_l = \sigma = \text{const.}$ and $\mathbf{k}(\mathbf{x}) = \text{const.}$ then the splitting errors exist only for the operators $\mathbf{A}_4$ and $\mathbf{A}_5$. Since for the linear elements $\mathbf{J_R} := \mathbf{R_c}(\mathbf{c}) \in \mathbb{R}^{m\times m}$ is a constant matrix therefore under the condition $\mathbf{g} \in \ker \mathbf{J_R}$ even the last commutator is equal to zero.

- The diurnal cycle strongly influences the commutators leading to a relatively small local splitting error over nightly periods. Specific circumstances determine actual values of the splitting error.

References

[1] I. DIMOV, Z. ZLATEV, *Testing the sensitivity of Air Pollution Levels to Variations of some Chemical Rate Constants,* in: *“Large-Scale Computations in Engineering and Environmental Sciences,” Notes on Numerical Fluid Mechanics, Vol. 62* (M. Griebel, O. Iliev, S. Margenov and P. S. Vassilevski, Eds), Ind Workshop on “Large-Scale Scientific Computations”, Varna, Bulgaria, June, 1997, Proceedings, 1997, pp. 167–175.

[2] I. DIMOV, TZ. OSTROMSKY, I. TZVETANOV, Z. ZLATEV, *Economical Estimation of the Loses of Crops Due to High Ozone Levels,* in: *“Large-Scale Computations in Engineering and Environmental Sciences”, Notes on Numerical Fluid Mechanics* (M. Griebel, S. Margenov and P. Yalamov, Eds), IInd Workshop on “Large-Scale Scientific Computations”, Sozopol, Bulgaria, June 2–6, 1999, Proceedings, Vieweg, Braunschweig, Germany, pp. 273–280.

[3] E. G. DYAKONOV, Difference schemes with splitting operators for multidimensional stationary problems, *Journal of Computational Math. and Math. Phys.,* **2** (1962), No 4 (in Russian).

[4] E. HILLE, *Analytic function theory,* Vol 1., Blaisdell Publishing Company, London, 1959.

[5] D. LANSER and J. G. VERWER, Analysis of operators splitting for advection-diffusion-reaction problems in air pollution modelling, *J. Comput. Appl. Math.,* **111** (1999), 201–216.

[6] G. I. MARCHUK, Some application of splitting-up methods to the solution of mathematical physics problems, *Applik. Mat.,* **13** (1968), No. 2.

[7] G. I. MARCHUK, *On the theory of the splitting-up method,* In: *Numerical solution of partial differential equations.* II. SYSSPADE, 1970, Academic Press, New York–London, 1971.

[8] G. I. MARCHUK, Methods of computational mathematics, Moscow, Nauka, 1980. (in Russian)

[9] G. I. MARCHUK, Mathematical modelling for the problem of the environment, *Studies in Mathematics and Applications, No. 16,* North-Holland, Amsterdam, 1985.

[10] G. I. MARCHUK, Methods of Splitting, Moscow, Nauka, 1988 (in Russian).

[11] G. I. MARCHUK, *Applications of Adjoint Equations to Problems of Global Change and Environmental Protection,* HERMIS'94, Vol. 1, (E.A. Lipitakis, Edr), 1994, Hellenic Mathematical Society, pp. 1–20.

[12] G. J. MCRAE, W. R. GOODIN, J. H. SEINFELD, Numerical solution of the atmospheric diffusion equations for chemically reacting flows, *Journal of Computational Physics,* **45** (1984), 1–42.

[13] V. V. PENENKO, N. N. OBRAZTSOV, A variational initialization model for the fields of meteorological elements, *Soviet Meteorol. Hydrol.*, **11** (1976), 1–11.

[14] W. RUDIN, *Functional Analysis,* Tata McGraw-Hill Publishing Company Ltd., New Delhi, 1973.

[15] J. M. SANZ-SERNA, M. P. CALVO, *Numerical Hamiltonian Problems,* Chapman and Hall, 1994.

[16] G. STRANG, On the construction and comparison of difference schemes, *SIAM Journal Numer. Anal,* **5** (1968), 505–517.

[17] B. SPORTISSE, An Analysis of Operator Splitting Techniques in the Stiff Case, CERMICS-INRIA, 2004, Route des Lucioles – BP 93 – Sophia-Antipolis CEDEX.

[18] A. N. TIKHONOV, A. A. SAMARSKI, *Equations of the Mathematical Physics,* Moscow, Nauka, 1977. (in Russian)

[19] V. S. VARADARAJAN, *Lie Groups, Lie Algebras and Their Representations,* Prentice-Hall Inc., EnglewoodCliffs, New Jersey, 1974.

[20] J. G. VERWER, W. H. HUNDSDORFER and J. G. BLOM, Numerical time integration for air pollution models, CWI Report MAS-R9825, 1998.

[21] N. N. YANENKO, On convergence of the splitting method for heat equation with variable coefficients, *Journal of Computational Math. and Math. Phys.,* **2** (1962), No 5 (in Russian).

[22] Z. ZLATEV, *Computer treatment of large air pollution models,* Kluwer Academic Publishers, Dordrecht–Boston–London, 1995.

[23] Z. ZLATEV, I. DIMOV, K. GEORGIEV, Three-dimensional version of the Danish Eulerian Model, *Zeitschrift für Angewandte Mathematik and Mechanik,* **76** (1996), 473–476.

ANNALES UNIV. SCI. BUDAPEST., **44** (2001), 151–156

THIN COMPLETE SUBSEQUENCE

By

NORBERT HEGYVÁRI

Department of Mathematics, TFK, Eötvös Loránd University, Budapest

(Received December 6, 2001)

1. Introduction

The well known theorem of Lagrange states that every non-negative integer n is the sum of four squares. In other words the sequence $S = \{1, 4, \dots \dots, n^2, \dots\}$ is bases of order four. WIRSING defined the notion of thin bases; A is *thin bases* of order h if $A(x) < c'x^{1/h}$; $(c' > 0)$, where $A(x)$ is the counting function of A. Let us note if A is a bases of order h then $A(x) >> x^{1/h}$. By a non-constructive method Wirsing proved [1] that S has a subbases S' which is almost thin, proving $S'(x) = O(x^{1/4}(\log x)^{1/4})$. Later J. SPENCER [2] gave a short proof for it, using the Janson's inequality, which is an important tool of probabilistic method.

Let us mention it is not even known an explicit subsequence S' of S for which $S'(x) = o(\sqrt{x})$.

A related question would be the following: let $A = \{a_1 < \dots < a_n < < \dots\} \subseteq \mathbb{N}$. A is said to be complete if there exists Δ_A such that for every $n \geq \Delta_A$ we have

$$n \in \Sigma(A) = \{S(B) : S(B) = \sum_{b \in B} b; \text{B is a finite subset of A}, S(\emptyset) = 0\}.$$

Clearly if $|A| = k$ then $|\Sigma(A)| \leq 2^k$. This implies if A is complete then $2^{A(x)} \geq x - \Delta_A$ i.e. for $x \geq x_0$ $A(x) \geq \log_2(x - \Delta_A)$.

---

[1] Research partially supported by Hungarian National Foundation for Scientific Research, Grant No. T025617 and No. TO29759 and by DIMACS (Center for Discrete Mathematics and Theoretical Computer Science) NSF-STC-91-19999.

The related notion of thin bases is the following

DEFINITION. A' is said to be *thin complete subsequence* of A if A' is complete and

$$A'(x) = (1 + o(1)) \log_2 x .$$

We shall show that a wide class of complete sequences have a thin complete subsequence and not merely the sequence of squares S . We prove

THEOREM. *Let $A = \{a_1 < a_2 < \dots\}$ be a complete sequence of integers. Assume that* $\lim_{n\to\infty} a_{n+1}/a_n = 1$. *Then A contains a thin complete subsequence.*

The proof will be completely constructive.

Let $X = \{x_1 < x_2 < \dots\}$. Let us denote

$$(X) = \sup_i \{x_{i+1} - x_i\}.$$

(So that if $G(X) < \infty$ then it indicates the size of the biggest gap in X)

PROOF OF THE THEOREM. We need some lemmas.

LEMMA 1. *Let $X = \{x_1 < \dots < x_n < \dots\} \in \mathbb{N}$. Assume that for every $i = 1, 2, \dots$*

(1) $$x_{i+1} \leq x_1 + \dots + x_i .$$

Then $G(\Sigma(X)) \leq x_1$.

The proof of Lemma 1 is easy or see [3].

LEMMA 2. *Let A be a complete sequence of integers. Let $A_1 = \{2\Delta_A < < a'_1 < a'_2 < \dots\}$ be an infinite subsequence of A for which $a'_{i+1}/a'_i \leq 2$ $i = = 1, 2, \dots$. Let $A_2 = A \cap [1, a'_1)$ and assume there are elements $b_1, b_2 \in \in A$ such that $\Delta_A \leq b_2 - b_1 < a'_1 - \Delta_A$. Furthermore assume the sets A_1, A_2, $\{b_1\}$, $\{b_2\}$ are pairwise disjoint. Then $B = A_1 \bigcup A_2 \bigcup \{b_1\} \bigcup \{b_2\}$ is complete.*

PROOF OF LEMMA 2. Let $A_1 = \{2\Delta_A < a'_1 < a'_2 < \dots\}$.

First we prove that for every $i = 1, 2, \dots$

(2) $$a'_{i+1} \leq a'_1 + a'_2 + \dots + a'_i ,$$

by induction on i. For $i = 1$ (2) is trivial. Furthermore by $a'_{i+2} \leq 2a'_{i+1}$ we get

$$a'_{i+2} \leq 2a_{i+1} \leq (a'_1 + \ldots + a'_i) + a'_{i+1},$$

which provides the inductive steps. So A_1 fulfills (1), hence

$$G(\Sigma(A_1)) \leq a'_1. \tag{3}$$

We claim that for every $n \geq \Delta_A + a'_1 + b_1 + b_2 := \Delta_B$, $n \in \Sigma(B)$. Assume now contrary to the assertion there exists an $n \geq \Delta_B$ and $n \notin \Sigma(B)$. Let t be the subsript defined by

$$a'_t < n - (b_1 + b_2) < a'_{t+1} \tag{4}$$

(clearly the equality cannot hold). Now $(n - b_1) - (n - b_2) = b_2 - b_1$, thus

$$\Delta_A < (n - b_1) - (n - b_2) < a'_1 - \Delta_A. \tag{5}$$

Furthermore $\Sigma(A_2) = \Sigma(A) \cap [1, a'_1)$ and so $\Sigma(A_1) + \Sigma(A_2) \supseteq \Sigma(A_1) + [1, a'_1)$. Thus by (3) we conclude that $\Sigma(A_1)+\Sigma(A_2)$ contains a set which is the union of blocks of consecutive integers with length at lest $a'_1 - \Delta_A$ and gaps at most Δ_A. So (5) implies that there is an $i = 0$ or 1 such that $n - b_i \in \Sigma(A_1)+\Sigma(A_2)$ and thus $n \in \Sigma(A_1) + \Sigma(A_2) + b_i \subseteq \Sigma(B)$ a contradiction.

LEMMA 3. Let $A = \{a_1 < a_2 < \ldots < a_n < \ldots\}$ be an infinite sequence of integers. Assume that $\lim_{n\to\infty} a_{n+1}/a_n = 1$. Then for every $\lambda \geq 1$ real number and $K \in \mathbb{N}$ there exits a subsequence $A_\lambda = \{K < a_{k_1} < a_{k_2} < \ldots < < a_{k_n} < \ldots\}$ of A such that

$$a_{k_{n+1}} - a_{k_n} \geq K \tag{6}$$

and

$$\lim_{n\to\infty} \frac{a_{k_{n+1}}}{a_{k_n}} = 1. \tag{7}$$

PROOF OF LEMMA 3. Let

$$A^* = \{a_K < a_{2K} < \ldots < a_{mK} < \ldots\}. \tag{8}$$

It is obvious that $a_K \geq K$. Furthermore for every m, $a_{(m+1)K} - a_{mK} \geq K$ and

$$\lim_{n\to\infty} \frac{a_{(m+1)K}}{a_{mK}} = \lim_{n\to\infty} \frac{a_{mK+1}}{a_{mK}} \cdots \frac{a_{(m+1)K}}{a_{(m+1)K-1}} = 1.$$

Hence every subsequence of A^* satisfies (6). Let now $a_{k_1} = a_K$ and assume that the elements $a_{k_1} < a_{k_2} < \ldots < a_{k_{n-1}}$ have been defined. Then let $a_{k_n} = \min\{a_t | a_t \leq \lambda a_{k_{n-1}}, a_t \in A\}$. By the definition of a_{k_n} we have

$$a_{k_n} \leq \lambda a_{k_{n-1}} \leq a_{k_n+1}. \tag{9}$$

Hence $\frac{a_{k_n}}{a_{k_{n-1}}} \leq \lambda$. Furthermore

$$\lambda a_{k_{n-1}} \leq a_{k_n} \frac{a_{k_n+1}}{a_{k_n}}$$

and so

$$\lambda \frac{a_{k_n}}{a_{k_n+1}} \leq \frac{a_{k_n}}{a_{k_{n-1}}} \leq \lambda.$$

Since $\lim_{n\to\infty} \frac{a_{k_n}}{a_{k_n+1}} = 1$ we prove the lemma.

PROOF OF THE THEOREM. Let $K = 5\Delta_A$ and $\lambda = 2$. By Lemma 3 we can select a subsequence A_1 of A for which $A_1 = \{2\Delta_A < a'_1 < a'_2 < \ldots < a'_n < < \ldots\}$ and $a'_{n+1} \leq 2a'_n$. Furthermore by (6) and (8) we can choose elements b_1, b_2 of A for which $a'_1 < b_1 < b_2 < a'_1 - \Delta_A$ (say let $b_1 = a_{5\Delta+1}$ and $b_2 = a_{10\Delta-1}$) Then $b_2 - b_1 > \Delta_A$. Finally let $A_2 = A \cap [1, a'_1)$. Clearly the elements b_1, b_2 and the sequences A_1, A_2 satisfies the conditions of Lemma 2 and hence $B = A_1 \bigcup A_2 \bigcup \{b_1\} \bigcup \{b_2\}$ is complete.

In the rest of the proof we shall show tha B is thin.

Since A_2 is a finite sequence, thus

$$B(x) \leq |A_2| + 2 + A_1(X),$$

which means that if A_1 is thin so is B.

By Lemma 3 for the elements of A_1 we have $\lim_{n\to\infty} a'_{n+1}/a'_n = 2$, which implieses the theorem.

Concluding remarks

We can now apply the theorem for some “classical” sequences which have thin complete subsequences.

We shall investigate the following three type of sequences: let

$$P = \{2 < 3 < \ldots < p_n < \ldots\}$$

be the sequence of prime numbers, let

$$B(p,q) = \{p^k q^m : (p,q) = 1;,\ p,q > 1,\ p,q,m,k \in \mathbb{N}\}$$

be the sequence the Birch-sequence.

Let $g_m(x) \in \mathbb{Z}[x]$. Assume $g_m(x)$ has positive leading coeffitient and

$$\text{g.c.d.}\,\{g_m(1), \ldots, g_m(n), \ldots\} = 1$$

and finally consider the sequence

$$G = \{g_m(1), \ldots, g_m(n), \ldots\}.$$

The sequence of P**.** Richert proved in [8] that $\Delta_P = 7$ and it is well known that $p_{n+1}/p_n \to 1$ as $n \to \infty$.

The sequence $B(p,q)$**.** Erdős conjectured and Birch proved that $B(p,q)$ is complete sequence (see [5]). By the irrationality of $(\log p/\log q)$ we infer that the quotient of consecutive terms of this sequence tends to 1.

The sequence G**.** Finally the completeness of the sequences G were investigated by many authors. In 1948 Sprague proved for the sequence of squares that $\Delta_S = 129$ [6]. Further he proved in [7] that for every k the sequence $\{n^k : n \in \mathbb{N}\}$ is complete. A far-reaching generalization of Birch's and Sprague's results was published by J. W. Cassels (see [4] and (5)). This result gives in the general case that the sequence G is complete. Furthermore since $\lim_{n\to\infty} g_m(n+1)/g_m(n) = 1$ we conclude that for these sequences fulfills the conditions of the Theorem. Hence we obtain the following:

COROLLARY. *The sequences P, $B(p,q)$ and G have thin complete subsequence.*

Certainly there exists complete sequence which has no thin complete subsequence. For instance if $\Phi = \{F_1 < \ldots\}$ where $F_1 = 1$, $F_2 = 2$ is the sequence of Fibonacci then it is well known that Φ is complete and $F(x) = c\log_2 x;\ c > 1$. But if we omit at least two elements from Φ then the remaining sequence cannot be complete (see [5]).

References

[1] E. WIRSING, Thin Subbases, *Analysis,* **6** (1986), 285–308.

[2] J. SPENCER, Four Squares with Few Squares, Number Theory, (New York, 1991–1995) 295–297, Springer Verlag, New York

[3] R. L. GRAHAM, Complete sequences of polynomial values, *Duke Math. J.,* **31** (1964), 275–286.

[4] J. W. CASSELS, On the representation of integers as the sums of distinct summands taken from a fixed set, *Acta Sci. Math.,* **21** (1960), 111–124.

[5] R. L. GRAHAM, *On sums of integers taken from a fixed sequence,* Proc. Wash. Univ. Conf. on Number Theory, (1971) 22–40.

[6] R. SPRAGUE, Über Zerlegungen in ungleiche Quadratzahlen, *Math. Z.,* **51** (1948), 289–290.

[7] R. SPRAGUE, Über Zerlegungen in n-te Potenzen mit lauter verschidenen Grundzahlen, *Math. Z.,* **51** (1948), 466–468.

[8] H. E. RICHERT, Über Zerfällungen in ungleiche Primzahlen, *Math. Z,* (1949), 342–343.

ANNALES UNIV. SCI. BUDAPEST., **44** (2001), 157–170

WARPED PRODUCT OF FINSLER MANIFOLDS

By

LÁSZLÓ KOZMA, RADU PETER and CSABA VARGA

Institute of Mathematics and Informatics, University of Debrecen, Debrecen, Hungary
Technical University of Cluj-Napoca, Cluj-Napoca, Romania
Babeş Bolyai University, Faculty of Mathematics and Informatics, Cluj-Napoca, Romania

(Received March 23, 2002)

1. Introduction

Recently the notion of warped product is playing an important role in Riemannian geometry (see [3, 8, 9, 11, 14, 16]), moreover in geodesic metric spaces [9]. This construction can be extended for Finslerian metrics with some minor restriction. This is motivated by Asanov's papers ([4, 5]) where some models of relativity theory are described through the warped product of Finsler metrics. For example, Asanov [4] studied the property of the generalized Schwarzschild metric on $\mathbb{R} \times M$.

2. Preliminaries

Let M be a real manifold of dimension n and (TM, π, M) the tangent bundle of M. The vertical bundle of the manifold M is the vector bundle $(\mathcal{V}, \overline{\pi}, M)$ given by $\mathcal{V} = \operatorname{Ker} d\pi \subset T(TM)$. (x^i) will denote local coordinates on an open subset U of M, and (x^i, y^i) the induced coordinates on $\pi^{-1}(U) \subset \subset TM$. The radial vector field ι is locally given by $\iota(x, y) = y^a \frac{\partial}{\partial x^a}$.

A Finsler metric on M is a function $F : TM \to \mathbb{R}_+$ satisfying the following properties:

1. F^2 is smooth on $\widetilde{M}$ where $\widetilde{M} = TM \setminus (0)$
2. $F(u) > 0$ for all $u \in \widetilde{M}$
3. $F(\lambda u) = |\lambda| F(u)$ for all $u \in TM, \lambda \in \mathbb{R}$
4. For any $p \in M$ the indicatrix $I_p = \{u \in T_pM |\ F(u) < 1\}$ is strongly convex.

A manifold M endowed with a Finsler metric F is called a Finsler manifold (M, F).

Condition 4. implies that the quantities $g_{ij}(x, y) = \frac{1}{2}\frac{\partial^2 F(x,y)}{\partial y^i \partial y^j}$ forms a positive definite matrix so a Riemannian metric $\langle \cdot, \cdot \rangle$ can be introduced in the vertical bundle $(\mathcal{V}, \widetilde{\pi}, TM)$.

On a Finsler manifold there does not exist, in general, a linear metrical connection. The analogue of the Levi-Civita connection lives just in the vertical bundle, however, there are several ones.

In this paper we use the Cartan connection which is a good vertical connection on $\mathcal{V}$, i.e. an $\mathbb{R}$-linear map

$$\nabla^v : \mathfrak{X}(\widetilde{M}) \times \mathfrak{X}(\mathcal{V}) \to \mathfrak{X}(\mathcal{V})$$

having the usual properties of a covariant derivations, metrical with respect to $\langle \cdot, \cdot \rangle$, and 'good' in the sense that the bundle map $\Lambda : T\widetilde{M} \to \mathcal{V}$ defined by $\Lambda(Z) = \nabla^v_Z \iota$ is a bundle isomorphism when ∇^v is restricted to $\mathcal{V}$. The latter property induces the horizontal subspaces $H_u = \mathrm{Ker}\, \Lambda$ for all $u \in \widetilde{M}$ which are direct summands of the vertical subspaces $V_u = \mathrm{Ker}\,(d\pi)_u$:

$$T\widetilde{M} = \mathcal{H} \oplus \mathcal{V}.$$

For a tangent vector field X on M we have its vertical lift X^V and its horizontal lift X^H to $\widetilde{M}$.

$\Theta : \mathcal{V} \to \mathcal{H}$ denotes the horizontal map associated to the horizontal bundle $\mathcal{H}$. Using Θ, first we get the radial horizontal vector field $\chi = \Theta \circ \iota$. In our case $\sigma^H = \chi(\dot{\sigma})$. Secondly we can extend the covariant derivation ∇^v of the vertical bundle to the whole tangent bundle of $\widetilde{M}$. Denoting it with ∇, for horizontal vector fields we have

$$\nabla_Z H = \Theta(\nabla^v_Z(\Theta^{-1}(H))), \ \forall\, Z \in \mathfrak{X}(\widetilde{M}).$$

An arbitrary vector field $Y \in \mathfrak{X}(\widetilde{M})$ is decomposed into vertical and horizontal parts:

$$\nabla_Z Y = \nabla_Z Y^V + \nabla_Z Y^H.$$

Thus $\nabla : \mathfrak{X}(T\widetilde{M}) \times \mathfrak{X}(T\widetilde{M}) \to \mathfrak{X}(T\widetilde{M})$ is a linear connection on $\widetilde{M}$ induced by a good vertical connection. Its torsion θ and curvature Ω are defined as usual:

$$\nabla_X Y - \nabla_Y X = [X, Y] + \theta(X, Y)$$
$$R_Z(X, Y) = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z$$

and the torsion has the property that for horizontal vectors $\theta(X, Y)$ is a vertical vector [2]. The metrical property of the Cartan connection is also important [2]:

$$X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle.$$

The Cartan connection does not verify the Koszul formula for all vectors, but this formula is true for the horizontal ones, as is shown in the next Lemma:

LEMMA 1. *Let (M, F) be a Finsler manifold with its Cartan connection ∇. For $X, Y, Z \in \mathcal{H}$ the following relation holds:*

$$2\langle \nabla_X Y, Z\rangle =$$
$$= X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle - \langle X, [Y, Z]\rangle + \langle Y, [Z, X]\rangle + \langle Z, [X, Y]\rangle.$$

PROOF. For the first three terms we use the metrical property of the Cartan connection, and for the last three terms we use the relation satisfied by the torsion as follows:

$$\begin{aligned} X\langle Y, Z\rangle &= \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle; \\ Y\langle Z, X\rangle &= \langle \nabla_Y Z, X\rangle + \langle Z, \nabla_Y X\rangle; \\ Z\langle X, Y\rangle &= \langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle; \\ [Y, Z] &= \nabla_Y Z - \nabla_Z Y - \theta(Y, Z); \\ [Z, X] &= \nabla_Z X - \nabla_X Z - \theta(Z, X); \\ [X, Y] &= \nabla_X Y - \nabla_Y X - \theta(X, Y). \end{aligned}$$

Summing up and using the fact that for horizontal vectors $\langle X, \theta(Y, Z)\rangle$ is zero because $\theta(Y, Z)$ is vertical for horizontal vectors Y, Z we obtain the Koszul formula. ∎

We are interested in some properties of the curvature of Cartan connection listed below.

LEMMA 2. *Let (M, F) be a Finsler manifold. The curvature of the Cartan connection satisfies the following properties for horizontal vectors X, Y, Z, V, W:*

1. $R(X, Y) = -R(Y, X)$
2. $\langle R_V(X, Y), W\rangle = -\langle R_W(X, Y), V\rangle$
3. $R_Z(X, Y) + R_X(Y, Z) + R_Y(Z, X) = 0$
4. $\langle R_V(X, Y), W\rangle = \langle R_X(V, W)X, Y\rangle$.

The proof of the previous Lemma can be found in [2, p. 31], and [12, p. 72].

Let P be a submanifold of M of dimension $p < n$ and let us consider $F^* = F|_{TP}$; it is a Finsler metric and thus P becomes a Finsler space. Let $\widetilde{x} \in \widetilde{P}$ and let $P^*_{\widetilde{x}}$ be the $\langle\cdot,\cdot\rangle_{\widetilde{x}}$ orthogonal complement of $T_{\widetilde{x}}TP$ in $T_{\widetilde{x}}TM$. Let $P^{\perp}$ be the disjoint union of all $P^{\perp}_{\widetilde{x}}$, $\widetilde{x} \in \widetilde{P}$ and let $\pi^{\perp} : \widetilde{P}^{\perp} \to \widetilde{P}$ the natural projection. Then $(P^{\perp}, \pi^{\perp}, \widetilde{P})$ admits a natural structure of real differentiable vector bundle, $\operatorname{rank} P^{\perp} = n - p$. It is the normal bundle of the submanifold P.

Let $\widetilde{X}^*, \overline{Y}$ be respectively a tangent vector field on $\widetilde{P}$ and a cross section in $T\widetilde{P}$ and $\widetilde{X}^*, \overline{Y}^*$ prolongations to $T\widetilde{M}$. Then the restriction of $\nabla_{\widetilde{X}^*}\overline{Y}$ to $T\widetilde{P}$ does not depend upon the choice of prolongations and is denoted by $\nabla^*_{\widetilde{X}}\overline{Y}$. The bundle direct sum decomposition

$$T\widetilde{M} = T\widetilde{P} \oplus P^{\perp}$$

leads to the Gauss–Weingarten formulae:

$$\nabla_{\widetilde{X}}\overline{Y} = \nabla^*_{\widetilde{X}}\overline{Y} + \mathbb{I}(\widetilde{X}, \overline{Y})$$

$$\nabla_{\widetilde{X}}\overline{\xi} = -\widetilde{A}_{\overline{\xi}}\widetilde{X} + \nabla^{\perp}_{\widetilde{X}}\overline{\xi}.$$

Here $\xi \in \operatorname{Sec}(\widetilde{P}, P^{\perp})$ and a similar argument (independence of extensions of $\widetilde{X}, \overline{\xi}$ to $T\widetilde{P}$) leads to the notation $\nabla_{\widetilde{X}}\overline{\xi}$. Then ∇^* is the induced connection, $\mathbb{I}$ the second fundamental form, $\widetilde{A}_{\overline{\xi}}$ the operators of Weingarten and $\nabla^{\perp}$ is the normal connection ([7, 1, 10]). Next we define the umbilical point of a Finsler submanifold and the umbilical submanifold.

DEFINITION 3. A point $q \in P$ is an umbilical point if there exists a vector $Z \in \mathcal{H}^{\perp}(P)$ such that $\mathbb{I}(X, Y) = \langle X, Y\rangle Z$. The submanifold P is said to be totally umbilical if every point of P is an umbilical point.

3. Construction of the warped product

Let (M, F_1) and (N, F_2) be Finsler manifolds with their Cartan connections ∇^1 and ∇^2, and let $f : M \to \mathbb{R}_+$ be a smooth function. Let $p_1 : M \times N \to M$, and $p_2 : M \times N \to N$. We consider the product manifold $M \times N$ endowed with the metric $F : \widetilde{M} \times \widetilde{N} \to \mathbb{R}$,

$$F(v_1, v_2) = \sqrt{F_1^2(v_1) + f^2(\pi_1(v_1))F_2^2(v_2)}.$$

We show that the metric defined above is a Finsler metric. First it is clear that F is smooth on $\widetilde{M} \times \widetilde{N}$, because F_1 and F_2 are. F is not necessarily smooth at the vectors of the form $(v_1, 0)$ and $(0, v_2) \in TM \times TN$. This means that F is not a really Finsler metric on the product manifold $M \times N$, therefore the study should be restricted to the domain $\widetilde{M} \times \widetilde{N}$. Secondly F is homogeneous with respect to the vector variables because F_1 and F_2 are. Third, the Hessian of F with respect to the vector variables is of the form:

$$\begin{pmatrix} A & 0 \\ 0 & f^2 B \end{pmatrix}$$

where A and B are the Hessians of the Finsler metrics F_1 and F_2. So the Hessian of F is positive because the Hessians of F_1 and F_2 are. It means that the indicatrix of F is strongly convex. The difference between this metric and a classical Finsler metric is that it is not smooth at the vectors of the form $(v_1, 0)$ and $(0, v_2)$.

The product manifold $M \times N$ with the metric $F(v) = F(v_1, v_2)$, for $v = (v_1, v_2) \in \widetilde{M} \times \widetilde{N}$ defined above will be called the warped product of the manifolds M, N, and f will be called the warping function. We denote this warped product by $M \times_f N$. We just showed that $(M \times_f N, F)$ is a Finsler manifold in the restricted sense above.

Our goal is to express the geometry of warped product by the geometries of M, N and the warping function f. The study follows the line adopted in Riemannian and semi-Riemannian cases [13], with the specific situation due to the Finslerian context.

The manifold M will be called base and the manifold N will be called fiber as in [13].

4. The gradient of a function in Finsler geometry

In this section we define the gradient of the smooth function $f : M \to \mathbb{R}_+$ with $df_x \neq 0$. We follow the line of SHEN [15, p. 43]. Define ∇f_x by

$$\nabla f_x := L_x^{-1}(df_x)$$

where $L_x : T_x M \to T_x^* M$ is the Legendre transformation. Shen proves that

$$\nabla f^H = \widehat{\nabla} f$$

where $\widehat{\nabla} f$ is the gradient of f with respect to Riemannian metric induced by the Finsler metric, and

$$F(\nabla f) = \sqrt{\langle \widehat{\nabla} f, \widehat{\nabla} f \rangle_{\nabla f}}.$$

We work with ∇f^H, the horizontal lifting of ∇f which has the property that $F^2(\nabla f) = \langle \nabla f^H, \nabla f^H \rangle_{\nabla f^H}$.

Next we define the Hessian of a function.

DEFINITION 4. The Hessian of a function $f \in \mathcal{F}(M)$ is its second covariant differential $\mathcal{H}^f = \nabla(\nabla f)$.

LEMMA 5. *The Hessian $\mathcal{H}^f$ satisfies the following relation:*

$$\mathcal{H}^f(X, Y) = XYf - (\nabla_X Y)f = \langle \nabla_X(\nabla f^H), Y \rangle$$

for $X, Y \in \mathcal{H}$.

PROOF.

$$\mathcal{H}^f(X, Y) = \nabla(df^H)(X, Y) = \langle \nabla_X \nabla f^H, Y \rangle$$

since $Yf = \langle \nabla f^H, Y \rangle$ and it follows that

$$\begin{aligned} XYf = X\langle \nabla f^H, Y \rangle &= \langle \nabla_X \nabla f^H, Y \rangle + \langle \nabla f^H, \nabla_X Y \rangle \\ &= \langle \nabla_X(\nabla f^H), Y \rangle + (\nabla_X Y)f \end{aligned}$$

which implies the assertion. ∎

If f is smooth on M (i.e. $f : M \to \mathbb{R}$ is smooth), the lift of f to $M \times N$ is the map $\widehat{f} := f \circ p_1 : M \times N \to \mathbb{R}$. If $a \in T_pM$ and $q \in N$ then the lift $\widehat{a}$ of a to (p, q) is the unique vector in $T_{(p,q)}(M \times q)$ such that $dp_1(\widehat{a}) = a$. If $X \in \mathfrak{X}(M)$ the lift of X to $M \times N$ is the vector field $\widehat{X}$ whose value at each (p, q) is the lift of X_p to (p, q). Because of the product coordinate systems it is clear that $\widehat{X}$ is smooth. It follows that the lift of $X \in \mathfrak{X}(M)$ is the unique element of $\mathfrak{X}(M \times N)$ that is p_1-related to X and p_2-related to the zero vector field on N. The same method could be used to lift objects defined on N to $M \times N$.

Now we prove a Lemma needed in what follows:

LEMMA 6. *If h is a smooth function on M, then the gradient of the lift $h \circ p_1$ of h to $M \times_f N$ is the lift to $M \times_f N$ of the gradient of h on M.*

PROOF. Let $v \in TN$. Now $\langle \nabla(h \circ p_1), v^H \rangle = v^H(h \circ p_1) = 0$. Next for $x \in TM$ we have that

$$\begin{aligned} &\langle d\widetilde{p}_1((\nabla(h \circ p_1))^H), d\widetilde{p}_1(x) \rangle = \\ &= \langle (\nabla(h \circ p_1))^H, x^H \rangle = (x(h \circ p_1))^H = \langle (\nabla h)^H, dp_1(x)^H \rangle. \end{aligned}$$

From these two properties we obtain the assertion in the lemma. ∎

Due to this lemma there will be no confusion if we denote h and ∇h instead of $h \circ p_1$ and $\nabla(h \circ p_1)$, resp.

5. Properties of warped metrics

Let (M, F_1) and (N, F_2) be two Finsler manifolds, with Finsler metrics F_1, F_2 resp. We consider the product manifold $M \times N$ and the warped metric defined above. We consider the projections $p_1 : M \times N \to M$ and $p_2 : M \times \times N \to N$ and the canonical projections $\pi_1 : TM \to M$ and $\pi_2 : TN \to N$. The projections p_1, p_2 resp. generate the projections $dp_1 : TM \times TN \to TM$ and $dp_2 : TM \times TN \to TN$, for $v = (v_1, v_2) \in TM \times TN, dp_i(v_1, v_2) = v_i$, $i = 1, 2$.

It is obvious that the fibers $p \times N = p_1^{-1}(p), p \in M$ and the leaves $M \times q = p_2^{-1}(q), q \in N$ are Finsler submanifolds of $M \times_F N$ and the warped metric has the properties:

1. for each $q \in N$ the map $p_1|_{(M \times q)}$ is an isometry onto M.
2. for each $p \in M$ the map $p_2|_{(p \times N)}$ is a positive homothety onto N with scale factor $\frac{1}{f}$.
3. for each $(p, q) \in M \times N$ the leaf $M \times q$ and the fiber $p \times N$ are orthogonal with respect to the Riemannian metrics induced by the Finsler metrics.

The canonical projection π_1 gives rise to the vertical bundle

$$(\mathcal{V}_1, \widetilde{\pi_1}, TM),$$

where $\mathcal{V}_1 = \operatorname{Ker}(d\pi_1)$ and $\widetilde{\pi_1} = d\pi_1 : TTM \to TM$. The same is true for the manifold N. Now we have that

$$d\pi_1 \times d\pi_2 = d(\pi_1 \times \pi_2) : TTM \times TTN = T(TM \times TN) \to TM \times TN$$

and $\operatorname{Ker} d(\pi_1 \times \pi_2) = \operatorname{Ker} d\pi_1 \oplus \operatorname{Ker} d\pi_2$. It follows that the vertical space of the manifold $M \times N$, $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2$, so the Riemannian metrics $\langle \cdot, \cdot \rangle^1$ and $\langle \cdot, \cdot \rangle^2$, defined on $\mathcal{V}_1$ and $\mathcal{V}_2$ as in the introduction give rise to a Riemannian metric $\langle \cdot, \cdot \rangle$ on $\mathcal{V}$ as follows: $\langle \cdot, \cdot \rangle_v = \langle \cdot, \cdot \rangle^1_{v_1} + f^2(\pi_1(v_1))\langle \cdot, \cdot \rangle^2_{v_2}$. Now let $\mathcal{H}_1$ and $\mathcal{H}_2$ be the horizontal spaces with respect to the Cartan connections ∇^1 and ∇^2 on the Finsler manifolds (M, F_1) and (N, F_2), resp.

We have the direct sum decomposition

$$TT(M \times N) = TTM \oplus TTN = \mathcal{V}_1 \oplus \mathcal{H}_1 \oplus \mathcal{V}_2 \oplus \mathcal{H}_2.$$

Next the Finsler metrics F_1, F_2 on the manifolds M and N resp. generate the Riemannian metrics $\langle , \rangle^1$ and $\langle , \rangle^2$ on the vertical spaces $\mathcal{V}_1$ and $\mathcal{V}_2$, resp. By the horizontal maps these Riemannian metrics are mapped onto horizontal spaces $\mathcal{H}_1, \mathcal{H}_2$ resp. Finally these Riemannian metrics generates a Riemannian metric on $T(TM \times TN)$. In what it follows we work mostly on the direct sum $\mathcal{H}_1 \oplus \mathcal{H}_2$ the direct sum of the liftings of $\mathcal{H}_1$ and $\mathcal{H}_2$ to the $TTM \times TTN$.

The following theorem relates the Cartan connections of M and N to the Cartan connection of $M \times_f N$.

THEOREM 7. *On $B = M \times_f N$ if $X, Y \in \mathfrak{X}(\mathcal{H}_1)$ and $V, W \in \mathfrak{X}(H_2)$ the following relations are true:*

1. $\nabla_X Y$ *on* $\mathcal{H}_1 \oplus \mathcal{H}_2$ *is the lift of* $\nabla_X Y$ *on* $\mathcal{H}_1$.
2. $\nabla_X V = \nabla_V X = (Xf/f)V$.
3. $\operatorname{nor} \nabla_V W = \mathbb{I}(V, W) = -(\langle V, W\rangle / f)\nabla f^H$.
4. $\theta(X, V) = \theta(V, X) = 0$.
5. $\tan \nabla_V W \in \mathfrak{X}(N)$ *is the lift of* $\nabla_V W$ *on* N.

PROOF. We apply the Koszul formula (see Lemma 1) for $2\langle \nabla_X Y, V\rangle$ and we obtain that it is equal to $-V\langle X, Y\rangle + \langle V, [X, Y]\rangle$ because $[X, V] = [Y, V] = 0$. Because X, Y are lifts from M, $\langle X, Y\rangle$ is constant on fibers (liftings on N), and because $V \in T\widetilde{N}$ follows that $V\langle X, Y\rangle = 0$. Analogously $\langle V, [X, Y]\rangle = 0$. Thus $\langle \nabla_X Y, V\rangle = 0$ for all $V \in \mathfrak{X}(N)$ and it follows formula (1).

First we prove the first equality from (2). The second one will be proved after (3). We have that $X\langle V, Y\rangle = \langle \nabla_X V, Y\rangle + \langle V, \nabla_X Y\rangle = 0$, so $\langle \nabla_X V, Y\rangle = -\langle V, \nabla_X Y\rangle$. We apply the Koszul formula for $2\langle \nabla_X V, W\rangle$, and we observe that all the terms vanish except $X\langle V, W\rangle$.

It follows from the expression of the Riemannian metric induced by the warped metric that $\langle V, W\rangle(v, w) = f^2(\pi_1(v))\langle V_w, W_w\rangle$. This term is constant on leaves. Thus $X\langle V, W\rangle = X(f^2(\pi_1(v))\langle V_w, W_w\rangle) = 2f X(f(\pi_1(v)))\langle V_w, W_w\rangle = 2\left(\frac{Xf}{f}\right)\langle V, W\rangle$. From these relations we have that $\nabla_X V = \left(\frac{Xf}{f}\right) V$. Now $\nabla_X V - \nabla_V X = [X, V] + \theta(X, V)$. We can assume that $[X, V] = 0$.

It is obvious that $V\langle W, X\rangle = 0$. But this means that

$$\langle \nabla_V W, X\rangle = -\langle W, \nabla_V X\rangle = -\langle W, (Xf/f)V + \theta(X, V)\rangle = -(Xf/f)\langle V, W\rangle$$

because $\theta(X, V)$ is vertical. Now $\langle \nabla f^H, X\rangle = Xf$. Thus

$$\langle \nabla_V W, X\rangle = -\langle (\langle V, W\rangle/f)\nabla f^H, X\rangle.$$

This yields (3).

$$\begin{aligned}\langle \nabla_V X, W\rangle &= -\langle X, \nabla_V W\rangle = -\langle X, \langle V, W\rangle/f \nabla f^H\rangle \\ &= \frac{1}{f}\langle X, \nabla f^H\rangle\langle V, W\rangle = \langle \langle X, \nabla f^H\rangle/f\, V, W\rangle.\end{aligned}$$

The above gives the second part of (2) and it follows that

$$\nabla_V X = \nabla_X V = \left(\frac{Xf}{f}\right) V,$$

and the mixed part of the torsion vanishes $\theta(X, V) = \theta(V, X) = 0$. The last assertion (5) is trivial. ∎

It is a remarkable fact that the torsion vanishes on the mixed part. This will let us to compute the curvature of warped product.

Now the next Corollary easily follows:

COROLLARY 8. *The leaves $M \times q$ of a warped product are totally geodesic; the fibers $p \times M$ are totally umbilical.*

PROOF. By the claim (1) in Theorem 7 it follows that for a geodesic α in M its lifting on $M \times_f N$ is also a geodesic. The second assertion comes from (3) of Theorem 7. ∎

6. Geodesics of warped product manifolds

In a warped product manifold a curve γ can be written as $\gamma(s) = (\alpha(s), \beta(s))$ where the curves α and β are the projections of γ into M and N, resp. Now we give conditions for a curve in the warped product to be geodesic with respect to the warped metric.

THEOREM 9. *A curve $\gamma = (\alpha, \beta)$ in $M \times_f N$ is a geodesic if and only if*

1. $\nabla_{\alpha'^H} \alpha'^H = \frac{\|\beta'^H\|^2}{f} \nabla f^H$,
2. $\nabla_{\beta'^H} \beta'^H = \frac{-2}{f \circ \alpha} \frac{(d(f \circ \alpha))^H}{ds} \beta'^H$

PROOF. We work in an interval around $s = 0$.

Case 1. $\gamma'(0)$ is neither in $T_{\alpha(0)}M$ nor in $T_{\beta(0)}N$. Then $\alpha'(0) \neq 0$ and $\beta'(0) \neq 0$. So we can suppose that α is an integral curve for X in M and β is an integral curve for V in N. Also we denote by X and V the lifts on $M \times_f N$. It follows that γ is a geodesic curve if and only if $\nabla_{X^H + V^H}(X^H + V^H) = 0$. But this means that

$$\nabla_{X^H} X^H + \nabla_{X^H} V^H + \nabla_{V^H} X^H + \nabla_{V^H} X^H = 0.$$

Now we use Theorem 7 from the previous section and we have that

$$\nabla_{X^H} X^H - \frac{\|V^H\|^2}{f} \nabla f^H = 0$$

and

$$2\frac{X^H f}{f} V + \nabla_{V^H} V^H = 0.$$

Case 2. Suppose that $\gamma'(0) \in T_{\alpha(0)}M$. If γ is a geodesic, because $M \times \beta(0)$ is totally geodesic, it follows that γ remains in $M \times \beta(0)$. Thus β is constant and the assertions of the theorem are trivial. Conversely if condition (2) from Theorem 7 holds, since $\beta'(0) = 0$ it follows that β is constant. Then condition (1) in Theorem 7 implies that α is a geodesic, and so is γ.

Case 3. Suppose that $\gamma'(0) \in T_{\beta(0)}N$ and nonzero. Suppose that ∇f is not zero, because otherwise $\alpha(0) \times N$ is totally geodesic and the conclusion follows as in *Case 1*. Now if γ is a geodesic, it follows that on no interval around 0 γ remains in the totally umbilical fiber $p \times N$. It follows that there is a sequence $\{s_i\} \to 0$ such that for all i, $\gamma'(s_i)$ is neither in $T_{\alpha(s_i)}M$ or in $T_{\beta(s_i)}N$. The assertions in the theorem follows by continuity from the first case. Conversely, if (1) in the theorem is true it follows that $\nabla_{\alpha'(0)^H}\alpha'(0)^H \neq 0$ hence there exists a sequence $\{s_i\}$ as above, and using again the first case it follows that γ is a geodesic. ∎

7. Curvature of warped product manifolds

Now we express the curvature of the warped product. The curvature tensor is defined by the relation

$$R_Z(X,Y) = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z.$$

Because the projection p_1 is an isometry it follows that the lift of the curvature on M is equal to the curvature of the warped product when is computed for vectors from on $\mathcal{H}_1$.

THEOREM 10. *Let $M \times_f N$ be a warped product of Finsler manifolds with curvature tensor R and let $X, Y, Z \in \mathcal{H}_1$ and $U, V, W \in \mathcal{H}_2$. Let R^M_Z and R^N_U denote the curvature tensors of the manifolds (M, F_1) and (N, F_2) resp. The following relations are true:*

1. $R_Z(X,Y) \in \mathfrak{X}(\mathcal{H}_1)$ *is the lift of* $R^M_Z(X,Y)$ *on* M.
2. $R_Y(V,X) = -\left(\frac{H^f(X,Y)}{f}\right)V$, *where* H^f *is the Hessian of* f.
3. $R_X(V,W) = (Xf/f)\theta(V,W)$.
4. $R_W(X,V) = \left(\frac{\langle V,W\rangle}{f}\right)\nabla_X(\nabla f)$.
5. $R_U(V,W) = R^N_U(V,W) - \left(\frac{\langle\nabla f,\nabla f\rangle}{f^2}\right)\{\langle V,U\rangle W - \langle W,U\rangle V\}$.

PROOF. (1) This is true because the projection p_1 is an isometry and the leaves are totally geodesic.

(2) Because $[V,X] = 0$ it follows that $\nabla_V\nabla_X Y - \nabla_X\nabla_V Y = R_Y(V,X)$. By Theorem 7 we have that $\nabla_V\nabla_X Y = \left(\frac{(\nabla_X Y)f}{f}\right)V$ because

$\nabla_X Y \in \mathfrak{X}(\mathcal{H}_1)$. The second term

$$\nabla_X \nabla_V Y = \nabla_X \left(\frac{Yf}{f} V \right) = X(Yf/f)V + (Yf/f)\nabla_X V$$
$$= [(XY)f/f + Yf X(1/f)]V + (Yf/f)(Xf/f)V.$$

Because $X(1/f) = -Xf/f^2$ the last expression reduces to $(XYf/f)V$. Thus

$$R_Y(V, X) = -[(XYf - (\nabla_X Y)f)/f]V = -(H^f(X, Y)/f)V.$$

(3) We can assume that $[V, W] = 0$. It follows that

$$R_X(V, W) = \nabla_V \nabla_W X - \nabla_W \nabla_V X.$$

But

$$\nabla_V \nabla_W X = \nabla_V((Xf/f)W) = V(Xf/f)W + (Xf/f)\nabla_V W.$$

Now $V(Xf/f) = 0$ because Xf/f is constant on the fibers. This implies that

$$R_X(V, W) = (Xf/f)[\nabla_V W - \nabla_W V] = (Xf/f)\theta(V, W).$$

We note that $R_X(V, W) \in \mathcal{V}_2$ by the properties of the Cartan connection.

By the symmetry of curvature $\langle R_V(X, Y), W\rangle = \langle R_X(V, W), Y\rangle = 0$ because $R_X(V, W)$ is vertical. Now we use (2), the curvature symmetries, and then we obtain that relation (3) is true.

(4) We have that $\langle R_W(X, V), U\rangle = \langle R_X(W, U), W\rangle = 0$ because of the point above. We use here the properties from Lemma 2. Now $R_X(V, W)$ is vertical and it follows that

$$\langle R_W(V, X), Y\rangle = \langle R_Y(V, X), W\rangle = H^f(X, Y)\langle V, W\rangle$$
$$= (\langle V, W\rangle/f)\langle \nabla_X(\nabla f), Y\rangle,$$

which gives assertion (4).

(5) Again we can assume that $[U, V]$ is zero.

$$
\begin{aligned}
&R(V, W) = \\
&= \nabla_V \nabla_W U - \nabla_W \nabla_V U = \nabla_V \{-(\langle W, U\rangle/f)\nabla f^H + \nabla^N_V U\} \\
&\quad - \nabla_W \{-(\langle V, U\rangle/f)\nabla f^H + \nabla^N_V U\} = -(\langle \nabla_V W, U\rangle \\
&\quad + \langle W, \nabla_V U\rangle)(\nabla f^H/f) - (\langle W, U\rangle/f)\nabla_V(\nabla f^H) \\
&\quad + \nabla_V \nabla^N_W U + (\langle \nabla_W V, U\rangle + \langle V, \nabla_W U\rangle)(\nabla f^H/f) \\
&\quad + (\langle V, U\rangle/f)\nabla_W(\nabla f^H) - \nabla_W \nabla^N_V U = (\langle \nabla_W V - \nabla_V W, U\rangle \\
&\quad - \langle W, \nabla_V U\rangle - \langle V, \nabla_W U\rangle)(\nabla f^H/f) + \nabla^N_V \nabla^N_W U - \nabla^N_W \nabla^N_V U \\
&\quad - (\langle V, \nabla^N_W U\rangle/f)\nabla f^H + (\langle W, \nabla^N_V U\rangle)(\nabla f^H) \\
&\quad + (\langle V, U\rangle/f)(\langle \nabla f^H, \nabla f^H\rangle/f) - (\langle W, U\rangle/f)(\langle \nabla f^H, \nabla f^H\rangle/f)V \\
&= R^N(V, W)U + \frac{\langle \nabla f^H, \nabla f^h\rangle}{f^2}(\langle V, U\rangle W - \langle W, U\rangle V).
\end{aligned}
$$

We use that $\langle V, \nabla_W U\rangle = \langle V, \nabla^N_W U\rangle$, and the properties from Theorem 7. Thus we have

$$
R_U(V, W) = R^N_U(V, W) + \left(\frac{\langle \nabla f^H, \nabla f^h\rangle}{f^2}\right)(\langle V, U\rangle W - \langle W, U\rangle V). \quad \blacksquare
$$

References

[1] L. MARIA ABATANGELO, SORIN DRAGOMIR and SHUN ICHI HOJO, On submanifolds of Finsler Spaces, *Tensor, N.S.*, **47** (1988), 272–285.

[2] M. ABATE and G. PATRIZIO, *Finsler Metrics – A Global Approach*, Lecture Notes in Mathematics, 1591, Springer Verlag, Berlin, Heidelberg, 1994.

[3] STEPHANIE B. ALEXANDER, RICHARD L. BISHOP, Warped products of Hadamard spaces, *Manuscripta Math.,* **96** (1998), 487–505.

[4] G. S. ASANOV, Finslerian metric functions over the product $R \times M$ and their potential applications, *Rep. Math. Phys.,* **41** (1998), 117–132.

[5] G. S. ASANOV, Finslerian extensions of Schwarzschild metric, *Fortschr. Phys.,* **40** (1992), 667–693.

[6] D. BAO, S.S. CHERN, Z. SHEN, *An Introduction to Riemann Finsler Geometry*, Springer Verlag, 2000.

[7] AUREL BEJANCU, On the theory of Finsler submanifolds, *Meetings of Minds*, eds. P. L. Antonelli, Kluwer Academic Press, 1999, 111–129.

[8] BANG-YEN CHEN, Geometry of warped product CR-submanifolds in Kaehler manifolds, *Monatsh. Math.,* **133**, (2001), 177–195.

[9] CHIEN-HSIUNG CHEN, Warped products of metric spaces of curvature bounded from above, *Trans. Amer. Math. Soc.*, **351** (1999), 4727–4740.

[10] S. DRAGOMIR, Submanifolds of Finsler Spaces. *Conf. Sem. Mat. Univ. Bari*, **271** (1986), 1–15.

[11] BYUNG HAK KIM, Warped producs with critical Riemannian metric, *Proc. Japan Acad. Ser. A Math. Sci.,* **71** (1995), 117–118.

[12] M. MATSUMOTO, *Foundations of Finsler geometry and special Finsler spaces*, Kasheisha Press, Japan, 1986.

[13] BARRETT O'NEILL, *Semi-Riemannian geometry with applications to relativity*, Pure and Applied Mathematics. Academic Press, New-York, 1983.

[14] S. NÖLKER, Isometric immersions of warped products, *Diff. Geom. Appl.,* **6** (1996), 1–30.

[15] Z. SHEN, *Lectures on Finsler Geometry*, World Scientific, 2001.

[16] M ULANOVSKII, Lorentzian warped products and singularity, *Gen. Relativity Gravitation*, **31** (1999), 1813–1820.

ANNALES UNIV. SCI. BUDAPEST., **44** (2001), 171–182

$\mathcal{K}$-STRUCTURES AND FOLIATIONS

By

LUIGIA DI TERLIZZI, JERZY KONDERAK, ANNA MARIA PASTORE
and ROBERT WOLAK

Dipartimento di Matematica, Università di Bari, Italy
Instytut Matematyki, Uniwersytet Jagielloński, Poland

(Received March 28, 2002)

1. Preliminaries

Let M be $2n+s$ dimensional manifold on which is defined an f-structure of rank $2n$ with complemented frames. This means that there exist vector fields $\xi_1,\dots,\xi_s$ on M such that if $\eta_1,\dots,\eta_s$ are dual 1-forms then

$$f(\xi_i)=0 \tag{1}$$

$$\eta_i\circ f=0 \tag{2}$$

for any $i=1,\dots,s$ and

$$f^2=-I+\sum_{i=1}^{s}\xi_i\otimes\eta_i. \tag{3}$$

Let $\Gamma(TM)$ be the module of differentiable sections of TM. It is well known that in such conditions we can define a Riemannian metric g on M such that for any $X,Y\in\Gamma(TM)$ the following equality holds:

$$g(X,Y)=g(fX,fY)+\sum_{i=1}^{s}\eta_i(X)\eta_i(Y). \tag{4}$$

We suppose also that the f-structure is a $\mathcal{K}$-structure, i.e. $[f,f]+\sum_{i=1}^{s}\xi_i\otimes d\eta_i=0$, where $[f,f]$ is the Nijenhuis torsion of f (cf. [2]) and the fundamental 2-form, F defined as $F(X,Y)=g(X,fY)$ is closed, i.e. $dF=0$.

If $d\eta_1=\dots=d\eta_s=F$ and $\eta_1\wedge\dots\wedge\eta_s\wedge(d\eta_i)^n\neq0$ we say that the $\mathcal{K}$-structure is an $\mathcal{S}$-structure and M is an $\mathcal{S}$-manifold. Finally, if $d\eta_i=0$ for all $i=1,\dots,s$ then the $\mathcal{K}$-structure is called a $\mathcal{C}$-structure and M is said a $\mathcal{C}$-manifold.

We recall some facts that will be used in the sequel (cf. [2]).

1. On a $\mathcal{K}$-manifold M the vector fields $\xi_1, \ldots, \xi_s$ are Killing.
2. If f is an $\mathcal{S}$-structure then for any $Y \in \Gamma(TM)$ and for any $i = 1, \ldots, s$ we have that

$$\nabla_Y \xi_i = -\frac{1}{2} f Y; \tag{5}$$

if f is a $\mathcal{C}$-structure then

$$\nabla_Y \xi_i = 0 \tag{6}$$

where we denote by ∇ the Levi–Civita connection of the Riemannian metric g.

3. A $\mathcal{K}$-structure is a $\mathcal{C}$-structure if and only if $\nabla F = 0$ or $\nabla f = 0$.
4. On $\mathcal{S}$-manifolds we have:

$$(\nabla_X F)(Y, Z) = \frac{1}{2} \sum_{i=1}^{s} \left[\eta_i(X) g(X, Z) - \eta_i(Z) g(X, Y)\right]$$
$$-\frac{1}{2} \sum_{i,j=1}^{s} \eta_j(X) \left[\eta_i(Y)\eta_j(Z) - \eta_i(Z)\eta_j(Y)\right].$$

2. Transversally holomorphic foliations

THEOREM 1. *Let M be a $2n+s$-manifold with an f-structure of rank $2n$. Then f is a $\mathcal{K}$ if and only if the foliation* $\ker f$ *is a transversely Kähler foliation given by an isometric action of $\mathbb{R}^s$.*

PROOF. Let f be a $\mathcal{K}$ structure. We know that the vector fields $\xi_1, \xi_2, \ldots, \xi_s$ are Killing vector fields and $[\xi_i, \xi_j] = 0$, $i, j = 1, 2, \ldots, s$. Therefore the subbundle spanned by $\xi_1, \xi_2, \ldots, \xi_s$ is integrable and defines the foliation $\mathcal{F}_\xi$. In fact, the condition for the $\mathcal{K}$-structure, i.e.

$$0 = \{[f, f] + \sum_{i=1}^{s} \xi_i \otimes d\eta_i\} = 0$$

easily implies that $[\xi_i, \xi_j] = 0$ for any $i, j = 1, \ldots, s$. Moreover, the fact that the vector fields ξ_i are Killing ensures that $\nabla_{\xi_i} \xi_j = 0$ for any $i, j = 1, \ldots, s$. As the vector fields ξ_i are Killing the foliation $\mathcal{F}_\xi$ is Riemannian and the Riemannian metric g is bundle-like. The normal bundle of $\mathcal{F}_\xi$ can be identified with Imf. First, notice that Imf $\perp \mathcal{F}_\xi$. In fact,

$$g(\xi_i, f(X)) = g(f(\xi_i), f^2(X)) + \sum_{j=1}^{s} \eta_j(\xi_i)\eta_j(f(X)) = 0$$

Moreover, rank(Imf) = codim $\mathcal{F}_\xi$ so indeed Imf is the normal bundle of $T\mathcal{F}_\xi$. Now we should prove that $f|_{\mathrm{Imf}}$ is foliated, which is equivalent to the property

$$L_{\xi_i}(f|_{\mathrm{Imf}}) = 0 \quad \forall i = 1, \ldots, s$$

i.e. $L_{\xi_i} f(Y) = 0$ for any section of Imf. We may assume that Y is an infinitesimal automorphism of $\mathcal{F}_\xi$. First compute $\{[f,f] + \sum_{i=1}^{s} \xi_i \otimes d\eta_i\} = 0$ on ξ_r and an infinitesimal automorphism X of $\mathcal{F}_\xi$, a section of Imf. Thus $[\xi_r, X]$ is a section of $T\mathcal{F}_\xi$. Therefore

$$\{[f,f] + \sum_{i=1}^{s} \xi_i \otimes d\eta_i\}(\xi_r, X) = f([\xi_r, f(X)]) + \sum_{i=1}^{s} d\eta_i(\xi, X)\xi_i.$$

Hence $f([\xi_r, fX]) = 0$ and $d\eta_i(\xi_r, X) = 0$. So $[\xi_r, f(X)] \in T\mathcal{F}_\xi$ for any X an infinitesimal automorphism of $\mathcal{F}_\xi$. Let Y be an infinitesimal automorphism of $\mathcal{F}_\xi$. Then:

$$L_{\xi_i} f(f(Y)) = [\xi_i, -Y] - f([\xi_i, f(Y)] = 0$$

In fact, we may assume that Y is an infinitesimal automorphism commuting with ξ_i as the subbundle Imf is ξ_i-invariant. So $f|$Imf is constant along leaves of $\mathcal{F}_\xi$ ("foliated"). Moreover, the forms $d\eta_i$ are base-like as $d\eta_i(\xi_j, \ldots) = 0$. So is the form F as

$$F(\xi_i, Y) = 0, \quad g(\xi_i, f(Y)) = 0$$

and $dF = 0$. So our foliation $\mathcal{F}_\xi$ is transversely Kähler as the metric g, tensor field f and the 2-form F project along leaves to the Riemannian metric $\overline{g}$, tensor field $\overline{J}$ and the 2-form $\overline{\Omega}$, respectively. The structure $(\overline{g}, \overline{J}, \overline{\Omega})$ is Kählerian.

Now, assume that we have an $\mathbb{R}^s$ isometric action on M which is transversely Kählerian. Then we have Killing vector fields $\xi_1, \ldots, \xi_s$ pair-wise commuting which define a transversely Kähler foliation $\mathcal{F}_\xi$. As $\xi_1, \ldots, \xi_s$ are Killing vector fields they leave invariant the subbundle Q orthogonal to $T\mathcal{F}_\xi$. We define the tensor field forms η_i as follows:

$$\eta_i(\xi_j) = \delta_i^j, \quad \eta_i|_Q = 0 \;\; f(\xi_i) = 0.$$

On Q the tensor field f is defined as follows: let $f_i : U_i \to N$ be a local submersion defining the foliation $\mathcal{F}_\xi$; then $d_x f_i|_Q : Q_x \to T_{f_i(x)}N$ is an isomorphism for any $x \in U_i$. If $X \in Q_x$ then $d_x f_i(X)$ by $\overline{X}$; if $X \in T_{f_i(x)}N$ we denote $(dxf_i|_Q)^{-1}$ by $\hat{X}$. With this notation in mind we put

$$f(X) = \widehat{J(\overline{X})}$$

for any $X \in Q$. So for any $X \in Q_x$ $f(X) \in Q_x$ as well. If necessary we can modify the Riemannian metric by assuming on Q the following values:

$$g(X, Y) = \overline{g}(\overline{X}, \overline{Y})$$

for all $X, Y \in Q$. With $\xi_1, \ldots, \xi_s$, $\eta_1, \ldots, \eta_s$, g and f defined as above it is straightforward to verify that they define a $\mathcal{K}$-structure. ∎

EXAMPLE 1. We are going to use the well-known construction of a suspension to produce examples of $\mathcal{K}$-manifolds, cf. [11]. Let $(F, \bar{g}, \bar{J})$ be a compact Kähler manifold. Let

$$h : \pi_1(T^s) = \mathbb{Z}^s \to Iso(F, \bar{g}, \bar{J})$$

be a homomorphism of groups, which is equivalent to choosing s commuting holomorphic isometries of $(F, \bar{g}, \bar{J})$. The group $\mathbb{Z}^s$ acts on the product $\mathbb{R}^s \times F$ as follows. Let $p \in \mathbb{Z}^s$, $v \in \mathbb{R}^s$, $w \in F$, $\phi(p)(v, w) = (v + p, h(p)(w))$. The action ϕ is locally free and commutes with the standard action of R^s. If we endow $\mathbb{R}^s \times F$ with the product metric $\tilde{g} = g_0 \times \bar{g}$, ($g_0$-the Euclidean metric of $\mathbb{R}^s$), then the action ϕ is isometric. The quotient manifold $T^s \times_h F$ is a compact fibre bundle over T^s with standard fibre F. The Riemannian metric $\tilde{g}$ defines a Riemannian metric g on $T^s \times_h F$ for which the induce $\mathbb{R}^s$-action is isometric. Moreover, the foliation defined by this action is transversely Kähler, so any such manifold $T^s \times_h F$ is equipped with a $\mathcal{K}$-structure.

We finish the section with a very useful proposition whose proof is straightforward.

PROPOSITION 1. *Let W be a foliated submanifold of M (i.e. if $x \in W$, then the leaf $L_x \subset W$). Let $\{U_i, f_i, g_{ij}\}_{i \in I}$ be a cocycle defining foliation $\mathcal{F}_\xi$ and N the transverse manifold, H the holonomy pseudogroup associated to this cocycle. Then there exists W_0 an H-invariant submanifold of N such that $W|_{U_i} = f_i^{-1}(W_0)$ for any index $i \in I$.*

This proposition can be used to find properties of geodesics orthogonal to $\mathcal{F}_\xi$. In fact, the submanifold W_0 is totally geodesic iff W is totally geodesic in the orthogonal direction to $\mathcal{F}_{\xi_i}$.

3. General properties

Having proved the fact that our structure is a transversely Kähler foliation let us draw some general conclusions.

The set of points of leaves without holonomy is open and dense in the manifold M, cf. [14]. Unless all leaves are compact, there is no compact leaf among them. If the foliation has a compact leaf without holonomy, then this leaf covers any other leaf, thus all leaves are compact, cf. [13, 10, 17]. Leaves without holonomy are stable, which means that for any leaf L without holonomy there exists an $\epsilon > 0$ such that any leaf at the distance smaller than ϵ from L is diffeomorphic to L.

The holonomy pseudogroup of our foliations consists of local isomorphisms of the Kähler structure of the transverse manifolds, i.e. hermitian isometries which preserve the complex structure. The Molino structure theorem for Riemannian foliations, cf. [10], has its hermitian version, cf. [16, 19].

THEOREM 2. *Let $\mathcal{F}$ be a transversely hermitian codimension $2q$ foliation on a compact manifold M. Then the bundle of transverse orthonormal frames $B(M, O(2q); \mathcal{F})$ admits an $U(q)$ reduction $B(M, U(q); \mathcal{F})$ which is its foliated subbundle. The lifted foliation $\mathcal{F}_U$ is transversely parallelisable. The closures of its leaves are fibres of locally trivial fibration (called the basic foliation) onto a compact manifold. The foliation of any closure by leaves of $\mathcal{F}_U$ is a Lie foliation. The projections on M of the fibres of the basic fibration are the closure of the foliation $\mathcal{F}$.*

The structure theorem permits us to define the commuting sheaf, cf. [20, 19]. Local foliated transformations which preserve the complex structure and the hermitian metric of the normal bundle, or equivalently which define local isomorphisms of the induced Kähler (hermitian) structure on any traverse manifold are lifted to the bundle $B(M, U(q); \mathcal{F})$ and preserve the transverse parallelism. So local foliated infinitesimal automorphisms of the complex structure and the hermitian metric of the normal bundle are the foliated vector fields which when lifted to bundle $B(M, U(q); \mathcal{F})$ commute with the transverse parallelism. Therefore we can formulate the following proposition, cf. [19, 20].

PROPOSITION 2. *Let $\mathcal{F}$ be a transversely hermitian codimension $2q$ foliation on a compact manifold M. Then the closures of leaves are submanifolds,* *are "orbits" of the commuting sheaf of the foliation and form a singular foliation.*

Now we apply the above results to our $\mathcal{K}$-structures. The vector fields ξ_i are defined by an isometric action of R^s, which therefore defined a representation of $\mathbb{R}^s$ into the group Isom(M, g) of isometries of the Riemannian manifold (M, g). Therefore the closures of leaves are the orbits of some toral action – the action of T^r-the closure of $\mathbb{R}^s$ in Isom(M, g)-the smallest compact abelian subgroup of Isom(M, g) containing the image of $\mathbb{R}^s$. Let V be the orthogonal complement in $\mathbb{R}^r = Lie(T^r)$ of $\mathbb{R}^s = Lie(\mathbb{R}^s)$. As the flows of ξ_i, $i = 1, \ldots, s$, preserve the hermitian metric and the complex structure on the normal bundle, so the flows of the characteristic vector fields of the toral action. Thus the vector fields corresponding to vectors of V define the global trivialisation of the commuting sheaf

As the vector fields ξ_i commute it is most natural to recall the notion of the rank of a manifold, cf. [4, 15]. This fact will help us determine the dimension of the closures of leaves.

PROPOSITION 3. *Let $\mathcal{F}$ be a codimension q foliation determined by an isometric action of the group R^s. Then the closures of leaves have at most dimension $s + rk(q-1)$, where $rk(q)$ is the rank of the q-sphere.*

Let us choose a leaf L whose closure is of maximal dimension. The same property have neighbouring leaves. The fact that the foliation is Riemannian ensures that these neighbouring leaves live on sphere bundles of a tubular neighbourhood of L, So do their closures. Therefore on the $(q-1)$-spheres there are $(r-s)$ commuting vector fields. Thus $r-s$ must be smaller or equal to $rk(q-1)$. ∎

The closures of leaves define a singular Riemannian foliation $\mathcal{F}_b$. The set of points M_0 where the closures are of maximal dimension is open and dense in M, and on this set the closures form a regular Riemannian foliation. Let us look closer The tangent bundle TM on M_0 admits the orthogonal decomposition $T\mathcal{F}_\xi \oplus Q_b \oplus Q_N$ where $T\mathcal{F}_b = T\mathcal{F} \oplus Q_b$. Denote by π_b the orthogonal projection of TM onto Q_N. Then define the $(1,1)$-tensor field f_b by $\pi_b f$. It is a kind of f-structure on the open set M_0. We will study its properties in relation to the initial $\mathcal{K}$-structure in a subsequent paper.

4. Submanifolds in $\mathcal{K}$-manifolds

The theory of submanifolds tangent to the characteristic foliation developed for various types of f-structures can be also treated in a "foliated" way, so very often these results are straightforward generalisations of properties of submanifolds of Kähler manifolds. We assume that the ambient manifold M is compact, although in most cases this condition can be weakened to "complete".

Let W be an $m+s$ dimensional submanifold of M tangent to the characteristic foliation $\mathcal{F}_\xi$, i.e. for any $x \in W$, $T_x\mathcal{F}_\xi \subset T_x W$ or equivalently $\xi_i(x) \in T_x W$ for $i = 1, \ldots, s$. The following lemma is a simple generalization of the Frobenius theorem.

LEMMA 1. *Let x be a point of a submanifold W of dimension $m+s$ tangent to the foliation $\mathcal{F}_\xi$. Then there exists an adapted chart $\psi\colon V \longrightarrow R^{2n+s}, \psi = (\psi_1, \ldots, \psi_{2n+s})$, at x such that the set $U = \{y \in V \mid \psi_{m+s+1}(y) = = \ldots \psi_{2n+s}(y) = 0\}$ is a connected component of $V \cap W$ containing x and $(\psi_1 | U, \ldots \psi_{m+s}$ is an adapted chart for the induced foliation of W.*

As a corollary we obtain the following:

PROPOSITION 4. *Let W be a submanifold tangent to the characteristic foliation of a $\mathcal{K}$-manifold M. Then for any point x of W there exist neighbourhoods U and V of x in W and M, respectively, having the following properties:*

i) *U is a connected component of $V \cap W$ containing x;*
ii) *U is a foliated subset of V (for the characteristic foliation);*
iii) *there exists a Riemannian submersion with connected fibres $f\colon V \to N_0$ onto a Kähler manifold N_0 defining the characteristic foliation;*
iv) *there exists a submanifold $\bar{W}$ of N_0 such that $U = f^{-1}(\bar{W})$.*

Now we turn our attention to CR-submanifolds, cf. [21, 9, 12, 3, 7, 5, 6].

Let W be a connected submanifold of a $\mathcal{K}$-manifold M. Assume that W is tangent to the characteristic foliation. Then W is called a contact CR-submanifold of M if there exists a differentiable distribution D on W of constant dimension, $D\colon x \mapsto D_x \subset T_x W$, satisfying the following conditions:

i) D is invariant with respect to f, i.e. for any $x \in W$ $f(D_x) \subset D_x$;
ii) the complementary orthogonal distribution $D^\perp\colon x \mapsto D_x^\perp \subset T_x W$ is anti-invariant with respect to f, i.e. for any $x \in W$;$f(D_x^\perp) \subset T_x W^\perp$.

A contact CR-submanifold W is non-trivial if $\dim D = h > 0$ and $\dim D^{\perp} = q > O$; cf. [21] p. 48. Let $f(D_x) = T_x W \cap f(T_x W) = D_0$ and $f(D_x^{\perp}) = T_x W^{\perp} \cap f(T_x W)$. Then the distribution D_0 has constant dimension and $D = D_0$ or $D = D_0 \oplus T\mathcal{F}$, and $D^{\perp} = D_0^{\perp} \oplus T\mathcal{F}$ or $D_0^{\perp}$, respectively, where $D_0^{\perp}$ is the orthogonal complement of $D_0 \oplus T\mathcal{F}$. This means that the tangent bundle TW of W admits the following decomposition: $T\mathcal{F}_\xi \oplus D_0 \oplus D_0^{\perp}$ and $\ker\eta|_{TW} = imf|_{TW} = D_0 \oplus D_0^{\perp}$. For the rest of the paper we assume that $D = D_0 \oplus T\mathcal{F}$.

Our previous considerations lead to the following:

PROPOSITION 5. *Let W be a submanifold tangent to the characteristic foliation of a $\mathcal{K}$-manifold. Then W is a contact CR-submanifold iff the corresponding submanifolds in any transverse manifold are the characteristic foliation CR-submanifolds.*

Our distributions D and $D^{\perp}$ have the following properties.

THEOREM 3. *Let W be a contact CR-submanifold of a $\mathcal{K}$-manifold M. Then the distribution $D^{\perp} \oplus T\mathcal{F}$ is completely integrable and its integral submanifolds are anti-invariant submanifolds (tangent to the characteristic foliation).*

For the proof see Theorem III.3.1 of [21] or [9], Theorem 3.1. Similarly we have the following version of Theorem III.3.2 of [21], or [9], Th.3.5 , where B is the second fundamental form of the submanifold W in M:

THEOREM 4. *Let W be a contact CR-submanifold of a K-manifold M. Then the distribution D is integrable iff $B(X,fY) = B(Y,fX)$ for any $X, Y \in D$. Its integral submanifolds are invariant submanifolds of M.*

REMARK 1. As the properties described by the above theorems are local, they can be derived from the corresponding theorems for CR-submanifolds of Kähler manifolds, compare Theorems IV.4.1 and IV.4.2 of [21].

Having proved these basic properties let us turn our attention to geodesics:

PROPOSITION 6. *Let W be a contact CR-submanifold tangent to the characteristic foliation $\mathcal{F}_\xi$ of a $\mathcal{K}$-manifold M. If $g(B(X, Y),fZ) = 0$ for any $X, Y \in D_0$, $Z \in D_0^{\perp}$ then any geodesic of W tangent to D_0 at one point remains tangent to D_0 at any point of its domain.*

PROOF. The foliation $\mathcal{F}_\xi | W$ is a Riemannian foliation and the distribution $D_0 \oplus D_0^\perp$ is the orthogonal complement of the bundle tangent to the foliation. Therefore a geodesic orthogonal to $\mathcal{F}_\xi$, i.e. tangent to $D_0 \oplus D_0^\perp$ at one point is tangent to $D_0 \oplus D_0^\perp$ at any point of its domain. Moreover, such orthogonal geodesics are $D_0 \oplus D_0^\perp$-horizontal lift of the corresponding geodesics in the transverse manifold, cf. [10]. Let us consider a geodesic $\alpha\colon (a,b) \to W$ tangent to D_0 at 0 and the set $A = \{t \in (a,b)\colon \dot{\alpha}(t) \in D_0\}$. The set A is closed and $0 \in A$. We shall show that it is also open. As the problem is local we can reduce our considerations to a foliated submanifold of a $\mathcal{K}$-anifold with the characteristic foliation given by a global submersion with connected fibres, i.e. the characteristic fibration $f\colon M \to N$ and $W = h^{-1}(\bar{W})$ where $\bar{W}$ is a CR-submanifold of the Kähler manifold N. Therefore $T\bar{W}$ admits a decomposition into orthogonal distributions $\bar{D}$ and $\bar{D}^\perp$ such that $D = h^{-1}(\bar{D})$ and $D_0 = \ker\eta \cap h^{-1}(\bar{D}), D_0^\perp = \ker\eta \cap h^{-1}(\bar{D}^\perp$. Let B be the second fundamental form of the submanifold W in M and $\bar{B}$ be the second fundamental form of the submanifold $\bar{W}$ in N. Then $B(X^*, Y^*) = \bar{B}(X, Y)^*$, cf. [21], p. 101, where for any vector X tangent to $\bar{W}$ X^* is its $\ker\eta$ ($D_0 \oplus D_0^\perp$)-lift to M, and hence $\bar{g}(\bar{B}(X, Y), \bar{f}Z) = 0$ for any $X, Y \in \bar{D}$ and $Z \in \bar{D}^\perp$. Then Proposition IV.4.2 of [21] ensures that $\bar{D}$ is a totally geodesic foliation of $\bar{W}$. Let $\bar{\alpha}$ be the geodesic in $\bar{W}$ corre then $\bar{\alpha}$ is tangent to $\bar{D}$ at this point. Since the foliation $\bar{D}$ is totally geodesic $\bar{\alpha}$ must be contained in some leaf of $\bar{D}$. Hence α being the $D_0 \oplus D_0^\perp$-horizontal lift of $\bar{\alpha}$, it must be tangent to D_0. Therefore the set A is open, and thus $A = (a,b)$. ∎

Taking as a model Kähler manifolds we can introduce the following notions:

DEFINITION 1. We say that a contact CR-submanifold W is:

i) D_0-totally geodesic iff $B(X, Y) = 0$ for any $X, Y \in D_0$;

ii) contact mixed foliate if $B(X, Y) = 0$ for any $X \in D$ and $Y \in D^\perp$, and $B(PX, Y) = B(X, PY)$ for any $X, Y \in D_0$.

It is not difficult to verify the following:

LEMMA 2.

i) *W is a D_0-totally geodesic iff $\bar{W}$ is $\bar{D}$-totally geodesic;*

ii) *W is contact mixed foliate iff $\bar{W}$ is mixed foliate.*

Then we can prove:

PROPOSITION 7. *Let W be a contact CR-submanifold tangent to the characteristic foliation of a $\mathcal{K}$-manifold M. If W is D_0-totally geodesic, then D is a foliation and any geodesic of W tangent to D_0 at one point remains tangent to D_0 at any point of its domain.*

PROOF. It is a consequence of Lemma 2, Corollary IV.4.3 of [21] and of the considerations similar to those of the second part of the proof of Proposition 4. ∎

Another property of Kähler manifolds gives us the following theorem, cf. Theorem IV.6.1 of [21] or [1].

THEOREM 5. *Let W be a contact totally umbilical non-trivial contact CR-submanifold of a K-manifold M. If $\dim D_0^{\perp} > 1$, then a geodesic orthogonal to $\mathcal{F}_\xi$ and tangent to W at one point has this property on an open subset of its domain.*

PROOF. The corresponding submanifold $\bar{W}$ in the transverse manifold is totally umbilical. Since the characteristic foliation is Riemannian we have to show that the geodesic is tangent to W on an open subset of its domain. This property is a local one and therefore we can reduce our considerations to the canonical fibration. The geodesic is the $\ker\eta$-horizontal lift of a geodesic in N. Therefore it is sufficient to know that the submanifold $\bar{W}$ is totally geodesic. This is precisely the fact which Bejancu's theorem ensures. ∎

Finally we have the following theorem about totally geodesic CR-submanifolds, cf. Theorem 3.4 of [7] for $\mathcal{S}$-structures.

THEOREM 6. *Let W be a totally geodesic contact CR-submanifold of a $\mathcal{K}$-manifold M. Then D and $D^{\perp} \oplus T\mathcal{F}$ are Riemannian foliations, and locally:*

i) *W is diffeomorphic to $\mathbb{R}^s \times N_0 \times N_1$,*

ii) *the foliation D is given by the projection $\mathbb{R}^s \times N_0 \times N_1 \to N_1 \subset N$,*

iii) *the foliation $D^{\perp} \oplus T\mathcal{F}$ is given by the projection $\mathbb{R}^s \times N_0 \times N_1 \to N_0 \subset N$,*

iv) *the submanifold $\bar{W} \subset N$ is a Riemannian product of $N_0 \times N_1$ of a totally geodesic invariant submanifold N_0 and a totally geodesic anti-invariant submanifold N_1 of N.*

PROOF. The problem is local and we can reduce our considerations to the case of canonical fibration. Therefore we can assume that $W = f^{-1}(\bar{W})$ for some CR-submanifold $\bar{W}$ of the Kähler manifold N and that the submersion $f : M \to N$ is a Riemannian submersion. The orthogonal complement of $T\mathcal{F}$

on W is equal to $\ker\eta = D_0 \oplus D_0^\perp$. Therefore $D_0 = (df\,|\,W)^{-1}(\bar{D}) \cap \ker\eta$ and $D_0^\perp = (df\,|\,W)^{-1}(\bar{D}^\perp) \cap \ker\eta$ where $\bar{D}$ and $\bar{D}^\perp$ are invariant and anti-invariant distributions, respectively, of the CR-submanifold $\bar{W}$ of N.

Since $\bar{W}$ is totally geodesic, cf. [21], Prop. V.2.5, Theorem IV.6.2 of [21] assures that the submanifold $\bar{W}$ of N is a Riemannian product of $N_0 \times N_1$ of a totally geodesic invariant submanifold N_0 and a totally geodesic anti-invariant submanifold N_1 of N. Therefore it remains to prove that the foliations D and $D^\perp \oplus T\mathcal{F}$ are Riemannian foliations of the submanifold W. The subbundle $D_0^\perp$ is the orthogonal complement of D, therefore the foliation D is Riemannian iff any a geodesic of W which is tangent to $D_0^\perp$ at one point remains tangent to $D_0^\perp$ at any point of its domain, cf. [22, 10]. Likewise the subbundle D_0 is the orthogonal complement of $D^\perp \oplus T\mathcal{F}$, therefore the foliation $D^\perp \oplus T\mathcal{F}$ is Riemannian iff any a geodesic of W which is tangent to D_0 at one point remains tangent to D_0 at any point of its domain.

Let us take a geodesic γ of W which is tangent to $D_0^\perp$ at one point x . Since f is a Riemannian submersion γ is a horizontal geodesic, i.e. tangent to $\ker\eta$. Its image $f\gamma$ is a geodesic in $\bar{W}$, cf. [8], which is tangent to $\bar{D}^\perp$ at one point. As both distributions, $\bar{D}$ and $\bar{D}^\perp$ are totally geodesic, the geodesic $f\gamma$ remains tangent to $\bar{D}$ throughtout its domain. The $\ker\eta$-orthogonal lift γ' passing through the point x of $f\gamma$ is a geodesic in M and W which is tangent to $D_0^\perp$. Both geodesics, γ and γ', have the same tangent vector at the point x, therefore they must be equal.

Similar considerations are valid for the other distribution. ∎

References

[1] A. Bejancu, Umbilical CR-submanifolds of a Kähler manifold, *Rend. Mat.,* **15** (1980), 29–45.

[2] D. E. Blair, Geometry of manifolds with structural group $\mathcal{U}(n) \times \mathcal{O}(s)$, *J. Diff. Geom.,* **4** (1970), 155–167.

[3] J. L. Cabrerizo, Some results on normal CR-submanifolds of S-manifolds, *Bull. Math. Soc. Sci. Roum.,* **36(84)** (1992), 103–107.

[4] C. Camacho, A. Neto, *Geometric Theory of Foliations,* Birkhäuser, Boston 1985.

[5] L. Di Terlizzi, On invariant submanifolds of $\mathcal{C}$-manifolds and $\mathcal{S}$-manifolds, *Acta Math. Hungar.,* **85** (1999), 229–239.

[6] K. DUGGAL, S. IANUS, A. M. PASTORE, Maps interchanging f-structures and their harmonicity, preprint 2000

[7] L. M. FERMANDEZ, CR-products of S-products, *Port. Math.,* **47** (1990), 167–181.

[8] R. HERMANN, A sufficient condition that a map of Riemannian manifolds be a fiber bundle, *Proc. AMS,* **11** (1960), 236–242.

[9] I. MIHAI, CR-subvarietati ale unei f-varietati cu repere complementare, *Stud. Cerc. Mat.,* **35** (1983), 127–136.

[10] P. MOLINO, *Riemannian Foliations, Progress in Math.,* vol. 73, Birkhäuser 1988.

[11] C. C. MOORE, CL. SCHOCHET, *Global Analysis on Foliated Spaces,* MSRI LN no 9, Springer 1988.

[12] L. ORNEA, Subvarietati Cauchy–Riemann generice in S-varietati, *Stud. Cerc. Mat.,* **36** (1984), 435–443.

[13] B. REINHART, Foliated manifolds with bundle-like metrics, *Ann. of Math.,* **69** (1959), 119–132.

[14] B. REINHART, *Differential Geometry of Foliations,* Springer Verlag 1983.

[15] V. ROVENSKII, *Foliations on Riemannian Manifolds and Submanifolds,* Birkhäuser 1998.

[16] R. A. WOLAK, On $\nabla - G$-foliations, *Suppl. Rend. Cir. Mat. Palermo,* **6** (1984), 329–341.

[17] R. A. WOLAK, Foliations admitting transverse systems of differential equations, *Comp. Math.,* **67** (1988), 89–101.

[18] R. A. WOLAK, Foliated and associated geometric structures on foliated manifolds, *Ann. Fac. Sc. Toulouse,* **10** (1989), 337–360.

[19] R. A. WOLAK, *Geometric Structures on Foliated Manifolds,* Santiago de Compostela 1989.

[20] R. A. WOLAK, Foliated G-structures and Riemannian foliations, *Manus. Math.,* **66** (1989), 45–59.

[21] K. YANO, M. KON, *CR-submanifolds of Kaehlerian and Sasakian Manifolds, Progress in Math.,* vol 30, Birkhäuser 1983.

[22] S. YOROZU, Behaviour of geodesics in foliated manifolds with bundle-like metrics, *J. Math. Soc. Japan,* **35** (1983), 251–272.

ANNALES UNIV. SCI. BUDAPEST., **44** (2001), 183–198

CONVERGENCE OF MOVING AVERAGE PROCESSES DEDUCED BY NEGATIVELY ASSOCIATED RANDOM VARIABLES

By

H.-Y. LIANG and JONG-IL BAEK

Department of Applied Mathematics, Tongji University, Shanghai P. R. China
School of Mathematics and Informational Statistics, Wonkwang University, South Korea

(Received October 29, 2001)
(Revised April 15, 2002)

1. Introduction

Let $\{Y_i, -\infty < i < \infty\}$ be a sequence of identically distributed random variables and $\{a_i, -\infty < i < \infty\}$ a sequence of real numbers with $\sum_{i=-\infty}^{\infty} |a_i| < < \infty$. Put

$$X_k = \sum_{i=-\infty}^{\infty} a_{i+k} Y_i, \qquad k \geq 1.$$

When $\{Y_i, -\infty < i < \infty\}$ is a sequence of independent random variables, there have been some authors who studied limit properties for the moving average process $\{X_k, k \geq 1\}$. In particular, IBRAGIMOV (1962) had established the Central Limit Theorem for $\{X_k, k \geq 1\}$, BURTON and DEHLING (1990) had obtained large deviation priciple for $\{X_k, k \geq 1\}$ assuming $E \exp(t Y_1) < < \infty$ for all t, and LI et al. (1992) had obtained the following result on complete convergence.

THEOREM A. *Suppose $\{Y_i, -\infty < i < \infty\}$ is a sequence of independent and identically distributed (i.i.d.) random variables. Let $\{X_k, k \geq 1\}$ be defined as above and $1 \leq t < 2$. Then $E Y_1 = 0$ and $E|Y_1|^{2t} < \infty$ imply*

$$\sum_{n=1}^{\infty} P\left(\left|\sum_{k=1}^{n} X_k\right| \geq \epsilon n^{1/t}\right) < \infty, \qquad \forall \epsilon > 0. \tag{1.1}$$

Recently, ZHANG (1996) gave a general version of Theorem A under identically distributed ϕ-mixing assumptions. Clearly, Theorem A implies

$$\sum_{n=1}^{\infty} P\left(\left|\sum_{i=1}^{n} Y_i\right| \geq \epsilon n^{1/t}\right) < \infty, \qquad \forall \epsilon > 0. \tag{1.2}$$

While, by Kolmogorov's law of iterated logarithm, we know

$$\limsup_{n\to\infty} \frac{\left|\sum\limits_{i=1}^{n} Y_i\right|}{n^{1/2}} = \infty \quad \text{a.s..}$$

Therefore, (1.2) is not true for $t = 2$, further (1.1) does not hold for $t = 2$.

The main aim of this note is to extend and generalize Theorem A to NA random variables; discuss the result for $t = 2$ in NA setting, which had not been settled by LI et al. (1992) in i.i.d. setting.

A finite family of random variables $\{X_i, 1 \leq i \leq n\}$ is said to be negatively associated (NA) if for every pair of disjoint subsets A and B of $\{1, 2, \ldots, n\}$,

$$\mathrm{Cov}(f_1(X_i, i \in A),\ f_2(X_j, j \in B)) \leq 0$$

whenever f_1 and f_2 are coordinatewise increasing and such that the covariance exists. An infinite family of random variables is NA if every finite subfamily is NA. This definition is introduced by ALAM and SAXENA (1981) and carefully studied by JOAG-DEV and PROSCHAN (1983) and BLOCK, SAVITS and SHAKED (1982). As pointed out and proved by JOAG-DEV and PROSCHAN (1983), a number of well known multivariate distributions possess the NA property, such as (a) multinomial, (b) convolution of unlike multinomials, (c) multivariate hypergeometric, (d) Dirichlet, (e) Dirichlet compound multinomial, (f) negatively correlated normal distribution, (g) permutation distribution, (h) random sampling without replacement, and (i) joint distribution of ranks. Because of its wide applications in multivariate statistical analysis and reliability, the notion of NA have received considerable attention recently. We refer to JOAG-DEV and PROSCHAN (1983) for fundamental properties, NEWMAN (1984) for the central limit theorem, MATULA (1992) for the three series theorem, SU et al. (1997) for a moment inequality, a weak invariance principle and an example to show that there exists infinite family of non-degenerate non-independent strictly stationary NA random variables, SHAO and SU (1999) for the law of the iterated logarithm, LIANG and SU (1999a) for convergence rates of law of the logarithm, LIANG and SU (1999b) and LIANG

(2000) for complete convergence of weighted sums, ROUSSAS (1994) for the central limit theorem of random fields, some examples and applications.

2. Main Result

Here, let $\{Y_i, -\infty < i < \infty\}$ be a sequence of identically distributed NA random variables (r.v.'s) with $EY_1 = 0$ and $\{X_k, k \geq 1\}$ be defined as in section 1. Denote by $L(x) = \max(1, \log x)$.

THEOREM 2.1. *Let $h(x) > 0$ be a slowly varying function as $x \to \infty$ and $r \geq 1$, $1 \leq t < 2$, $h(x)$ is non-decreasing when $r = 1$. If $E\left[|Y_1|^{rt} h(|Y_1|^t)\right] < \infty$, then $\forall \epsilon > 0$,*

$$\sum_{n=1}^{\infty} n^{r-2} h(n) P\left(\left|\sum_{k=1}^{n} X_k\right| \geq \epsilon n^{1/t}\right) < \infty.$$

THEOREM 2.2. *Let $r > 1$. If $E[Y_1^2 / L(|Y_1|)]^r < \infty$, then there exsits some $\epsilon_0 > 0$ such that $\epsilon > \epsilon_0$,*

$$\sum_{n=1}^{\infty} n^{r-2} P\left(\left|\sum_{k=1}^{n} X_k\right| \geq \epsilon (nL(n))^{1/2}\right) < \infty.$$

THEOREM 2.3. *For $\eta > 0$, if $E\left[Y_1^2 / (L(|Y_1|))^{1-\eta}\right] < \infty$, then $\forall \epsilon > 0$,*

$$\sum_{n=1}^{\infty} \frac{1}{n} P\left(\left|\sum_{k=1}^{n} X_k\right| \geq \epsilon (nL(n))^{1/2}\right) < \infty.$$

REMARK 2.1. Since i.i.d. r.v.'s are a special case of NA r.v.'s, Theorem 2.1 generalizes and extends Theorem A. Theorems 2.2–2.3 complement the results for $t = 2$ in NA setting, which had not been discussed by LI et al. (1992) in i.i.d. setting.

REMARK 2.2. GUT (1980) conjectured that under $\{Y_i\}$ is a sequence of i.i.d. symmetric random variables, for $\eta > 0$, if $E\left[Y_1^2 / (L(|Y_1|))^{1\eta}\right] < \infty$, then $\forall \epsilon > 0$,

$$\sum_{n=1}^{\infty} \frac{1}{n} P\left(\left|\sum_{i=1}^{n} Y_i\right| \geq \epsilon (n(L(n))^{1/2}\right) < \infty.$$

Clearly, Theorem 2.3 extends and generalizes (taking $a_0 = 1$, $a_j = 0$, $j \neq 0$) GUT's (1980) above conjecture.

3. Proof of Main Result

In this section, $a \ll b$ means $a = O(b)$. C and C_q $(q \geq 1)$ will represent positive constants, their value may change from one place to another.

LEMMA 1 (BURTON and DEHLING (1990)). *Let* $\sum_{i=-\infty}^{\infty} a_i$ *be an absolutely convergent series of real numbers* $a = \sum_{i=-\infty}^{\infty} a_i$, $b = \sum_{i=-\infty}^{\infty} |a_i|$. *Suppose* $\Phi : [-b, b] \to \mathbf{R}$ *is a function satisfying the following conditions:*

(i) Φ *is bounded and continuous at* a.

(ii) *There exist* $\delta > 0$ *and* $C > 0$ *such that for all* $|x| \leq \delta$, $|\Phi(x)| \leq C|x|$.

Then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=-\infty}^{\infty} \Phi\left(\sum_{j=i+1}^{i+n} a_j\right) = \Phi(a).$$

REMARK 3.1. Taking $\Phi(x) = |x|^q$, $q \geq 1$, from Lemma 1 we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=-\infty}^{\infty} \left|\sum_{j=i+1}^{i+n} a_j\right|^q = |a|^q. \tag{3.1}$$

LEMMA 2. (SU et al. (1997), SHAO and SU (1999)). *Let* $p \geq 2$ *and let* $\{X_i, i \geq 1\}$ *be a sequence of* NA r.v.'s *with* $EX_i = 0$ *and* $E|X_i|^p <$. *Then, there exist constant* $A_p > 0$ *and* $B_p > 0$ *such that*

$$E\left|\sum_{i=1}^{n} X_i\right|^p \leq A_p \left\{\left(\sum_{i=1}^{n} EX_i^2\right)^{p/2} + \sum_{i=1}^{n} E|X_i|^p\right\},$$

$$E \max_{1\leq k\leq n} \left|\sum_{i=1}^{k} X_i\right|^p \leq B_p \left\{\left(\sum_{i=1}^{n} EX_i^2\right)^{p/2} + \sum_{i=1}^{n} E|X_i|^p\right\}.$$

LEMMA 3. (SHAD and SU (1999)). *Let* $\{X_j, 1 \le j \le n\}$ *be mean zero* NA r.v.'s, *finite variance. Denote by* $B_n = \sum_{j=1}^{n} EX_j^2$. *Then for any* $x > 0$, $\alpha > 0$ *and* $0 < \beta < 1$,

$$P\left(\max_{1\le k\le n} |S_k| \ge x\right) \le 2P\left(\max_{1\le k\le n} |X_k| > \alpha\right) + $$
$$+\frac{2}{1-\beta}\exp\left\{-\frac{x^2\beta}{2(\alpha x + B_n)}\left(1+\frac{2}{3}\ln\left(1+\frac{\alpha x}{B_n}\right)\right)\right\}.$$

REMARK 3.2. If $\{Z_i; -\infty < i < \infty\}$ is a sequence of identically distributed mean zero NA r.v.'s with $E|Z_1| < \infty$, finite variance and $\{a_i; -\infty < i < \infty\}$ a sequence of real numbers with $\sum_{i=-\infty}^{\infty} |a_i| < \infty$. Put $B = \sum_{i=-\infty}^{\infty} E|a_i Z_i|^2$. From Lemma 3, we have

$$P\left(\left|\sum_{i=-\infty}^{\infty} a_i Z_i\right| \ge 2x\right) \le P\left(\left|\sum_{i=-m}^{m} a_i Z_i\right| \ge x\right) + P\left(\left|\sum_{|i|>m} a_i Z_i\right| \ge x\right) \le$$
$$\le 2P\left(\sup_i |a_i Z_i| > \alpha\right) + \frac{2}{1-\beta}\exp\left\{-\frac{x^2\beta}{2(\alpha x + B)}\right\} + \frac{E|Z_1|}{x}\sum_{|i|>m}|a_i|.$$

Since $\sum_{i=-\infty}^{\infty} |a_i| < \infty$, $\forall \epsilon > 0$, choose m such that $\frac{E|Z_1|}{x}\sum_{|i|>m}|a_i| < \epsilon$. Thus, we get

$$P\left(\left|\sum_{i=-\infty}^{\infty} a_i Z_i\right| \ge 2x\right) \le \tag{3.2}$$
$$\le 2P\left(\sup_i |a_i Z_i| > \alpha\right) + \frac{2}{1-\beta}\exp\left\{-\frac{x^2\beta}{2(\alpha x + B)}\right\}.$$

PROOF OF THEOREM 2.1. Note that

$$\sum_{k=1}^{n} X_k = \sum_{i=-\infty}^{\infty}\left(\sum_{k=1}^{n} a_{i+k}\right) Y_i = \sum_{i=-\infty}^{\infty} a_{ni} Y_i. \tag{3.3}$$

It suffices to show that for every $\epsilon > 0$,

$$\sum_{n=1}^{\infty} n^{r-2} h(n) P\left(\left|\sum_{i=-\infty}^{\infty} a_{ni}^{+} Y_i\right| > \epsilon n^{1/t}\right) < \infty, \tag{3.4}$$

$$\sum_{n=1}^{\infty} n^{r-2} h(n) P\left(\left|\sum_{i=-\infty}^{\infty} a_{ni}^{-} Y_i\right| > \epsilon n^{1/t}\right) < \infty, \tag{3.5}$$

where $a_{ni}^{+} = a_{ni} \vee 0$, $a_{ni}^{-} = (-a_{ni}) \vee 0$. We prove only (3.4), the proof of (3.5) is analogous. Let

$Y_{ni} =$

$= -n^{1/t} I(a_{ni}^{+} Y_i < -n^{1/t}) + a_{ni}^{+} Y_i I(|a_{NI}^{+} Y_i| \leq n^{1/t}) + n^{1/t} I(a_{ni}^{+} Y_i > n^{1/t})$.

Then

$$\sum_{n=1}^{\infty} n^{r-2} h(n) P\left(\left|\sum_{i=-\infty}^{\infty} a_{ni}^{+} Y_i\right| > \epsilon n^{1/t}\right) \leq$$

$$\leq \sum_{n=1}^{\infty} n^{r-2} h(n) P\left(\left|\sum_{i=-\infty}^{\infty} Y_{ni}\right| \geq \frac{\epsilon}{2} n^{1/t}\right) +$$

$$+ \sum_{n=1}^{\infty} h(n) \sum_{i=-\infty}^{\infty} P(|a_{ni}^{+} Y_i| > n^{1/t} =: I_1 + I_2.$$

From (3.1) we can assume, without loss of generality, that $\sum\limits_{i=-\infty}^{\infty} a_{ni}^{+} \leq n$, $a_{ni}^{+} \leq 1$ and denote by $I_{nj} = \{i \in Z : (j+1)^{-1/t} < a_{ni}^{+} \leq j^{-1/t}\}$. It is easy to verify from Lemma 1 that

$$\sum_{j=1}^{k} \# I_{nj} \leq C n (k+1)^{1/t}. \tag{3.6}$$

For I_2, using (3.6) we have

$$I_2 \leq \sum_{n=1}^{\infty} n^{r-2} h(n) \sum_{j=1}^{\infty} (\# I_{nj}) \sum_{k=nj}^{\infty} P(k \leq |Y_1|^t < k+1) \leq$$

$$\leq \sum_{n=1}^{\infty} n^{r-2} h(n) \sum_{k=n}^{\infty} \sum_{j=1}^{[k/n]} (\# I_{nj}) P(k \leq |Y_1|^t < k+1) \ll$$

$$\ll \sum_{n=1}^{\infty} n^{r-1} h(n) n^{-1/t} \sum_{k=n}^{\infty} k^{1/t} P(k \leq |Y_1|^t < k+1) \ll$$
$$\ll E[|Y_1|^{rt} h(|Y_1|^t)] < \infty.$$

By $EY_1 = 0$, we get

$$\frac{\left|\sum_{i=-\infty}^{\infty} EY_{ni}\right|}{n^{1/t}} \leq 2E|Y_1|^t I(|Y_1| > n^{1/t}) \to 0, \quad \text{as } n \to \infty.$$

Thus, to prove $I_1 < \infty$, we need only to show that

$$I_1^* =: \sum_{n=1}^{\infty} n^{r-2} h(n) P\left(\left|\sum_{i=-\infty}^{\infty} (Y_{ni} - EY_{ni})\right| \geq \epsilon n^{1/t}\right) < \infty, \quad \forall \epsilon > 0.$$

In fact, we use the Markov's inequality for a suitably large M, which will be determined later, Lemma 2 and note that for each $n \geq 1$, $\{Y_{ni}, -\infty < i < \infty\}$ is still a sequence of NA r.v.'s from the definition, we have

$$I_1^* \ll \sum_{n=1}^{\infty} n^{r-2} h(n) n^{-M/t} \left\{ \left(\sum_{i=-\infty}^{\infty} E|Y_{ni}|^2 \right)^{M/2} + \sum_{i=-\infty}^{\infty} E|Y_{ni}|^M \right\} =$$
$$=: I_3 + I_4.$$

If $r > 2$, note that $E|Y_1|^2 < \infty$ and $\sum_{i=-\infty}^{\infty} |a_{ni}|^q \leq Cn$ for $q \geq 1$, taking $M > 2t(r-1)/(2-t)$, we get

$$I_3 \leq \sum_{n=1}^{\infty} n^{r-2} h(n) n^{-M/t} \left\{ \sum_{i=-\infty}^{\infty} \left[n^{2/t} P(|a_{ni}^+ Y_i| > n^{1/t}) + \right.\right.$$
$$\left.\left. + E|a_{ni}^+ Y_i|^2 I(|a_{ni}^+ Y_i| \geq n^{1/t}\right] \right\}^{M/2} \ll$$
$$\ll \sum_{n=1}^{\infty} n^{r-2-(1/t-1/2)M} h(n) < \infty.$$

If $1 < r \leq 2$ and $1 < rt \leq 2$, then there exists some s such that $r > s > 1$, taking $M > 2(r-1)/(s-1)$, we have

$$I_3 \leq \sum_{n=1}^{\infty} n^{r-2} h(n) n^{-M/t} \left\{ \sum_{i=-\infty}^{\infty} \left[n^{2/t} P(|a_{ni}^{+} Y_i| > n^{1/t}) + \right.\right.$$

$$\left.\left. + E|a_{ni}^{+} Y_i|^2 I(|a_{ni}^{+} Y_i| \leq n^{1/t} \right] \right\}^{M/2} =$$

$$= \sum_{n=1}^{\infty} n^{r-2-M/t} h(n) \left[\sum_{i=-\infty}^{\infty} \int_0^{2/t} P(|a_{ni}^{+} Y_i|^2 > x) dx \right]^{M/2} \ll$$

$$\ll \sum_{n=1}^{\infty} n^{r-2-(s-1)M/2} h(n) < \infty.$$

If $r = 1$. Choose $M = 2$. Similarly to the below proof of $I_4 < \infty$, we get $I_3 < \infty$.

As to I_4, we have

$$I_4 \ll \sum_{n=1}^{\infty} n^{r-2} h(n) \sum_{i=-\infty}^{\infty} P(|a_{ni}^{+} Y_i| > n^{1/t}) +$$

$$+ \sum_{n=1}^{\infty} n^{r-2-M/t} h(n) \sum_{i=-\infty}^{\infty} E|a_{ni}^{+} Y_i|^M I(|a_{ni}^{+} Y_i| \leq n^{1/t}) =: I_5 + I_6.$$

From the proof of $I_2 < \infty$ we know $I_5 < \infty$.

$$I_6 \leq \sum_{n=1}^{\infty} n^{r-2-M/t} h(n) \sum_{j=1}^{\infty} (\# I_{nj}) j^{-M/t} \sum_{k=1}^{2n} E|Y_1|^M I(k_1 < |Y_1|^t \leq k) +$$

$$+ \sum_{n=1}^{\infty} n^{r-2-M/t} h(n) \sum_{j=1}^{\infty} (\# I_{nj}) j^{-M/t} \sum_{k=2n+1}^{n(j+1)} E|Y_1|^M I(k_1 < |Y_1|^t \leq k) =$$

$$=: I_7 + I_8.$$

Note that for $M \geq 1$ and $k \geq 1$, we have

$$\sum_{j=k}^{\infty} (\# I_{nj}) j^{-M/t} \leq C n k^{-(M-1)/t}.$$

Hence, taking $M > rt$, we get

$$I_7 \ll \sum_{k=1}^{\infty} \sum_{n=[k/2]}^{\infty} n^{r-1-\frac{M}{t}} h(n) E|Y_1|^M I(k_1 < |Y_1|^t \le k) \ll$$

$$\ll \sum_{k=1}^{\infty} k^{r-\frac{M}{t}} h(k) E|Y_1|^M I(k_1 < |Y_1|^t \le k) \ll$$

$$\ll E[|Y_1|^{rt} h(|Y_1|^t)] < \infty.$$

$$I_8 \ll \sum_{n=1}^{\infty} n^{r-2-\frac{M}{t}} h(n) \sum_{k=2n+1}^{\infty} n \left(\frac{k}{n}\right)^{\frac{-(M-1)}{t}} E|Y_1|^M I(k_1 < |Y_1|^t \le k) \ll$$

$$\ll \sum_{k=2}^{\infty} \sum_{n=1}^{[k/2]} n^{r-1-\frac{M}{t}} h(n) E|Y_1|^M I(k_1 < |Y_1|^t \le k) \ll$$

$$\ll E[|Y_1|^{rt} h(|Y_1|^t)] < \infty.$$

PROOF OF THEOREM 2.2. We need only to prove that for $\epsilon > \epsilon_0/2$,

$$\sum_{n=1}^{\infty} n^{r-2} P\left(\left|\sum_{i=-\infty}^{\infty} a_{ni}^{+} Y_i\right| > \epsilon (nL(n))^{1/2}\right) < \infty. \tag{3.7}$$

$$\sum_{n=1}^{\infty} n^{r-2} P\left(\left|\sum_{i=-\infty}^{\infty} a_{ni}^{-} Y_i\right| > \epsilon (nL(n))^{1/2}\right) < \infty. \tag{3.8}$$

We give proof of (3.7), the proof of (3.8) is analogous. Let

$$\lambda_n = \frac{10}{2\epsilon}\sqrt{n/L(n)}, \quad \rho_n = \frac{\epsilon}{4N}\sqrt{nL(n)},$$

$$Y_{ni}^{(1)} = -\lambda_n I(a_{ni}^{+} Y_i < -\lambda_n) + a_{ni}^{+} Y_i I(|a_{ni}^{+} Y_i| \le \lambda_n) + \lambda_n I(a_{ni}^{+} Y_i > \lambda_n),$$

$$Y_{ni}^{(2)} = (a_{ni}^{+} Y_i - \lambda_n) I(\lambda_n < a_{ni}^{+} Y_i < \rho_n),$$

$$Y_{ni}^{(3)} = (a_{ni}^{+} Y_i + \lambda_n) I(-\lambda_n > a_{ni}^{+} Y_i > -\rho_n),$$

$$Y_{ni}^{(4)} = (a_{ni}^{+} Y_i + \lambda_n) I(a_{ni}^{+} Y_i \le -\rho_n) + (a_{ni}^{+} Y_i - \lambda_n) I(a_{ni}^{+} Y_i \ge \rho_n),$$

where N is some large positive integer, which will be specified later on. Then

$$\sum_{n=1}^{\infty} n^{r-2} P\left(\left|\sum_{i=-\infty}^{\infty} a_{ni}^{+} Y_i\right| > \epsilon (nL(n))^{1/2}\right) \le$$

$$\leq \sum_{n=1}^{\infty} n^{r-2} P\left(\left|\sum_{i=-\infty}^{\infty} Y_{ni}^{(1)}\right| \geq \frac{\epsilon}{4}(nL(n))^{1/2}\right) +$$

$$+\sum_{n=1}^{\infty} n^{r-2} P\left(\left|\sum_{i=-\infty}^{\infty} Y_{ni}^{(2)}\right| \geq \frac{\epsilon}{4}(nL(n))^{1/2}\right) +$$

$$+\sum_{n=1}^{\infty} n^{r-2} P\left(\left|\sum_{i=-\infty}^{\infty} Y_{ni}^{(3)}\right| \geq \frac{\epsilon}{4}(nL(n))^{1/2}\right) +$$

$$+\sum_{n=1}^{\infty} n^{r-2} P\left(\left|\sum_{i=-\infty}^{\infty} Y_{ni}^{(4)}\right| \geq \frac{\epsilon}{4}(nL(n))^{1/2}\right) =: J_1 + J_2 + J_3 + J_4.$$

From (3.1), we can assume $a_{ni}^{+} \leq (2L(2))^{-1/2}$, denote by

$$I_{nj} = \{i \in \mathbf{Z} : ((j+2)L(j+2))^{-1/2} < a_{ni}^{+} \leq ((j+1)L(j+1))^{-1/2}\}.$$

Note that $\sum_{j=1}^{k} \# I_{nj} \leq Cn((k+2)L(k+2))^{1/2}$. Similarly to the proof of $I_2 < \infty$, we get

$$J_4 \leq \sum_{n=1}^{\infty} n^{r-2} \sum_{i=-\infty}^{\infty} P\left(|a_{ni}^{+} Y_i| \geq \frac{\epsilon}{4N}(nL(n))^{1/2}\right) \leq$$

$$\leq \sum_{n=1}^{\infty} n^{r-2} \sum_{j=1}^{\infty} (\# I_{nj}) \sum_{k=nj}^{\infty} P(|Y_1|^2 / L(|Y_1|) \geq Cnj) \ll$$

$$\ll \sum_{k=1}^{\infty} k^{1/2} \sum_{n=1}^{k} n^{r-3/2} (L(3k/n))^{1/2} P\left(k \leq \frac{|Y_1|^2}{CL(|Y_1|)} < k+1\right).$$

Choose $\beta > 0$ such that $r - 1/2 > \beta$, $L(x) \leq Cx^{2\beta}$ when $x \geq 2k_0$ for some $k_0 > 0$. Hence,

$$\sum_{n=1}^{k} n^{r-3/2} (L(2k/n))^{1/2} \ll \int_{1}^{k} x^{r-3/2} (L(2k/x))^{1/2} dx \ll$$

$$\ll \int_1^{k/k_0} x^{r-3/2}(k/x)^\beta dx + \int_{k/k_0}^{k} x^{r-3/2}(L(2k_0))^{1/2} dx \ll k^{r-1/2}.$$

Therefore,

$$J_4 \ll \sum_{k=1}^{\infty} k^r P\left(k \le \frac{|Y_1|^2}{CL(|Y_1|)} < k+1\right) \ll E[|Y_1|^2/L(|Y_1|)]^r < \infty.$$

Choose r_0 such that $r > r_0 > 1$, hence $E|Y_1|^{2r_0} < \infty$. From the definition of $Y_{ni}^{(2)}$, we know that $Y_{ni}^{(2)} > 0$, taking $N > (r-1)/(r_0-1)$, by the property of NA, we have

$$J_2 = \sum_{n=1}^{\infty} n^{r-2} \left(\sum_{i=-\infty}^{\infty} Y_{ni}^{(2)} \ge \frac{\epsilon}{4}(nL(n))^{1/2}\right) \le$$

$$\le \sum_{n=1}^{\infty} n^{r-2} P(\text{there are at least N } i\text{'s such that } Y_{ni}^{(2)} \ne 0) \le$$

$$\le \sum_{n=1}^{\infty} n^{r-2} \left[\sum_{i=-\infty}^{\infty} P(a_{ni}^{+} Y_i > \lambda_n)\right]^N \ll$$

$$\ll \sum_{n=1}^{\infty} n^{r-2-(r_0-1)N} (L(n))^{Nr_0} < \infty.$$

Similarly, $Y_{ni}^{(3)} < 0$ and $J_3 < \infty$. By $EY_1 = 0$ and $E|Y_1|^{2r_0} < \infty$, $\sum_{i=-\infty}^{\infty} (a_{ni}^{+})^{2r_0} \le Cn$, we have

$$\left|\sum_{i=-\infty}^{\infty} EY_{ni}^{(1)}\right| /(nL(n))^{1/2} \le$$

$$\le \sum_{i=-\infty}^{\infty} [\lambda_n P(|a_{ni}^{+} Y_i| > \lambda_n) + E|a_{ni}^{+} I_i| I(|a_{ni}^{+} Y_i| > \lambda_n)]/(nL(n))^{1/2} \ll$$

$$\ll 1/n^{-(r_0-1)}(L(n))^{-(r_0-1)} \to 0, \qquad \text{az } n \to \infty.$$

Therefore, to prove $J_1 < \infty$, it suffices to show that

$$J_1^* =: \sum_{n=1}^{\infty} n^{r-2} P\left(\left|\sum_{i=-\infty}^{\infty} (Y_{ni}^{(1)} - EY_{ni}^{(1)})\right| \geq \frac{\epsilon}{5}(nL(n))^{1/2}\right) < \infty.$$

Note that for each $n \geq 1$, $\{Y_{ni}^{(1)}, -\infty < i < \infty\}$ is still a sequence of NA r.v.'s. Taking $\alpha = \frac{10}{\epsilon}\sqrt{nL(n)}$, $x = \frac{\epsilon}{10}\sqrt{nL(n)}$, $\beta = \frac{1}{2}$ and it is easy to verify that

$$\sup_i |Y_{ni}^{(1)} - EY_{ni}^{(1)}| \leq 2\lambda_n = \alpha, \qquad B = \sum_{i=-\infty}^{\infty} E(Y_{ni}^{(1)} - EY_{ni}^{(1)})^2 \leq C_0 nEY_1^2,$$

where C_0 satisfies $\sum_{i=-\infty}^{\infty} a_{ni}^2 \leq nC_0/2$. Hence, by using (3.2) we get

$$J_1^* \leq 4\sum_{n=1}^{\infty} n^{r-2} \exp\left(-\frac{\frac{1}{2}\cdot\frac{\epsilon^2}{100}\cdot nL(n)}{2(n + C_0 nEY_1^2)}\right) = 4\sum_{n=1}^{\infty} n^{r-2-\frac{\epsilon}{400(1+C_0EY_1^2)}} < \infty,$$

here $\epsilon_0 = 40\sqrt{(r-1)(1 + C_0 EY_1^2)}$.

PROOF OF THEOREM 2.3. Similarly to the proof of Theorem 2.2, we prove only that

$$\sum_{n=1}^{\infty} \frac{1}{n} P\left(\left|\sum_{i=-\infty}^{\infty} a_{ni}^+ Y_i\right| > \epsilon(nL(n))^{1/2}\right) < \infty, \qquad \forall \epsilon > 0.$$

We mas assume $\eta < 1$ and choose $\alpha > 0$ such that $\alpha < \eta$. Denote by $\lambda_n = n^{1/2}(L(n))^{(1-\alpha)/2}$,

$$Y_{ni}^{(1)} = -\lambda_n I(a_{ni}^+ Y_i < -\lambda_n) + a_{ni}^+ Y_i I(|a_{ni}^+ Y_i| \leq \lambda_n) + \lambda_n I(a_{ni}^+ Y_i > \lambda_n),$$

$$Y_{ni}^{(2)} = (a_{ni}^+ Y_i - \lambda_n) I\left(\lambda_n < a_{ni}^+ Y_i \leq \frac{\epsilon}{4N}(nL(n))^{1/2}\right),$$

$$Y_{ni}^{(3)} = (a_{ni}^+ Y_i + \lambda_n) I\left(-\lambda_n > a_{ni}^+ Y_i \geq \frac{\epsilon}{4N}(nL(n))^{1/2}\right),$$

$$Y_{ni}^{(4)} = (a_{ni}^+ Y_i + \lambda_n) I\left(a_{ni}^+ Y_i < -\frac{\epsilon}{4N}(nL(n))^{1/2}\right) + \\ + (a_{ni}^+ Y_i - \lambda_n) I\left(a_{ni}^+ Y_i > \frac{\epsilon}{4N}(nL(n))^{1/2}\right),$$

where N is some large positive integer, which will be specified later on. Then

$$\sum_{n=1}^{\infty} \frac{1}{n} P\left(\left|\sum_{i=-\infty}^{\infty} a_{ni}^{+} Y_i\right| > \epsilon(n(L(n))^{1/2})\right) \leq$$

$$\leq \sum_{n=1}^{\infty} \frac{1}{n} P\left(\left|\sum_{i=-\infty}^{\infty} Y_{ni}^{(1)}\right| \geq \frac{\epsilon}{4}(nL(n))^{1/2}\right) +$$

$$+\sum_{n=1}^{\infty} \frac{1}{n} P\left(\left|\sum_{i=-\infty}^{\infty} Y_{ni}^{(2)}\right| \geq \frac{\epsilon}{4}(nL(n))^{1/2}\right) +$$

$$+\sum_{n=1}^{\infty} \frac{1}{n} P\left(\left|\sum_{i=-\infty}^{\infty} Y_{ni}^{(3)}\right| \geq \frac{\epsilon}{4}(nL(n))^{1/2}\right) +$$

$$+\sum_{n=1}^{\infty} \frac{1}{n} P\left(\left|\sum_{i=-\infty}^{\infty} Y_{ni}^{(4)}\right| \geq \frac{\epsilon}{4}(nL(n))^{1/2}\right) =: Q_1 + Q_2 + Q_3 + Q_4.$$

From (3.1), similarly to the proof of $J_4 < \infty$ we can get $Q_4 < \infty$. From the definition of $Y_{ni}^{(2)}$ we know that $Y_{ni}^{(2)} > 0$, hence taking $N > 1/(\eta - \alpha)$ and noticing that $\sum_{j=1}^{\infty} (\#I_{nj}) j^{-\delta} \ll n$ for $\delta > 0$ (the definition of I_{nj} is as in the proof of Theorem 2.2),

$$Q_2 = \sum_{n=1}^{\infty} \frac{1}{n} P\left(\sum_{i=-\infty}^{\infty} Y_{ni}^{(2)} \geq \frac{\epsilon}{4}(nL(n))^{1/2}\right) \leq$$

$$\leq \sum_{n=1}^{\infty} \frac{1}{n} \left[\sum_{i=-\infty}^{\infty} P(a_{ni}^{+} Y_i > \lambda_n)\right]^N \leq$$

$$\leq \sum_{n=1}^{\infty} \frac{1}{n} \left[\sum_{j=1}^{\infty} (\#I_{nj}) P\left(\frac{Y_1^2}{(L(|Y_1|))^{1-\eta}} \geq \right.\right.$$

$$\left.\left. \geq \frac{CnL^{1-\alpha}(n)(j+1)L(j+1)}{(L(nL^{1-\alpha}(n)) + L((j+1)L(j+1)))^{1-\eta}}\right)\right]^N \ll$$

$$\ll \sum_{n=1}^{\infty} \frac{1}{n} \left\{ \sum_{j=1}^{\infty} (\# I_{nj}) \Big[(nj L(j+1))^{-1} (L(n))^{-(\eta-\alpha)} + \right.$$

$$\left. + (nj L^{\eta}(j+1))^{-1} (L(n))^{-(1-\alpha)} \Big] \right\} \ll$$

$$\ll \sum_{n=1}^{\infty} \frac{1}{n} \left[(L(n))^{-N(-(\eta-\alpha)} + (L(n))^{-N(1-\alpha)} \right] < \infty.$$

Similarly, $Y_{ni}^{(3)} < 0$ and $Q_3 < \infty$. By $EY_1 = 0$ and $E[Y_1^2/(L(|Y_1|))^{1-\eta}] <$ $< \infty$, we get

$$\left| \sum_{i=-\infty}^{\infty} EY_{ni}^{(1)} \right| / (nL(n))^{1/2} \leq$$

$$\leq \frac{1}{(nL(n))^{1/2}} \sum_{i=-\infty}^{\infty} [n^{1/2} L^{(1-\alpha)/2}(n) P(|a_{ni}^{+} Y_i| > n^{1/2} L^{(1-\alpha)/2}(n)) +$$

$$+ E|a_{ni}^{+} Y_i| I(|a_{ni}^{+} Y_i| > n^{1/2} L^{(1-\alpha)/2}(n))] \leq$$

$$\leq \frac{1}{(nL(n))^{1/2}} \sum_{j=1}^{\infty} (\# I_{nj}) [n^{1/2} L^{(1-\alpha)/2}(n) P(Y_1^2 > nL^{1-\alpha}(n)(j+1)L(j+1)) +$$

$$+ ((j+1)L(j+1))^{-1/2} E|Y_1| I(Y_1^2 > nL^{1-\alpha}(n)(j+1)L(j+1)) \ll$$

$$\ll (L(n))^{-(\eta-\alpha/2)} + (L(n))^{-(1-\alpha/2)} \to 0 \qquad \text{as } n \to \infty.$$

Thus, to prove $Q_1 < \infty$, it suffices to show that

$$Q_1 * =: \sum_{n=1}^{\infty} \frac{1}{n} P \left(\left| \sum_{i=-\infty}^{\infty} (Y_{ni}^{(1)} - EY_{ni}^{(1)}) \right| \geq \epsilon \cdot (nL(N))^{1/2} \right) < \infty, \ \forall \epsilon > 0.$$

Since, for each $n \geq 1$, $\{Y_{ni}, -\infty < i < \infty\}$ remains a sequence of NA r.v.'s, using Lemma 2, choose $p > \max\{2/\eta, 2(1-\eta)/\alpha + 2\}$ we have

$$Q_1 * \ll \sum_{n=1}^{\infty} \frac{1}{n} \cdot (nL(n))^{-p/2} \left\{ \left(\sum_{i=-\infty}^{\infty} E|Y_{ni}^{(1)}|^2 \right)^{p/2} + \sum_{i=-\infty}^{\infty} E|Y_{ni}^{(1)}|^p \right\} =:$$

$$=: Q_5 + Q_6.$$

While

$$Q_5 \le$$

$$\le \sum_{n=1}^{\infty} \frac{1}{n} \cdot (nL(n))^{-p/2} \left\{ \sum_{i=-\infty}^{\infty} [nL^{1-\alpha}(n)P(|a_{ni}^{+} Y_i| > n^{1/2}L^{(1-\alpha)/2}(n)) + \right.$$

$$\left. +E|a_{ni}^{+} Y_i|^2 I(|a_{ni}^{+} Y_i| \le n^{1/2}L^{(1-\alpha)/2}(n))] \right\}^{p/2} \ll$$

$$\ll \sum_{n=1}^{\infty} \frac{1}{n} \cdot (nL(n))^{-p/2} \left\{ \sum_{j=1}^{\infty} (\# I_{nj})[nL^{1-\alpha}(n)P(Y_1^2 > CnL^{1-\alpha}(n)jL(j)) + \right.$$

$$\left. +(jL(j))^{-1}E(Y_1^2/(L(|Y_1|))^{1-\eta}) \cdot (L(|Y_1|))^{1-\eta} I(Y_1^2 \le CnL^{1-\alpha}(n)jL(j))] \right\}^{p/2} \ll$$

$$\ll \sum_{n=1}^{\infty} \frac{1}{n} [(L(n))^{-p\eta/2} + (L(n))^{-p/2}] < \infty.$$

$$Q_6 \le \sum_{n=1}^{\infty} \frac{1}{n} \cdot (nL(n))^{-p/2} \sum_{i=-\infty}^{\infty} [E|a_{ni}^{+} Y_i|^p I(|a_{ni}^{+} Y_i| \le n^{1/2}L^{(1-\alpha)/2}(n)) +$$

$$+n^{p/2}L^{(1-\alpha)p/2}(n)P(|a_{ni}^{+} Y_i| > n^{1/2}L^{(1-\alpha)/2}(n))] \ll$$

$$\ll \sum_{n=1}^{\infty} \frac{1}{n} \cdot (nL(n))^{-p/2} \sum_{j=1}^{\infty} (\# I_{nj})[(jL(j))^{-p/2}E|Y_1|^p$$

$$I(Y_1^2 \le CnL^{1-\alpha}(n)jL(j)) + n^{p/2}L^{(1-\alpha)p/2}(n)P(Y_1^2 > CnL^{1-\alpha}(n)jL(j))] \le$$

$$\le \sum_{n=1}^{\infty} \frac{1}{n} \cdot (L(n))^{-p/2} \sum_{j=1}^{\infty} (\# I_{nj}) \left[(jL(j))^{-p/2} \cdot E\left(\frac{Y^2}{(L(|Y_1|))^{1-\eta}} \right) \cdot \right.$$

$$\cdot |Y_1|^{p-2} (L(|Y_1|))^{1-\eta} I(Y_1^2 \le CnL^{1-\alpha}(n)jL(j)) + n^{p/2}L^{(1-\alpha)p/2}(n) \cdot$$

$$\left. \cdot P\left(\frac{Y_1^2}{(L(|Y_1|))^{1-\eta}} \ge \frac{CnL^{1-\alpha}(n)jL(j)}{\left(L\left(nL^{1-\alpha}(n)\right) + L(jL(j))\right)^{1-\eta}} \right) \right] \ll$$

$$\ll \sum_{n=1}^{\infty} \frac{1}{n} \left\{ (L(n))^{-\left[\frac{p\alpha}{2} + (\eta - \alpha)\right]} + (L(n))^{-\left[\frac{p\alpha}{2} + (1-\alpha)\right]} \right\} < \infty.$$

ACKNOWLEDGEMENT. This research was partially supported by the National Natural Science Foundation of China (No. 10171079) and No. 2000-2-10400-1001-3 from the Basic Research Program of the KOSEF as well as Wongwang University Grant in 2002.

References

[1] K. ALAM, and K. M. L. SAXENA, Positive dependence in multivariate distributions, *Commun. Statist. Theor. Meth.,* **A10** (1981), 1183–1196,

[2] H. W. BLOCK, T. H. SAVITS and M. SHARKED: Some concepts of negative dependence, *Ann. Probab.,* **10** (1982), 765–772.

[3] R. M. BURTON and H. DEHLING, Large deviations for some weakly dependent random process, *Statist. & Probab. Lett.,* **9** (1990), 397–401.

[4] A. GUT, Convergence rates for probabilities of moderate deviations for sums of random variables with multidimensional indices, *Ann. Probab.,* **8**(2) (1980), 298–313.

[5] I. A. IBIRAGIMOV, Some limit theorems for stationary processed, *Theory Probab. Appl.,* **7** (1962), 349–382.

[6] K. JOAG-DEV and F. PROSCHAN, Negative association of random variables with applications, *Ann. Statist.,* **11** (1983), 286–295.

[7] D. L. LI. M. B. RAO and X. C. WANG, Complete convergence of moving average processes, *Staitst. & Probab. Lett.,* **14** (1992), 111–114.

[8] H. Y. LIANG and C. SU, Convergence in the law of logarithm for NA sequences, *Chinese Sci. Bull.,* **41**(15) (1999a), 1356–1360.

[9] H. Y. LIANG and C. SU, Complete convergence for weighted sums of NA sequences, *Statist. & Probab. Lett.,*. **45** (1999b), 85–95.

[10] H. Y. LIANG, Complete convergence for weighted sums of negatively associated random variable, *Statist. & Probab. Lett.,* **48** (2000), 317–325.

[11] P MATULA, A note on the almost sure convergence of sums of negatively dependent random variables, *Statist. & Probab. Lett.,* **15** (1992), 209–213.

[12] C. M. NEWMAN, *Asymptotic independence and limit theorems for positively and negatively dependent random variables,* in, Y. L. Tong, ed., (IMS, Hayward, CA) pp. 127–140 (1984).

[13] G. G. ROUSSAS, Asymptotic normality of random fields of positively or negatively associated random variables, *J. Multiv. Analysis,* **50** (1994), 152–172.

[14] Q. M. SHAO and C. SU, The law of the iterated logarithm for negatively associated random variables, *Stochastic Process Appl.,* **83** (1999), 139–148.

[15] C. SU, L. C. ZHAO and Y. B. WANG, Moment inequalities and week convergence for negatively associated sequences, (Series A), *Science in China,* **40** (1997), 172–182.

[16] L. X. ZHANG, Complete convergence of moving average processes under dependence assumptions, *Statist. & Probab. Lett.,* **30** (1996), 165–170.

INDEX

ISSN 0524-9007

Address:
MATHEMATICAL INSTITUTE, EÖTVÖS LORÁND UNIVERSITY
ELTE, TTK, KARI KÖNYVTÁR, MATEMATIKAI TUDOMÁNYÁGI SZAKGYŰJTEMÉNY
PÁZMÁNY PÉTER SÉTÁNY 1/c, 1117 BUDAPEST, HUNGARY
Műszaki szerkesztő:
Dr. SCHARNITZKY VIKTOR

A kiadásért felelős: az Eötvös Loránd Tudományegyetem rektora
A kézirat a nyomdába érkezett: 2002. május. Megjelent: 2002. június
Terjedelem: 18,3 A/4 ív. Példányszám: 500
Készült az EMTEX szedőprogram felhasználásával
az MSZ 5601–59 és 5602–55 szabványok szerint
Az elektronikus tipografálás Juhász Lehel és Fried Katalin munkája
Nyomda: Haxel Bt.
Felelős vezető: Korándi József