ACTA MATHEMATICA

ACADEMIAE SCIENTIARUM HUNGARICAE

TOMUS XXXIV FASCICULI 1—2

AKADÉMIAI KIADÓ, BUDAPEST

1979

ACTA MATH. HUNG.

ACTA MATHEMATICA
ACADEMIAE SCIENTIARUM HUNGARICAE

A MAGYAR TUDOMÁNYOS AKADÉMIA III. OSZTÁLYÁNAK
MATEMATIKAI KÖZLEMÉNYEI

SZERKESZTŐSÉG: 1053 BUDAPEST, REÁLTANODA U. 13—15.

KIADÓHIVATAL: 1363 BUDAPEST, PF. 24.

Az Acta Mathematica angol, német, francia és orosz nyelven közöl értekezéseket a matematika köréből. Váltakozó terjedelmű füzetekben jelenik meg, több füzet alkot egy kötetet. A közlésre szánt kéziratok a szerkesztőség, minden más levelezés a kiadóhivatal címére küldendő.

Megrendelhető a belföld számára az Akadémiai Kiadónál (1054 Budapest, Alkotmány utca 21. Bankszámla 215-11488), a külföld számára pedig a Kultúra Külkereskedelmi Vállalatnál (1011 Budapest, Fő utca 32. Bankszámla 218-10990), vagy annak külföldi képviseleteinél és bizományosainál.

ACTA MATHEMATICA

ACADEMIAE SCIENTIARUM HUNGARICAE

TOMUS XXXIV

AKADÉMIAI KIADÓ, BUDAPEST

1979

ACTA MATH. HUNG.

INDEX

TOMUS XXXIV

79-245—Szegedi Nyomda — F.v.: Dobó József igazgató

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 1—2.

KANTENKRÜMMUNG UND UMKUGELRADIUS KONVEXER POLYEDER

Von

J. LINHART (Salzburg)

Sei P ein konvexes Polyeder im $\mathbf{R}^3$ mit Kantenlängen $l_1, \dots, l_k$. α_i sei der Außenwinkel bei der i-ten Kante. Dann heißt $M=\frac{1}{2}\sum \alpha_i l_i$ die Kantenkrümmung von P.

SATZ. *Für jedes konvexe Polyeder mit f Flächen, k Kanten, e Ecken und Umkugelradius R gilt*

$$M/R \leqq 2k \sin\frac{\pi f}{2k}\left(1-\cot^2\frac{\pi f}{2k}\cot^2\frac{\pi e}{2k}\right)^{1/2} \arccos\left(\cos\frac{\pi e}{2k}\operatorname{cosec}\frac{\pi f}{2k}\right)$$

mit Gleichheit genau für die fünf regulären Polyeder.

BEMERKUNG. Dieser Satz wurde von A. FLORIAN [2] unter der zusätzlichen Voraussetzung bewiesen, daß die Fußpunkte der vom Umkugelmittelpunkt auf die Flächen und Kanten von P gefällten Lote auf den entsprechenden Flächen bzw. Kanten liegen. Im folgenden wird mit Hilfe einer allgemeinen Ungleichung von L. FEJES TÓTH [1] ein relativ kurzer Beweis angegeben, der ohne diese „Fußpunktbedingung" auskommt. Es sei weiters bemerkt, daß M/R in [5] mit einer ähnlichen Methode nach unten abgeschätzt wird $M \geqq 2\pi R$, mit Gleichheit genau dann, wenn P zu einer Strecke ausgeartet ist.

BEWEIS. Wir wählen den Umkugelmittelpunkt O als Ursprung und identifizieren Punkte und Ortsvektoren. Ohne Beschränkung der Allgemeinheit sei $R=1$. Es ist

$$M = \int_U t(u)\,du,$$

wobei $t(u)=\max_{x\in P}\langle u, x\rangle$ die Stützfunktion von P, und U die Einheitssphäre (mit Mittelpunkt O) ist [4]. Sind $v_1, \dots, v_e$ die Ecken von P, so lassen sich die Flächen des polaren Polyeders $P^*=\{y|\langle y, x\rangle \leqq 1$ für alle $x\in P\}$ so darstellen [3]: $F_i=\{y\in P^*|\langle y, v_i\rangle=1\}$ $(1\leqq i\leqq e)$. Wenn O nicht im Innern von P liegt, so ist O Randpunkt von P, und P^* ist unbeschränkt. O kann jedoch keine Ecke von P sein, daher enthält dann P^* einen höchstens zweidimensionalen Kegel mit Spitze O. $F_i'=\{u\in U|u=\frac{y}{\|y\|}$ für ein $y\in F_i\}$ ist die Projektion von F_i auf U. Für $u\in F_i'$ ist $t(u)=\langle u, v_i\rangle \leqq \cos\sphericalangle(u, v_i)$, da $\|v_i\|\leqq 1$. Die F_i' sind konvexe sphärische Polygone, die auf U ein Mosaik mit e Flächen, k Kanten und f Ecken bilden. Da cos in $\left[0, \frac{\pi}{2}\right]$ streng abnehmend ist, erhalten wir auf Grund einer Ungleichung von L. FEJES TÓTH ([1], S. 137);

$$M \leqq \sum_i \int_{F_i'} \cos \sphericalangle (u, v_i)\, du \leqq 4k \int_{\Delta} \cos \sphericalangle (u, A)\, du,$$

wobei $\Delta = ABC$ ein rechtwinkeliges sphärisches Dreieck mit den Winkeln $\alpha = \frac{\pi e}{2k}$ und $\beta = \frac{\pi f}{2k}$ bei A bzw. B bedeutet. a, b, c seien die Seitenlängen von Δ ($a = \widehat{BC} = \sphericalangle (B, C)$, usw.). Das oben rechts stehende Integral ist gleich dem Inhalt der Normalprojektion Δ' von Δ auf eine zu A orthogonale Ebene. Δ' ist ein Sektor einer Ellipse, dessen Inhalt $|\Delta'|$ leicht berechnet werden kann:

$$|\Delta'| = \frac{a}{2} \sin b = \frac{1}{2} \sin \beta (1 - \cot^2 \beta \cot^2 \alpha)^{1/2} \arccos (\cos \alpha \operatorname{cosec} \beta).$$

Damit folgt sofort die Richtigkeit des Satzes, da der Fall der Gleichheit schon bei der zitierten Ungleichung von L. FEJES TÓTH dargelegt ist.

Literaturverzeichnis

[1] L. FEJES TÓTH, *Lagerungen in der Ebene, auf der Kugel und im Raum,* 2. Aufl., Springer-Verlag (Berlin, Heidelberg, New York, 1972).
[2] A. FLORIAN, Ungleichungen über konvexe Polyeder, *Monatsh. Math.,* **60** (1956), 288—297.
[3] B. GRÜNBAUM, *Convex Polytopes,* Interscience (London, New York, Sydney, 1967).
[4] H. HADWIGER, *Vorlesungen über Inhalt, Oberfläche und Isoperimetrie,* Springer-Verlag (Berlin, Göttingen, Heidelberg, 1957).
[5] J. LINHART, Kantenlängensumme, mittlere Breite und Umkugelradius konvexer Polytope, *Archiv der Math.* (eingereicht).

(Eingegangen am 28. März 1977)

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 3—15.

LIMITS OF BIFUNCTORS INTO A CATEGORY

By

I. A. ASSEM, C. G. CHEHATA and M. EL-GENDY (Alexandria)*

Introduction

Although covariant and contravariant functors of a small category into an arbitrary category $\mathscr{C}$, also known as inductive and projective systems, have been widely studied, nothing is known of bifunctors of the product of two small categories into $\mathscr{C}$. It is the purpose of this paper to study such bifunctors, which we call mixed systems. We shall prove that mixed systems have, in a certain natural manner, two limits, and that there exists a natural morphism from one of these limits into the other. The limits and the natural morphisms are proved to satisfy a certain universal property. We also define a category of mixed systems and investigate some of its properties. In forthcoming papers, we shall take a closer look at mixed systems of sets and of modules.

Notations have been made to follow [1] as closely as possible. For results about categories and functors, we refer the reader to [2], [3] and [5]. Results on projective and inductive systems can be found in [4].

In what follows, we shall work within a category $\mathscr{C}$ such that every functor of a small category into $\mathscr{C}$ has projective and inductive limits.

1. The mixed system and its limits

Let I and L be two pre-ordered sets. We define a mixed system in $\mathscr{C}$ over $I\times L$ to be a bifunctor of $I\times L$ (with the product pre-order) into $\mathscr{C}$, contravariant in the first and covariant in the second variable. More explicitly:

DEFINITION 1.1. A mixed system $(E_\alpha^\lambda, f_{\alpha\beta}^{\mu\lambda})_{I\times L}$ in $\mathscr{C}$ over $I\times L$ is defined by the following;

i) to each ordered pair $(\alpha, \lambda)\in I\times L$, we associate an object E_α^λ of $\mathscr{C}$.

ii) to each couple of pairs (α, μ) and (β, λ) of $I\times L$ such that $\alpha\leqq\beta$ and $\lambda\leqq\mu$, there is associated a morphism $f_{\alpha\beta}^{\mu\lambda}\colon E_\beta^\lambda\to E_\alpha^\mu$ of $\mathscr{C}$ such that:

(MS1) for every $(a, \lambda)\in I\times L$, we have $f_{\alpha\alpha}^{\lambda\lambda}=1_{E_\alpha^\lambda}$,

(MS2) if $\alpha\leqq\beta\leqq\gamma$ in I and $\lambda\leqq\mu\leqq\nu$ in L, then $f_{\alpha\gamma}^{\nu\lambda}=f_{\alpha\beta}^{\nu\mu}f_{\beta\gamma}^{\mu\lambda}$.

The assumption that I and L are sets is not essential. It was made for the simplicity of notations and the whole theory can be developed in case I and L are arbitrary categories. Of course, we would then have to assume that $\mathscr{C}$ is such that every fuctor into $\mathscr{C}$ has projective and inductive limits.

* Research was carried out at the Science Centre for Advancement of Post Graduate Studies, University of Alexandria. The authors wish to express their thanks to Professor R. Wiegandt who gave the idea of a fruitful generalization of their original work.

To simplify notations, we shall denote the morphisms $f_{\beta\beta}^{\lambda\lambda}$ by $g_{\alpha\beta}^{\lambda}$ and the morphisms $f_{\alpha\alpha}^{\mu\lambda}$ by $h_{\alpha}^{\mu\lambda}$. We now construct the limits of the mixed system $(E_{\alpha}^{\lambda}, f_{\alpha\beta}^{\mu\lambda})_{I\times L}$.

Let $\lambda\in L$ be fixed and $\alpha\leqq\beta\leqq\gamma$ in I, then by (MS1) and (MS2),

(1.1) $$g_{\alpha\alpha}^{\lambda} = 1_{E_{\alpha}^{\lambda}}$$

(1.2) $$g_{\alpha\gamma}^{\lambda} = g_{\alpha\beta}^{\lambda}\, g_{\beta\gamma}^{\lambda}.$$

Thus the system $(E_{\alpha}^{\lambda}, g_{\alpha\beta}^{\lambda})_I$ is projective. By our assumption on $\mathscr{C}$, it must have a projective limit:

$$(E^{\lambda}, g_{\alpha}^{\lambda}) = \varprojlim_{\alpha\in I} (E_{\alpha}^{\lambda}, g_{\alpha\beta}^{\lambda}).$$

By definition, the canonical morphisms $(g_{\alpha}^{\lambda}\colon E^{\lambda}\to E_{\alpha}^{\lambda})_{\alpha\in I}$ are such that if $\alpha\leqq\beta$,

(1.3) $$g_{\alpha}^{\lambda} = g_{\alpha\beta}^{\lambda}\, g_{\beta}^{\lambda}.$$

Similarly, for any $\alpha\in I$, the system $(E_{\alpha}^{\lambda}, h_{\alpha}^{\mu\lambda})_L$ is inductive and so has an inductive limit:

$$(E_{\alpha}, h_{\alpha}^{\lambda}) = \varinjlim_{\lambda\in L} (E_{\alpha}^{\lambda}, h_{\alpha}^{\mu\lambda}).$$

The canonical morphisms $(h_{\alpha}^{\lambda};\ E_{\alpha}^{\lambda}\to E_{\alpha})_{\lambda\in L}$ are such that $\lambda\leqq\mu$ implies

(1.4) $$h_{\alpha}^{\lambda} = h_{\alpha}^{\mu}\, h_{\alpha}^{\mu\lambda}.$$

Let now $\alpha\leqq\beta$ in I and $\lambda\leqq\mu$ in L, then (MS2) implies

(1.5) $$h_{\alpha}^{\mu\lambda}\, g_{\alpha\beta}^{\lambda} = g_{\alpha\beta}^{\mu}\, h_{\beta}^{\mu\lambda}.$$

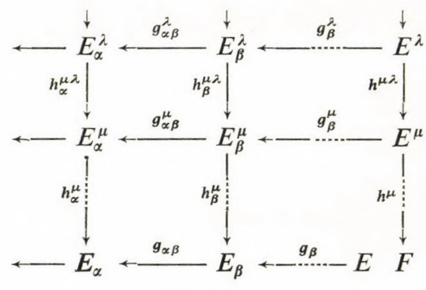


Fig. (1.1)

This equation may be interpreted in two ways. First it means that the $(h_{\alpha}^{\mu\lambda})_{\alpha\in I}$ form a projective system of morphisms of $(E_{\alpha}^{\lambda}, g_{\alpha\beta}^{\lambda})_I$ into $(E_{\alpha}^{\mu}, g_{\alpha\beta}^{\mu})_I$. Let its projective limit be:

$$h^{\mu\lambda} = \varprojlim_{\alpha\in I} h_{\alpha}^{\mu\lambda}.$$

By definition, $h^{\mu\lambda}$ is the only morphism of E^{λ} into E^{μ} such that, for any $a\in I$,

(1.6) $$g_{\alpha}^{\mu}\, h^{\mu\lambda} = h_{\alpha}^{\mu\lambda}\, g_{\alpha}^{\lambda}.$$

Next, (1.5) means also that the $(g^\lambda_{\alpha\beta})_{\lambda\in L}$ form an inductive system of morphisms of $(E^\lambda_\beta, h^{\mu\lambda}_\beta)_L$ into $(E^\lambda_\alpha, h^{\mu\lambda}_\alpha)_L$. Let its inductive limit be:

$$g_{\alpha\beta} = \varinjlim_{\lambda\in L} g^\lambda_{\alpha\beta}.$$

$g_{\alpha\beta}$ is the only morphism of E_β into E_α such that, for all $\lambda\in L$,

(1.7) $$g_{\alpha\beta}\, h^\lambda_\beta = h^\lambda_\alpha\, g^\lambda_{\alpha\beta}.$$

Now let $a\in I$ be arbitrary, and $\lambda\leqq\mu\leqq\nu$ in L. (MS2) gives

(1.8) $$h^{\nu\lambda}_\alpha = h^{\nu\mu}_\alpha h^{\mu\lambda}_\alpha.$$

Passing to the limit,

(1.9) $$h^{\nu\lambda} = h^{\nu\mu} h^{\mu\lambda}.$$

Also $h^{\lambda\lambda}_\alpha = 1_{E^\lambda_\alpha}$ implies $h^{\lambda\lambda} = 1_{E^\lambda}$. Therefore $(E^\lambda, h^{\mu\lambda})_L$ is an inductive system, with limit

$$(F, h^\lambda) = \varinjlim_{\lambda\in L} (E^\lambda, h^{\mu\lambda}).$$

Then $\lambda\leqq\mu$ in L implies

(1.10) $$h^\lambda = h^\mu h^{\mu\lambda}.$$

Similarly, the system $(E_\alpha, g_{\alpha\beta})_I$ is projective. Let us put:

$$(E, g_\alpha) = \varprojlim_{\alpha\in I} (E_\alpha, g_{\alpha\beta}).$$

$\alpha\leqq\beta$ in I implies

(1.11) $$g_\alpha = g_{\alpha\beta}\, g_\beta.$$

The objects E and F of $\mathscr{C}$ will be called the *limits* of the system $(E^\lambda_\alpha, f^{\mu\lambda}_{\alpha\beta})_{I\times L}$. Thus we have:

THEOREM 1.1. *A mixed system* $(E^\lambda_\alpha, f^{\mu\lambda}_{\alpha\beta})_{I\times L}$ *has two limits*

$$E = \varprojlim_{\alpha\in I} \big(\varinjlim_{\lambda\in L} E^\lambda_\alpha\big) \quad \textit{and} \quad F = \varinjlim_{\lambda\in L} \big(\varprojlim_{\alpha\in I} E^\lambda_\alpha\big).$$

2. The canonical morphism and the universal property

Equation (1.6) can be interpreted to mean that the family $(g^\lambda_\alpha)_{\lambda\in L}$ is an inductive system of morphisms of $(E^\lambda, h^{\mu\lambda})_L$ into $(E^\lambda_\alpha, h^{\mu\lambda}_\alpha)_L$. Hence there exists a unique morphism $g'_\alpha;\ F\to E_\alpha$ such that, for any $\lambda\in L$,

(2.1) $$g'_\alpha h^\lambda = h^\lambda_\alpha\, g^\lambda_\alpha.$$

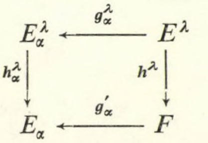


Fig. (2.1)

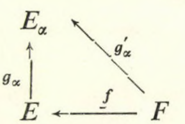


Fig. (2.2)

Let $(\alpha, \beta)\in I^2$, $\alpha\leqq\beta$, then (1.3) implies, by passing to the limit

$$g'_\alpha = g_{\alpha\beta}\, g'_\beta. \tag{2.2}$$

By definition of projective limit, there exists a unique morphism $\underline{f}\colon F\to E$ such that, for all $\alpha\in I$,

$$g'_\alpha = g_\alpha \underline{f}. \tag{2.3}$$

On the other hand, (1.7) means that $(h^\lambda_\alpha)_{\alpha\in I}$ is a projective system of morphisms mapping $(E^\lambda_\alpha, g^\lambda_{\alpha\beta})_I$ into $(E_\alpha, g_{\alpha\beta})_I$. Hence there exists a unique morphism $h'^\lambda\colon E^\lambda\to E$ such that, for all $\alpha\in I$,

$$g_\alpha h'^\lambda = h^\lambda_\alpha g^\lambda_\alpha. \tag{2.4}$$

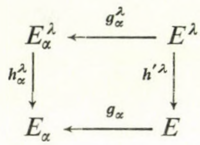


Fig. (2.3)

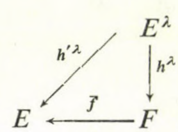


Fig. (2.4)

Now, if $(\lambda, \mu)\in L^2$, $\lambda\leqq\mu$, (1.4) implies, by passing to the limit,

$$h'^\lambda = h'^\mu h^{\mu\lambda}. \tag{2.5}$$

Therefore there exists a unique morphism $\bar f\colon F\to E$ such that, for all $\lambda\in L$,

$$\bar f h^\lambda = h'^\lambda. \tag{2.6}$$

We shall now prove that $\bar f=\underline{f}$. (2.1) and (2.4) imply, for any $(\alpha, \lambda)\in I\times L$, $g_\alpha h'^\lambda=g'_\alpha h^\lambda$; hence, by (2.3) and (2.6), $g_\alpha \bar f h^\lambda=g_\alpha \underline{f} h^\lambda$. Now the family $(g_\alpha)_{\alpha\in I}$ is monomorphic, and the family $(h^\lambda)_{\lambda\in L}$ is epimorphic (cf. [3], p. 164). Therefore $\bar f=\underline{f}$. This unique morphism will be denoted from now on by f and called the *canonical morphism*. Thus

THEOREM 2.1. *Let E, F be the limits of the mixed system $(E^\lambda_\alpha, f^{\mu\lambda}_{\alpha\beta})_{I\times L}$. There exists a unique morphism $f\colon F\to E$ which satisfies, for any $(\alpha, \lambda)\in I\times L$*

$$g_\alpha f h^\lambda = h^\lambda_\alpha g^\lambda_\alpha. \tag{2.7}$$

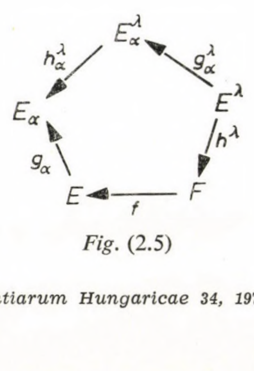


Fig. (2.5)

In other words, f is the only morphism which makes the pentagon of figure (2.5) commutative. This pentagon will be called in the sequel *canonical pentagon.* We note here that f is not an isomorphism in general (cf. [3], p. 197).

We now give another characterization of the limits and the canonical morphism:

DEFINITION 2.1. Let $\mathcal{E}=(E^\lambda_\alpha, f^{\mu\lambda}_{\alpha\beta})_{I\times L}$ be a mixed system in $\mathcal{C}$ over $I\times L$, and let:

$$(E_\alpha, h^\lambda_\alpha) = \varinjlim_{\lambda\in L}(E^\lambda_\alpha, f^{\mu\lambda}_{\alpha\alpha}), \quad g_{\alpha\beta} = \varinjlim_{\lambda\in L} f^{\lambda\lambda}_{\alpha\beta}$$

$$(E^\lambda, g^\lambda_\alpha) = \varprojlim_{\alpha\in I}(E^\lambda_\alpha, f^{\lambda\lambda}_{\alpha\beta}), \quad h^{\mu\lambda} = \varprojlim_{\alpha\in I} f^{\lambda\lambda}_{\alpha\alpha}.$$

A triple (E, F, f) where E and F are objects of $\mathcal{C}$, and $f\colon F\to E$ a morphism, will be called *a limit triple of* $\mathcal{E}$ if:

(ML1) There exist two families of morphisms $(g_\alpha\colon E\to E_\alpha)_{\alpha\in I}$ and $(h^\lambda; E^\lambda\to F)_{\lambda\in L}$ such that:

(i) $g_\alpha = g_{\alpha\beta}g_\beta$ if $\alpha\leqq\beta$ in I,

(ii) $h^\lambda = h^\mu h^{\mu\lambda}$ if $\lambda\leqq\mu$ in L,

(iii) $g_\alpha f h^\lambda = h^\lambda_\alpha g^\lambda_\alpha$ for any $(\alpha,\lambda)\in I\times L$.

(ML2) If (E', F', f') is another triple, with E', F' objects in $\mathcal{C}$, and $f'\colon F'\to E'$, and there exist families $(g'_\alpha\colon E'\to E_\alpha)_{\alpha\in I}$ and $(h'^\lambda\colon E^\lambda\to F')_{\lambda\in L}$ of morphisms of $\mathcal{C}$ which satisfy:

(i) $g'_\alpha = g_{\alpha\beta}g'_\beta$ if $\alpha\leqq\beta$ in I,

(ii) $h'^\lambda = h'^\mu h^{\mu\lambda}$ if $\lambda\leqq\mu$ in L,

(iii) $g'_\alpha f' h'^\lambda = h^\lambda_\alpha g^\lambda_\alpha$ for any $(\alpha,\lambda)\in I\times L$,

then there exists a unique pair (g, h) of morphisms of $\mathcal{C}$ such that $f=gf'h$.

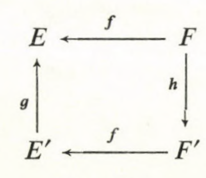


Fig. (2.6)

LEMMA 2.1. *If a limit triple exists, it is unique up to isomorphism.*

PROOF. Let (E, F, f) and (E', F', f') be two limit triples of the same mixed system. There exist unique pairs (g, h) and (g', h') of morphisms of $\mathcal{C}$ such that if $f=gf'h$ and $f'=g'fh'$ then $f=(gg')f(h'h)$.

By uniqueness, $gg'=1_E$ and $h'h=1_F$. Similarly, $g'g=1_{E'}$, and $hh'=1_{F'}$. Hence the result.

THEOREM 2.2. *The triple* (E, F, f) *where*

$$E=\varprojlim_{\alpha\in I}\varinjlim_{\lambda\in L}E^\lambda_\alpha, \quad F=\varinjlim_{\lambda\in L}\varprojlim_{\alpha\in I}E^\lambda_\alpha$$

and $f: F \to E$ is the canonical morphism, is the only (up to isomorphism) limit triple of the system $(E_\alpha^\lambda, f_{\alpha\beta}^{\mu\lambda})_{I\times L}$.

PROOF. It has been shown before that the triple (E, F, f) satisfies (ML1). We now prove (ML2). Assume that (E', F', f') satisfies the conditions of (ML2) then, since $E = \varprojlim_{\alpha\in I} E_\alpha$, there exists a unique morphism $g: E' \to E$ such that, for all $\alpha\in I$, $g'_\alpha = g_\alpha g$. Similarly, there exists a unique $h: F \to F'$ such that, for all $\lambda\in L$, $h'^\lambda = hh^\lambda$.

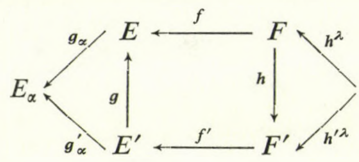


Fig. (2.7)

Hence, for any $(\alpha, \lambda)\in I\times L$,

$$g_\alpha gf' hh^\lambda = g'_\alpha f' h'^\lambda = h_\alpha^\lambda g_\alpha^\lambda = g_\alpha fh^\lambda.$$

Since the family $(g_\alpha)_{\alpha\in I}$ is monomorphic, and the family $(h^\lambda)_{\lambda\in L}$ is epimorphic, then $gf'h = f$. The uniqueness assertion follows from Lemma (2.1).

COROLLARY 2.1. *Let $\mathscr{C}$ be a category such that every functor of a small category into $\mathscr{C}$ has projective and inductive limits. Then every mixed system in $\mathscr{C}$ has a unique limit triple.*

3. The category of mixed systems

DEFINITION 3.1. Let $\mathscr{E} = (E_\alpha^\lambda, f_{\alpha\beta}^{\mu\lambda})_{I\times L}$ and $\mathscr{E}' = (E_\alpha'^\lambda, f_{\alpha\beta}'^{\mu\lambda})_{I\times L}$ be two mixed systems in $\mathscr{C}$ over $I\times L$. A family $u = (u_\alpha^\lambda)_{(\alpha,\lambda)\in I\times L}$ of morphisms of $\mathscr{C}$ will be called a *mixed system of morphisms* of $\mathscr{E}$ into $\mathscr{E}'$ if for every pair of relations $\alpha\leqq\beta$ in I and $\lambda\leqq\mu$ in L, we have:

$$u_\alpha^\mu f_{\alpha\beta}^{\mu\lambda} = f_{\alpha\beta}'^{\mu\lambda} u_\beta^\lambda. \tag{3.1}$$

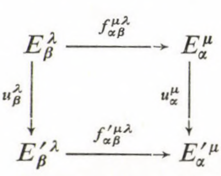


Fig. (3.1)

In other words, u is a functorial morphism of $\mathscr{E}$ into $\mathscr{E}'$.

DEFINITION 3.2. The category $\mathscr{M}(I\times L, \mathscr{C})$ of mixed systems in $\mathscr{C}$ over $I\times L$ has for objects all such systems, and for morphisms, the mixed systems of morphisms.

The product of two morphisms $u=(u_\alpha^\lambda)_{I\times L}\colon \mathscr{E}\to\mathscr{E}'$ and $v=(v_\alpha^\lambda)_{I\times L}\colon \mathscr{E}'\to\mathscr{E}''$ is defined to be $vu=(v_\alpha^\lambda u_\alpha^\lambda)_{I\times L}\colon \mathscr{E}\to\mathscr{E}''$.

We shall now prove that to a mixed system of morphisms correspond two limit morphisms in $\mathscr{C}$. Equation (3.1) shows that the $(u_\alpha^\lambda)_{\lambda\in L}$ form an inductive system of morphisms mapping the system $(E_\alpha^\lambda, h_\alpha^{\mu\lambda})_L$ into $(E_\alpha'^\lambda, h_\alpha'^{\mu\lambda})_L$. Let $u_\alpha=\varinjlim_{\lambda\in L} u_\alpha^\lambda$. This is, by definition, the only morphism of E_α into E_α' such that, for all $\lambda\in L$

(3.2) $$u_\alpha h_\alpha^\lambda = h_\alpha'^\lambda u_\alpha^\lambda .$$

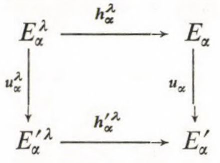


Fig. (3.2)

Let now $\alpha\leqq\beta$ in I and $\lambda\in L$ be arbitrary, then:

$$\begin{aligned} u_\alpha g_{\alpha\beta} h_\beta^\lambda &= u_\alpha h_\alpha^\lambda g_{\alpha\beta}^\lambda && (1.7)\\ &= h_\alpha' u_\alpha^\lambda g_{\alpha\beta}^\lambda && (3.2)\\ &= h_\alpha' g_{\alpha\beta}'^\lambda u_\beta^\lambda && (3.1)\\ &= g_{\alpha\beta}' h_\beta'^\lambda u_\beta^\lambda && (1.7)\\ &= g_{\alpha\beta}' u_\beta h_\beta^\lambda . && (3.2)\end{aligned}$$

Since the family $(h_\beta^\lambda)_{\lambda\in L}$ is epimorphic,

(3.3) $$u_\alpha g_{\alpha\beta} = g_{\alpha\beta}' u_\beta$$

and the family $(u_\alpha)_{\alpha\in I}$ is a projective system of morphisms of $(E_\alpha, g_{\alpha\beta})_I$ into $(E_\alpha', g_{\alpha\beta}')_I$. Therefore there exists a unique morphism $u_- =\varprojlim_{\alpha\in I} u_\alpha$ of E into E' such that, for all $\alpha\in I$,

(3.4) $$g_\alpha' u_- = u_\alpha g_\alpha .$$

Thus, by definition, $u_- =\varprojlim_{\alpha\in I}\varinjlim_{\lambda\in L} u_\alpha^\lambda$.

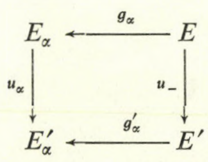


Fig. (3.3)

Similarly, (3.1) shows also that $(u_\alpha^\lambda)_{\alpha\in I}$ is a projective system of morphisms mapping $(E_\alpha^\lambda, g_{\alpha\beta}^\lambda)_I$ into $(E_\alpha'^\lambda, g_{\alpha\beta}'^\lambda)_I$, hence there exists a unique morphism

$u^\lambda = \varprojlim_{\alpha \in I} u_\alpha^\lambda \colon E^\lambda \to E'^\lambda$ such that for all $\alpha \in I$,

(3.5) $$g_\alpha'^\lambda u^\lambda = u_\alpha^\lambda g_\alpha^\lambda.$$

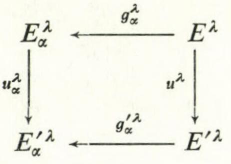


Fig. (3.4)

Let now $\lambda \leqq \mu$ in L and $\alpha \in I$ be arbitrary, then we have

$$g_\alpha'^\mu u^\mu h^{\mu\lambda} = u_\alpha^\mu g_\alpha^\mu h^{\mu\lambda} \quad (3.5)$$

$$= u_\alpha^\mu h_\alpha^{\mu\lambda} g_\alpha^\lambda \quad (1.6)$$

$$= h_\alpha'^{\mu\lambda} u_\alpha^\lambda g_\alpha^\lambda \quad (3.1)$$

$$= h_\alpha'^{\mu\lambda} g_\alpha'^\lambda u^\lambda \quad (3.5)$$

$$= g_\alpha'^\mu h'^{\mu\lambda} u^\lambda. \quad (1.6)$$

Since $(g_\alpha'^\mu)_{\alpha \in I}$ is a monomorphic family,

(3.6) $$u^\mu h^{\mu\lambda} = h'^{\mu\lambda} u^\lambda,$$

the $(u^\lambda)_{\lambda \in L}$ form thus an inductive system of morphisms of $(E^\lambda, h^{\mu\lambda})_L$ into $(E'^\lambda, h'^{\mu\lambda})_L$. Therefore there exists a unique morphism $u_+ = \varinjlim_{\lambda \in L} u^\lambda$ such that, for all $\lambda \in L$,

(3.7) $$u_+ h^\lambda = h'^\lambda u^\lambda.$$

We have thus, by definition, $u_+ = \varinjlim_{\lambda \in L} \varprojlim_{\alpha \in I} u_\alpha^\lambda$.

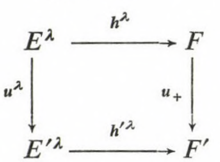


Fig. (3.5)

We have proved:

THEOREM 3.1. *A mixed system of morphisms $u\colon \mathscr{E} \to \mathscr{E}'$ defines two morphisms $u_-\colon E \to E'$ and $u_+\colon F \to F'$ of their limits, called the mixed limit morphisms.*

THEOREM 3.2. *The mixed limit morphisms and the canonical morphisms satisfy the commutativity relation*

(3.8) $$f' u_+ = u_- f.$$

PROOF.

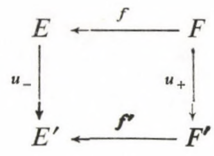


Fig. (3.6)

For any $(\alpha, \lambda)\in I\times L$, we have,

$$g'_\alpha f' u_+ h = g'_\alpha f' h'^\lambda u^\lambda \tag{3.7}$$

$$= h'^\lambda_\alpha g'^\lambda_\alpha u^\lambda \tag{2.7}$$

$$= h'^\lambda_\alpha u^\lambda_\alpha g^\lambda_\alpha \tag{3.5}$$

$$= u_\alpha h^\lambda_\alpha g^\lambda_\alpha \tag{3.2}$$

$$= u_\alpha g_\alpha f h^\lambda \tag{2.7}$$

$$= g'_\alpha u_- f h^\lambda. \tag{3.4}$$

Since the families $(g'_\alpha)_{\alpha\in I}$ and $(h^\lambda)_{\lambda\in L}$ are monomorphic and epimorphic, respectively, the result follows.

THEOREM 3.3. *Let* $u=(u^\lambda_\alpha)_{I\times L}\colon \mathscr{E}\to\mathscr{E}'$ *and* $v=(v^\lambda_\alpha)_{I\times L}\colon \mathscr{E}'\to\mathscr{E}''$ *be mixed systems of morphisms and* $w=vu$, *then:*

(3.9) $$w_- = v_- u_-,$$

(3.10) $$w_+ = v_+ u_+ .$$

Moreover, if $u^\lambda_\alpha = 1_{E^\lambda_\alpha}$ *for every pair* $(a, \lambda)\in I\times L$, *then*

(3.11) $$u_- = 1_E,$$

(3.12) $$n_+ = 1_F .$$

PROOF. This follows from the functorial properties of projective and inductive limits (cf. [4]).

Obviously, this implies:

COROLLARY 3.1. *If* $u\colon \mathscr{E}\to\mathscr{E}'$ *is a mixed system of isomorphisms, then* u_+ *and* u_- *are isomorphisms.*

COROLLARY 3.2. *Let* $\mathscr{C}$ *be the category of (left or right) modules over some ring R, and L be directed. Then if* $u=(u^\lambda_\alpha)_{I\times L}\colon \mathscr{E}'\to\mathscr{E}$ *and* $v=(v^\lambda_\alpha)_{I\times L}\colon \mathscr{E}\to\mathscr{E}''$ *are mixed systems of linear maps such that the sequence*

$$0 \to E'^\lambda_\alpha \xrightarrow{u^\lambda_\alpha} E^\lambda_\alpha \xrightarrow{v^\lambda_\alpha} E''^\lambda_\alpha$$

is exact for any $(\alpha, \lambda) \in I \times L$, *then the sequences:*

$$0 \to E' \xrightarrow{u_-} E \xrightarrow{v_-} E''$$

and

$$0 \to F' \xrightarrow{u_+} F \xrightarrow{v_+} F''$$

are exact.

PROOF. This follows from the exactness of the inductive limit functor and the left-exactness of the projective limit functor.

DEFINITION 3.3. We now define two covariant functors 1_+ and 1_- of $\mathcal{M}(I \times L, \mathcal{C})$ into $\mathcal{C}$ as follows:

i) $1_+ = \varinjlim_{\lambda \in L} \varprojlim_{\alpha \in I} (-)$ associates to a mixed system $\mathcal{E}$ its limit $F = \varinjlim_{\lambda \in L} \varprojlim_{\alpha \in I} E_\alpha^\lambda$ and to a morphism $u\colon \mathcal{E} \to \mathcal{E}'$ its limit $u_+ = \varinjlim_{\lambda \in L} \varprojlim_{\alpha \in I} u_\alpha^\lambda\colon F \to F'$.

ii) $1_- = \varprojlim_{\alpha \in I} \varinjlim_{\lambda \in L} (-)$ associates to $\mathcal{E}$ its limit $E = \varprojlim_{\alpha \in I} \varinjlim_{\lambda \in L} E_\alpha^\lambda$ and to u its limit $u_- = \varprojlim_{\alpha \in I} \varinjlim_{\lambda \in L} u_\alpha^\lambda\colon E \to E'$.

Theorem (3.3) shows that these are indeed covariant functors.

COROLLARY 3.3. *There exists a functorial morphism* $\varphi\colon 1_+ \to 1_-$ *defined as follows: to each mixed system* $\mathcal{E}$ *we associate its canonical morphism* $\varphi = f\colon 1_+(\mathcal{E}) \to 1_-(\mathcal{E})$.

PROOF. This is exactly Theorem (3.2).

We now associate to the category $\mathcal{C}$ the category $\mathcal{D}$ of triples (A, B, f) with $A, B \in \mathcal{C}$ and $f\colon B \to A$. A morphism $(g, h)\colon (A, B, f) \to (A', B', f')$ is a pair of morphisms of $\mathcal{C}$ such that $g\colon A \to A'$, $h\colon B \to B'$, and $f'g = hg$.

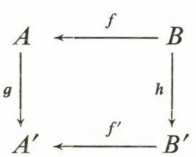


Fig. (3.7)

The product of $(g, h)\colon (A, B, f) \to (A', B', f')$ by $(g', h')\colon (A', B', f') \to (A'', B'', f'')$ is defined to be:

$$(g', h')(g, h) = (g'g, h'h).$$

DEFINITION 3.4. The covariant functor 1 of $\mathcal{M}(I \times L, \mathcal{C})$ into $\mathcal{D}$ is defined to be the functor which associates to a mixed system $\mathcal{E}$ its limit triple (E, F, f) and to a mixed system of morphisms $u\colon \mathcal{E} \to \mathcal{E}'$ the pair of morphisms (u_-, u_+).

1 is indeed a covariant functor by Theorems (3.2) and (3.3).

Finally, Corollary (3.3) shows that 1_+, 1_- and 1 are left-exact.

4. Properties of $\mathcal{M}(I\times L, \mathcal{C})$

THEOREM 4.1. *If M is an arbitrary index set and $(\mathcal{E}_i)_{i\in M}$ is a family of mixed systems in $\mathcal{C}$ over $I\times L$, with $\mathcal{E}_i=(E^\lambda_{i\alpha}, f^{\mu\lambda}_{i\alpha\beta})_{I\times L}$, then the product of the $\mathcal{E}_i$'s exists in the category $\mathcal{M}(I\times L, C)$ and is equal to*

$$\mathcal{E} = (\prod_{i\in M} E^\lambda_{i\alpha}, \prod_{i\in M} f^{\mu\lambda}_{i\alpha\beta})_{I\times L}.$$

PROOF. Since $\mathcal{C}$ is such that every projective system has a limit, every family of objects of $\mathcal{C}$ has a product in $\mathcal{C}$ (cf. [2], p. 54).

Hence $E^\lambda_\alpha = \prod_{i\in M} E^\lambda_{i\alpha}$ exists. Let the maps $(p^\lambda_{i\alpha}\colon E^\lambda_\alpha \to E^\lambda_{i\alpha})_{i\in M}$ denote the canonical projections. The product $f^{\mu\lambda}_{\alpha\beta} = \prod_{i\in M} f^{\mu\lambda}_{i\alpha\beta}$ is by definition the only morphism of E^λ_β into E^μ_α such that, for all $i\in M$,

$$p^\mu_{i\alpha} f^{\mu\lambda}_{\alpha\beta} = f^{\mu\lambda}_{i\alpha\beta}\, p^\lambda_{i\beta}. \tag{4.1}$$

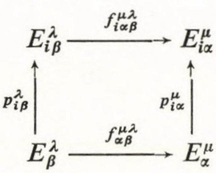


Fig. (4.1)

It is easy to prove that $\mathcal{E}=(E^\lambda_\alpha, f^{\mu\lambda}_{\alpha\beta})_{I\times L}$ is a mixed system. Let us put $p_i=(p^\lambda_{i\alpha})_{I\times L}$, then (4.1) shows that p_i is (for every $i\in M$) a mixed system of morphisms of $\mathcal{E}$ into $\mathcal{E}_i$. Let now $\mathcal{E}'=(E'^\lambda_\alpha, f'^{\mu\lambda}_{\alpha\beta})_{I\times L}$ be another mixed system, and $(q_i\colon \mathcal{E}'\to\mathcal{E}_i)_{i\in M}$ a family of morphisms, with $q_i=(q^\lambda_{i\alpha})_{I\times L}$. By definition of E^λ_α, for every pair $(\alpha, \lambda)\in_{I\times L}$, there exists a unique morphism t^λ_α; $E'^\lambda_\alpha \to E^\lambda_\alpha$ such that, for all $i\in M$,

$$p^\lambda_{i\alpha} t^\lambda_\alpha = q^\lambda_{i\alpha}. \tag{4.2}$$

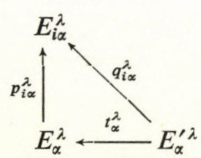


Fig. (4.2)

One can prove easily that $t=(t^\lambda_\alpha)_{I\times L}$ is a mixed system of morphisms of $\mathcal{E}'$ into $\mathcal{E}$. Finally, (4.2) shows that, for any $i\in M$,

$$p_i t = q_i, \tag{4.3}$$

and t is the only mixed system of morphisms which satisfies this relation (the uniqueness of t follows from that of each t^λ_α).

Dually,

THEOREM 4.2. *Let M be an arbitrary index set, and $(\mathcal{E}_i)_{i\in M}$ a family of mixed systems, with $\mathcal{E}_i=(E^{\lambda}_{i\alpha}, f^{\mu\lambda}_{i\alpha\beta})_{I\times L}$. Then the coproduct of the $\mathcal{E}_i$'s exists in $\mathcal{M}(I\times L, \mathcal{C})$ and is equal to* $\mathcal{E}=(\coprod_{i\in M} E^{\lambda}_{i\alpha}, \coprod_{i\in M} f^{\mu\lambda}_{i\alpha\beta})_{I\times L}$.

THEOREM 4.3. *Let $u=(u^{\lambda}_{\alpha})_{I\times L}$ and $v=(v^{\lambda}_{\alpha})_{I\times L}$ be two mixed systems of morphisms of $\mathcal{E}$ into $\mathcal{E}'$. If $r^{\lambda}_{\alpha}\colon K^{\lambda}_{\alpha}\to E^{\lambda}$ is the kernel of the pair $(u^{\lambda}_{\alpha}, v^{\lambda}_{\alpha})$ and $\mathcal{K}=(K^{\lambda}_{\alpha}, k^{\mu\lambda}_{\alpha\beta})_{I\in L}$ is the mixed system such $k^{\mu\lambda}_{\alpha\beta}$ is the only morphism of K^{λ}_{β} into K^{μ}_{α} which satisfies*

$$r^{\mu}_{\alpha}k^{\mu\lambda}_{\alpha\beta}=f^{\mu\lambda}_{\alpha\beta}r^{\lambda}_{\beta}, \tag{4.4}$$

then $r=(r^{\lambda}_{\alpha})_{I\times L}$; $\mathcal{K}\to\mathcal{E}$ is the kernel of the pair (u, v).

PROOF. Since the category $\mathcal{C}$ is such that every projective system has a limit, any pair of morphisms of $\mathcal{C}$ has a kernel (cf. [2] p. 54). Let $r^{\lambda}_{\alpha}\colon K^{\lambda}_{\alpha}\to E^{\lambda}_{\alpha}$ be the kernel of the pair $(u^{\lambda}_{\alpha}, v^{\lambda}_{\alpha})$. Moreover, if $\alpha\leqq\beta$ in I and $\lambda\leqq\mu$ in L, we have

$$u^{\mu}_{\alpha}f^{\mu\lambda}_{\alpha\beta}r^{\lambda}_{\beta}=f'^{\mu\lambda}_{\alpha\beta}u^{\lambda}_{\beta}r^{\lambda}_{\beta}=f'^{\mu\lambda}_{\alpha\beta}v^{\lambda}_{\beta}r^{\lambda}_{\beta}=v^{\mu}_{\alpha}f^{\mu\lambda}_{\alpha\beta}r^{\lambda}_{\beta};$$

hence there exists a unique morphism $k^{\mu\lambda}_{\alpha\beta}\colon K^{\lambda}_{\beta}\to K^{\mu}_{\alpha}$ such that (4.4) is satisfied. Clearly, $\mathcal{K}=(K^{\lambda}_{\alpha}, k^{\mu\lambda}_{\alpha\beta})_{I\times L}$ is a mixed system, and the morphism $r=(r^{\lambda}_{\alpha})_{I\times L}$ is such that $ur=vr$. Let now $s\colon \mathcal{L}\to\mathcal{K}$ be another mixed system of morphisms such that $us=vs$ with $\mathcal{L}=(L^{\lambda}_{\alpha}, 1^{\mu\lambda}_{\alpha\beta})_{I\times L}$ and $s=(s^{\lambda}_{\alpha})_{I\times L}$. By definition of kernel, there exists, for each $(\alpha, \lambda)\in I\times L$, a unique morphism $t^{\lambda}_{\alpha}\colon L^{\lambda}_{\alpha}\to K^{\lambda}_{\alpha}$ such that:

$$r^{\lambda}_{\alpha}t^{\lambda}_{\alpha}=s^{\lambda}_{\alpha}. \tag{4.5}$$

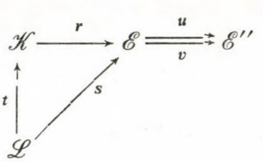


Fig. (4.3)

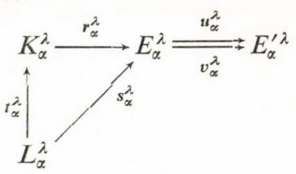


Fig. (4.4)

Again, one can prove easily that $t=(t^{\lambda}_{\alpha})_{I\times L}$ is a mixed system of morphisms of $\mathcal{L}$ into $\mathcal{K}$ and (4.5) shows that $rt=s$. Finally the uniqueness of t follows from that of each t^{λ}_{α}.

Dually,

THEOREM 4.4. *Let $u=(u^{\lambda}_{\alpha})_{I\times L}$ and $v=(v^{\lambda}_{\alpha})_{I\times L}$ be two mixed systems of morphisms of $\mathcal{E}$ into $\mathcal{E}'$. If $r'^{\lambda}_{\alpha}\colon E'^{\lambda}_{\alpha}\to K'^{\lambda}_{\alpha}$ is the cokernel of the pair $(u^{\lambda}_{\alpha}, v^{\lambda}_{\alpha})$ and $\mathcal{K}'=(K'^{\mu}_{\alpha}, k^{\mu\lambda}_{\alpha\beta})_{I\times L}$ is the mixed system such that $k'^{\mu\lambda}_{\alpha\beta}$ is the only morphism of K^{λ}_{β} into K^{μ}_{α} which satisfies $r'^{\mu}_{\alpha}f'^{\mu\lambda}_{\alpha\beta}=k'^{\mu\lambda}_{\alpha\beta}r'^{\lambda}_{\beta}$ then $r'\colon \mathcal{E}'\to\mathcal{K}'$ is the cokernel of the pair (u, v).*

COROLLARY 4.1. *Every functor from a small category into $\mathcal{M}(I\times L, \mathcal{C})$ has projective and inductive limits.*

THEOREM 4.5. *Let $\mathcal{C}$ be a category with zero object (respectively pre-additive, additive, abelian), then so is the category $\mathcal{M}(I\times L, \mathcal{C})$.*

This can be done by direct checking of the axioms.

References

[1] N. Bourbaki, *Eléments de Mathématique,* Hermann (Paris, 1970).
[2] I. Bucur and A. Deleanu, *Introduction to the theory of categories and functors,* John Wiley (1968).
[3] P. Jaffard et G. Poitou, *Introduction aux catégories et aux problèmes universels,* Ediscience (Paris, 1971).
[4] J. P. Lafon, *Les formalismes fondamentaux de l'algèbre commutative,* Hermann (Paris, 1974).
[5] B. Pareigis, *Categories and functors,* Academic Press (1970).

(Received May 5. 1977)

FACULTY OF SCIENCES
UNIVERSITY OF ALEXANDRIA
ALEXANDRIA
EGYPT

FACULTY OF ENGINEERING
UNIVERSITY OF ALEXANDRIA
ALEXANDRIA
EGYPT

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 17—22.

SYMMETRIC MONOTONICITY

By

C. L. BELNA (University Park), M. J. EVANS (Macomb), and P. D. HUMKE (Macomb)

§ 0. Introduction

A real valued function f defined on the real line R is said to be *nondecreasing at* x if there exists a positive number δ_x such that

$$f(x-h) \leqq f(x) \leqq f(x+h) \quad \text{for} \quad 0 < h < \delta_x.$$

The function f is said to be *symmetrically nondecreasing at* x if there exists a positive number δ_x such that

$$f(x-h) \leqq f(x+h) \quad \text{for} \quad 0 < h < \delta_x.$$

We set

$$\mathscr{I} = \{x\colon f \text{ is nondecreasing at } x\}$$

and

$$\mathscr{I}^s = \{x\colon f \text{ is symmetrically nondecreasing at } x\}.$$

The purpose of this paper is to determine the possible size of the set $\mathscr{I}^s - \mathscr{I}$ for various classes of functions. In § 1 we prove that this set is of Lebesgue measure zero for all measurable functions and of the first Baire category for those having the Denjoy property. In § 2 we present several examples to indicate the sharpness of these results.

We shall use the symbol $|E|(|E|^*)$ to denote the Lebesgue measure (outer measure) of a subset E of R.

§ 1. Functions for which $\mathscr{I}^s - \mathscr{I}$ is small

In this section we shall find it convenient to use the decomposition $\mathscr{I}^s = \bigcup_{n=1}^{\infty} S_n$ where $S_n = \{x\colon f(x-h) \leqq f(x+h) \text{ for } 0<h<1/n\}$. We begin with a lemma from which it will readily follow that $\mathscr{I}^s - \mathscr{I}$ is small in the measure-theoretic sense for measurable functions (cf. [3], pp. 217—219).

LEMMA 1. *If $f\colon R \to R$ is measurable and* 0 *is a point of outer density of the set S_n, then there is a positive number δ such that $f(a) \leqq f(b)$ whenever $-\delta<a<0<b<\delta$.*

PROOF. Choose a number $\delta \in (0, 1/n)$ such that, for each closed interval I containing 0 and having length less than δ,

(1) $$|S_n \cap I|^* > 7|I|/8.$$

Now, assume that $f(a)>f(b)$ for some pair of points a and b satisfying $-\delta<a<<0<b<\delta$. Furthermore, assume $|a|\leqq b$. (The proof in the case when $|a|>b$ is obtained by interchanging the roles of a and b in the ensuing arguments.)

It is a consequence of (1) that $|S_n\cap[b/2, 3b/4]|^*>b/8$. Thus, if we set

$$F=\{x\in[0, b/2]:\ (x+b)/2\in S_n\},$$

then $|F|^*>b/4$ and

$$F\subset G\equiv\{x\in[0, b/2]:\ f(x)<f(a)\}.$$

Since f is measurable, it follows that $|G|>b/4$. Now set $P=\{(a+x)/2:\ x\in G\}$, and let I_0 be the smallest closed interval containing 0 and the set P. It is evident that $|P|>b/8$ and $|I_0|\leqq b/2$; hence, $|P\cap I_0|=|P|>|I_0|/4$. But, since $P\cap S_n=\emptyset$, this contradicts (1) and the proof is complete.

THEOREM 1. *If f: $R\to R$ is measurable, then* $|\mathcal{I}^s-\mathcal{I}|=0$.

PROOF. From Lemma 1 we see that if f is approximately continuous at a point x_0 of outer density of S_n, then x_0 is necessarily in $\mathcal{I}$. Then since almost every point of S_n is both a point of approximate continuity of f and a point of outer density of S_n, it follows that $|S_n-\mathcal{I}|=0$ for each index n, and the theorem is proved.

We now proceed to establish the smallness of $\mathcal{I}^s-\mathcal{I}$ in the topological sense for functions possessing the Denjoy property. (Here we say that a function f: $R\to R$ has the Denjoy property if for every pair of open intervals I and J the set $I\cap f^{-1}(J)$ is either empty or of positive Lebesgue measure; we note that the class of functions having the Denjoy property contains the class of approximately continuous functions and the class of Baire*1 Darboux functions, recently introduced by R. J. O'MALLEY [5].) First we prove two lemmas which may have interest independent of the theorem they precede.

LEMMA 2. *Let f: $R\to R$ have the Denjoy property. If S_n is dense in an open interval I, then $I\subset S_n$.*

PROOF. Suppose I is not contained in S_n and for notational simplicity assume $0\in I-S_n$. Since $0\notin S_n$ there is a number h satisfying $0<h<1/n$ such that $f(-h)>>f(h)$. Choose δ such that $0<\delta<\min\{h, 1/n-h\}$ and $(-\delta, \delta)\subset I$. Employing the Denjoy property of f, we can find two sets A and B of positive measure such that $A\subset(-h-\delta, -h+\delta)$, $B\subset(h-\delta, h+\delta)$ and

$$f(a)>f(b) \quad \text{for all} \quad a\in A, \quad b\in B. \tag{2}$$

Let $C=\{(a+b)/2:\ a\in A, b\in B\}$. Since A and B have positive measure, C must contain an interval (see [2, Theorem B, p. 68]). Furthermore $C\subset(-\delta, \delta)\subset I$, and $C\cap S_n=\emptyset$ because of (2). This contradicts the hypothesis that S_n is dense in I, and the proof is complete.

The next lemma generalizes Theorem 1 of [4] where the same result is proved for continuous functions by different methods. We should also observe that based on the following result, it is easy to verify that all fifteen theorems and corollaries proved in [4] for continuous functions remain true for functions possessing the Denjoy property.

LEMMA 3. *If f: $R\to R$ has the Denjoy property and is nowhere monotone, then S_n is nowhere dense and hence $\mathcal{I}^s$ is of the first category.*

PROOF. Let I be any open interval. As f is nowhere monotone, to each positive integer n there correspond two points a_n and b_n in I such that $0<b_n-a_n<2/n$ and $f(a_n)>f(b_n)$. Hence $(a_n+b_n)/2$ is in $I-S_n$, and in view of Lemma 2 this lemma is proved.

THEOREM 2. *If f: $R\to R$ has the Denjoy property, then $\mathcal{I}^s-\mathcal{I}$ is of the first category.*

PROOF. Let I be any open interval. If f is nowhere monotone on I, then S_n is nowhere dense in I by Lemma 3. If f is monotone on an open subinterval J of I, then it is evident that $(S_n-\mathcal{I})\cap J=\emptyset$. Therefore, in either case, the set $S_n-\mathcal{I}$ is not dense in I, and the theorem obtains.

We conclude this section by establishing the following improved version of Lemma 2 for approximately continuous functions; this result will be used to determine a certain property of the function constructed in the third example of the next section.

LEMMA 4. *If f: $R\to R$ is approximately continuous, then S_n is a closed set.*

PROOF. Let x_0 be a limit point of S_n, and without loss of generality assume that $x_0=0$ and that there exists a sequence of positive numbers $x_1, x_2, \dots$ in S_n with $x_k\to 0$. Suppose $0\notin S_n$. Then there is an h_0 such that $0<h_0<1/n$ and $f(-h_0)>f(h_0)$. Set

$$T = \{x\colon f(x) \leqq [f(h_0)+f(-h_0)]/2\}.$$

Then, as f is approximately continuous at h_0, there exists a positive number δ_0 such that $|I_\delta\cap T|>\delta$ for each δ in $(0, \delta_0)$, where $I_\delta=(h_0-\delta, h_0+\delta)$. Set

$$J_\delta^k = \{x\colon (x+y)/2 = x_k \text{ for some } y\in I_\delta\}.$$

Since $x_k\in S_n$ we have

$$|J_\delta^k\cap T| \geqq |I_\delta\cap T| > \delta \quad (0<\delta<\delta_0^*),$$

where $\delta_0^*=\min(\delta_0, h_0, 1/n-h_0)$. Then, since $J_\delta^k\to I_\delta^-\equiv(-h_0-\delta, -h_0+\delta)$ as $k\to\infty$, it follows that $|I_\delta^-\cap T|\geqq\delta$ for each δ in $(0, \delta_0^*)$. Therefore, by the approximate continuity of f at $-h_0$, we have $-h_0\in T$. This is a contradiction, and the lemma is established.

§ 2. Examples

In this section we exhibit functions for which the set $\mathcal{I}^s-\mathcal{I}$ is not "small" in one sense or another. Before presenting these examples we observe that if G is an additive group of real numbers and if f is the characteristic function of G, then $G\subset\mathcal{I}^s$. Furthermore, if both G and its complement are dense in R, then $\mathcal{I}$ is empty. These observations are used in the first two examples given.

We begin by proving that in general the set $\mathcal{I}^s-\mathcal{I}$ need not be small in either the topological or the measure-theoretic sense, that is, neither Theorem 1 nor Theorem 2 is true for arbitrary functions.

EXAMPLE 1. *There is a function f such that $\mathcal{I}^s-\mathcal{I}$ is both nonmeasurable and of the second category in every interval.*

PROOF. We use transfinite induction to define an additive group G. Let $P_0, P_1, \dots, P_\alpha, \dots$ $(\alpha<\omega_c)$ be a transfinite sequence of all nowhere dense perfect subsets of R, where ω_c is the initial ordinal having the power $\mathfrak{c}$ of the continuum. For each subset A of R we use $\langle A\rangle$ to denote the smallest group containing all rational multiples of the elements of A.

Let x_0 and y_0 be elements of P_0 such that $\langle\{x_0\}\rangle\cap\langle\{y_0\}\rangle=\{0\}$. Now suppose that $\alpha<\omega_c$ and that for each $\beta<\alpha$ there exist elements x_β and y_β in P_β such that

$$\langle\{x_\sigma\colon \sigma\leqq\beta\}\rangle\cap\langle\{y_\sigma\colon \sigma\leqq\beta\}\rangle=\{0\}.$$

It then follows that

$$\langle\{x_\sigma\colon \sigma<\alpha\}\rangle\cap\langle\{y_\sigma\colon \sigma<\alpha\}\rangle=\{0\}.$$

Now $\operatorname{card}(\alpha)<\mathfrak{c}$ implies that $\langle\{x_\sigma\colon \sigma<\alpha\}\cup\{y_\sigma\colon \sigma<\alpha\}\rangle$ has cardinality less than $\mathfrak{c}$. Hence, as $\operatorname{card}(P_\alpha)=\mathfrak{c}$, there is a point $x_\alpha\in P_\alpha$ such that

$$x_\alpha\notin\langle\{x_\sigma\colon \sigma<\alpha\}\cup\{y_\sigma\colon \sigma<\alpha\}\rangle.$$

Similarly, the cardinality of $\langle\{x_\sigma\colon \sigma\leqq\alpha\}\cup\{y_\sigma\colon \sigma<\alpha\}\rangle$ is less than $\mathfrak{c}$ and there is a point $y_\alpha\in P_\alpha$ such that

$$y_\alpha\notin\langle\{x_\sigma\colon \sigma\leqq\alpha\}\cup\{y_\sigma\colon \sigma<\alpha\}\rangle.$$

It is easily verified that

$$\langle\{x_\sigma\colon \sigma\leqq\alpha\}\rangle\cap\langle\{y_\sigma\colon \sigma\leqq\alpha\}\rangle=\{0\},$$

and this completes the induction.

Now set $G=\langle\{x_\alpha\colon \alpha<\omega_c\}\rangle$ and $G'=\langle\{y_\alpha\colon \alpha<\omega_c\}\rangle$. Then $G\cap G'=\{0\}$ and both G and G' are additive groups of real numbers each of which contains at least one point from every nowhere dense perfect set. As such, both G and G' are nonmeasurable and of the second category in every interval. So if f is the characteristic function of G, then the example follows from the introductory remarks of this section.

In [1] it is shown that $\mathcal{I}$ is of type $G_{\delta\sigma}$ for an arbitrary function. Example 1 shows that $\mathcal{I}^s$ need not be measurable.

Example 1 can be used to make another interesting observation. Let f be a function from R to R and consider

$$\bar{f}^s(x)=\limsup_{h\to0}\frac{f(x+h)-f(x-h)}{2h}.$$

In 1927, A. KHINTCHINE [3] proved the following result.

THEOREM K. *If $f\colon R\to R$ is measurable, then f has a finite derivative $f'(x)$ at almost every point x where $\bar{f}^s(x)<\infty$.*

Letting f be the characteristic function of the group G constructed in Example 1, we see that Theorem K cannot be extended to arbitrary functions since $\bar{f}^s(x)=0$ for each $x\in G$ and yet $f'(x)$ exists for no $x\in G$.

The following apparently unpublished theorem of Paul Erdős (Budapest) was communicated to us by Jan Mycielski (Boulder) via James Foran (Kansas City). In particular, this theorem entails that the continuum hypothesis implies that the hypothesis on the function in Theorem 2 cannot be weakened to measurability.

THEOREM E. *The continuum hypothesis implies the existence of both* (1) *a group of real numbers that is of measure zero and of the second category, and* (2) *a group of real numbers that is of the first category and not of measure zero.*

PROOF. The proofs for (1) and (2) are essentially identical. We shall give the proof of (1) with the necessary changes for the proof of (2) indicated parenthetically. Let $A_0, A_1, \dots, A_\alpha, \dots$ $(\alpha<\omega_c)$ be a transfinite sequence of all subsets of R which are F_σ and of the first category [G_δ and of measure zero]. We construct a sequence of groups $G_0, G_1, \dots, G_\alpha, \dots$ $(\alpha<\omega_{\mathfrak{c}})$ such that each G_α satisfies the following properties:

(i) G_α is countable,

(ii) $G_\beta \subset G_\alpha, \quad G_\beta \neq G_\alpha$ for $\beta<\alpha$,

(iii) $G_\alpha \cap A_\beta = G_\beta \cap A_\beta$ for $\beta \leqq \alpha$.

If this is done, then the group $G = \bigcup_{\alpha<\omega_c} G_\alpha$ has the required properties; it is of measure zero because G has a countable intersection with any first category F_σ subset of R whose complement is of measure zero; it is not of the first category because G is not contained in any of the A_α $(0\leqq\alpha<\omega_c)$ [it is of the first category because G has a countable intersection with each residual G_δ subset of measure zero of R; it is not of measure zero because G is not contained in any of the $A_\alpha (0\leqq\alpha<\omega_c)$].

We construct $G_\alpha (0\leqq\alpha<\omega_c)$ by induction. Let G_0 be the group of rational numbers and suppose that for some $\hat\alpha<\omega_c$ the groups $G_\alpha (0\leqq\alpha<\hat\alpha)$ have been defined and satisfy the properties (i)—(iii).

Let $G^*_{\hat\alpha} = \bigcup_{\alpha<\hat\alpha} G_\alpha$ and choose $\hat x$ not in the union

$$(*) \qquad \bigcup_{\alpha<\hat\alpha} \{x\colon nx+g\in A_\alpha \text{ for some rational number } n\neq 0 \text{ and } g\in G^*_{\hat\alpha}\}.$$

Such an $\hat x$ exists since, under the assumption of the continuum hypothesis, the union in $(*)$ is a countable union of sets of the first category [measure zero].

Now define $G_{\hat\alpha} = \langle G^*_{\hat\alpha} \cup \{\hat x\}\rangle$, where the notation is the same as in Example 1. Each of the inductive properties (i)—(iii) is easily seen to be satisfied by $G_{\hat\alpha}$, and consequently the theorem is proved.

If f is the characteristic function of the group G constructed above to establish (1) in Theorem E, then f is measurable and $\mathcal{I}^s-\mathcal{I}$ is of the second category. Consequently, we have (or rather Erdős has) established the following example.

EXAMPLE 2. *There is a measurable function* $f\colon R\to R$ *for which* $\mathcal{I}^s-\mathcal{I}$ *is of the second category.*

From Theorems 1 and 2 it follows that if $f\colon R\to R$ has the Denjoy property then the set $\mathcal{I}^s-\mathcal{I}$ is of measure zero and of the first category; nevertheless, this set can be uncountable even if f is continuous as our closing example shows.

EXAMPLE 3. *There is a continuous function $f\colon R\to R$ for which $\mathcal{I}^s-\mathcal{I}$ is uncountable.*

PROOF. Let T denote the Cantor middle (3/5)ths set in [0, 1], and let $\{(a_n, b_n)\}_{n=1}^{\infty}$ be the open intervals in [0, 1] contiguous to T. Let $g\colon R\to R$ be the function given by

$$g(x)=\begin{cases} x, & x\in(-\infty, 1/6] \\ -x+1/3, & x\in(1/6, 1/3] \\ 0, & x\in(1/3, 2/3] \\ -x+2/3, & x\in(2/3, 5/6] \\ x-1, & x\in(5/6, \infty) \end{cases}$$

and then define the function $f\colon R\to R$ by

$$f(x)=\begin{cases} 0, & x\in T \\ (b_n-a_n)\,g\left(\dfrac{x-a_n}{b_n-a_n}\right), & x\in(a_n, b_n) \\ g(x), & x\in R-[0, 1]. \end{cases}$$

It is readily observed that f is continuous and satisfies the following properties; (i) $|f(x)|\leqq$distance from x to T, and (ii) for each $x\in T$, f assumes both positive and negative values in each open interval abutting x on the right or in each open interval abutting x on the left.

By (ii) and the fact that $f(x)=0$ for each $x\in T$, it follows that $\mathcal{I}\cap T=\emptyset$. However, $T\subset\mathcal{I}^s$; indeed, $T\subset S_1$. To see this it suffices to note that S_1 is closed by Lemma 4, that the set of endpoints of the intervals contiguous to T is dense in T, and that every such endpoint belongs to S_1. The last observation follows readily from (i). Then since T is uncountable, the example is established.

References

[1] R. O. DAVIES, L. EIFLER, and F. GALVIN (all independently), *Solution to Query* 9, Real Analysis Exchange **1** (1976), 78.

[2] P. R. HALMOS, *Measure Theory,* Van Nostrand (New York, 1950).

[3] A. KHINTCHINE, Recherches sur la structure des fonctions mesurables, *Fund. Math.,* **9** (1927), 212—279.

[4] N. K. KUNDU and S. N. MUKHOPADHYAY, Symmetric derivative of nowhere monotone functions, I, *Ann. Pol. Math.,* **23** (1970), 247—253.

[5] R. J. O'MALLEY, Baire*1, Darboux functions, *Proc. Amer. Math. Soc.,* **60** (1976), 187—192.

(Received May 27, 1977)

DEPARTMENT OF MATHEMATICS
PENNSYLVANIA STATE UNIVERSITY
UNIVERSITY PARK, PENNSYLVANIA 16802
U.S.A.

DEPARTMENT OF MATHEMATICS
WESTERN ILLINOIS UNIVERSITY
MACOMB, ILLINOIS 61455
U.S.A.

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 23—25.

ON A COMMUTIVITY THEOREM OF LUH

By
A. RICHOUX (Raleigh)

In his paper, J. LUH [1] showed that a primary ring with unity element 1 which satisfies the identities

$$(1)\qquad \begin{cases} (xy)^k = x^k y^k, \quad k = n, n+1, n+2 \quad \text{for all} \quad x, y \in R, \\ \text{where } n \text{ is fixed integer} \geqq 1, \end{cases}$$

is commutative. This result has subsequently been generalized by A. KAYA [2] and by S. LIGH and A. RICHOUX [3]. Ligh and Richoux showed that any ring with unity element 1 which satisfies the identities (1) is commutative. A. Kaya proved that a ring satisfying the identities

$$(2)\qquad \begin{cases} (xy)^k = x^k y^k, \; k = n(x, y), \; n(x, y), \; n(x, y)+1, \; n(x, y)+2 \\ \text{for all } x, y \in R, \text{ where } n(x, y) \text{ is an integer} \geqq 1 \\ \text{which depends on } x \text{ and } y, \end{cases}$$

is commutative if it is a primary ring with 1 or if it is a semi-prime ring with 1. In this paper we will show that *a ring which satisfies the identities* (2) *is commutative if it has* 1 *or if it has no nilpotents.* In the following R is understood to be any ring which satisfies (2).

We begin with a simple observation. Let $x, y \in R$ and suppose $x^i y^j = 0$ for some integers $i, j \geqq 1$. Then for $n_1 = n(x, y)+1 > 1$, $(xy)^{n_1} = x^{n_1} y^{n_1}$. If $\max\{i, j\} > n_1$, then for $n_2 = n(x^{n_1}, y^{n_1})+1 > 1$, we have $(xy)^{n_1 n_2} = (x^{n_1} y^{n_1})^{n_2} = x^{n_1 n_2} y^{n_1 n_2}$ with $n_1 n_2 > n_1$. Proceeding in this way we eventually reach the point where

$$(xy)^{n_1 n_2 \dots n_r} = x^{n_1 n_2 \dots n_r} y^{n_1 n_2 \dots n_r}$$

with $n_1 n_2 \dots n_r \geqq \max\{i, j\}$. Thus xy is nilpotent. But so is yx.

Now the identities

$$(3)\qquad x^k (y^k x - x y^k) y = 0, \quad k = n(x, y), \; n(x, y)+1$$

follow immediately from the identities (2) and will be used below.

THEOREM 1. *If R has no nilpotents, then R is commutative.*

PROOF. From the identities (3) and the observations noted above, for all $x, y \in R$, $(y^i x - x y^i) yx$ is nilpotent for $i = n(x, y)$, $n(x, y)+1$. Thus the identities

$$(y^i x - x y^i) yx = 0, \quad i = n(x, y), \; n(x, y)+1$$

hold for all $x, y \in R$. So for $i=n(x, y)$,

$$0 = (y^{i+1}x - xy^{i+1})yx - y[(y^i x - xy^i)yx] = (yxy^i - xy^{i+1})yx = (yx-xy)y^{i+1}x$$

holds for all $x, y \in R$. But then $x(yx-xy)y^{n(x,y)+1}$ is nilpotent and so $x(yx-xy)y^{n(x,y)+1}=0$ for all $x, y \in R$. Then again by our observation above, $y[x(yx-xy)]=0$ holds for all $x, y \in R$. By substituting x for y and y for x into this last identity, we get $xy(xy-yx)=0$ for all $x, y \in R$. Adding these last two identities, we reach $(yx-xy)^2=0$ for all $x, y \in R$ and so $xy=yx$ for all $x, y \in R$.

Lemma 1. *If R has unity element* 1, *nilpotents are in the centre Z of R.*

Proof. Let $x \in R$ be nilpotent. Note that for any $r \in R$, $r(x+1)-(x+1)r = = rx - xr$. So by (3), the identites

$$(x+1)^i(y^i x - xy^i)y = 0, \quad i = n(x+1, y),\ n(x+1, y)+1$$

hold for all $y \in R$. Then since $1+x$ is a unit in R, the identities

$$(y^i x - xy^i)y = 0, \quad i = n(x+1, y)\ n(x+1, y)+1$$

hold for all $y \in R$. But then for $i=n(x+1, y)$,

$$0 = (y^{i+1}x - xy^{i+1})y - y[(y^i x - xy^i)y] = (yxy^i - xy^{i+1})y = (yx-xy)y^{i+1}$$

holds for all $y \in R$.

Now let $y \in R$. Out of all integers $i \geqq 0$ such that $(yx-xy)y^i=0$, pick i_0 which is minimal with respect to this property. Suppose $i_0>0$. Then there exists an integer $j>0$ such that

$$0 = [(y+1)x - x(y+1)](y+1)^j = (yx-xy)(y+1)^j.$$

But

$$0=(yx-xy)(y+1)^j y^{i_0-1} = (yx-xy)\sum_{k=0}^{j}\binom{j}{k} y^{k+i_0-1} = (yx-xy)y^{i_0-1}$$

which contradicts the minimality of i_0. Thus $i_0=0$ and $yx=xy$.

Theorem 2. *If R has* 1, *R is commutative.*

Proof. Let N be the collection of nilpotents in R. By Lemma 1, N is an ideal. By Theorem 1, R/N is commutative. Thus for all $x, y \in R$, $xy-yx \in N \subseteqq Z$. Then by (3), the identities

$$0 = x^i[(y^i x - xy^i)y] = [(y^i x - xy^i)y]x^i, \quad i = n(x, y),\ n(x, y)+1$$

hold for all $x, y \in R$. So for $j=n(x, y)$, the identity

$$0 = (y^{j+1}x - xy^{j+1})yx^{j+1} - y[(y^j x - xy^j)yx^j]x = (yxy - xy^{j+1})yx^{j+1} =$$
$$= (yx-xy)y^{j+1}x^{j+1} = x^{j+1}(yx-xy)y^{j+1}$$

holds for all $x, y \in R$.

Now let $x, y \in R$. Of all such pairs $i,\ j \geqq 0$ of integers such that $x^i(yx-xy)y^j=0$, pick a pair i_0, j_0 such that i_0 is minimal. Suppose $i_0>0$. Then there exists an integer $k \geqq 0$ such that

$$0=(x+1)^k[y(x+1)-(x+1)y]y^k=(x+1)^k(yx-xy)y^k.$$

Then for $1=\max\{j_0, k\}$,

$$0=x^{i_0-1}(x+1)^k(yx-xy)y^1=\sum_{n=k}^{k}\binom{k}{n}x^{i_0-1+k}(yx-xy)y^1=x^{i_0-1}(yx-xy)y^1$$

which contradicts our choice of the pair i_0, j_0. Thus $i_0=0$ and $(yx-xy)y^{j_0}=0$. We now have for all $x, y \in R$, there exists an integer $i=i(x, y) \geqq 0$ such that $(xy-yx)y^i=0$. Let $x, y \in R$. Pick the integer $i_0 \geqq 0$ which is minimal with respect to the property $(yx-xy)y^{i_0}=0$. If $i_0>0$, then in a manner similar to the above, we reach a contradiction. Thus $xy=yx$.

References

[1] J. LUH, A commutivity theorem for primary rings, *Acta Math. Acad. Sci. Hungar.*, **22** (1971), 211—213.
[2] A. KAYA, On a commutivity theorem of Luh, *Acta Math. Acad. Sci. Hungar.*, **28** (1976), 33—36.
[3] S. LIGH and A. RICHOUX, A commutivity theorem for rings, *Bulletin of the Australian Math. Society* (to be published).

(Received May 27, 1977)

NORTH CAROLINA STATE UNIVERSITY
DEPARTMENT OF MATHEMATICS
RALEIGH, NORTH CAROLINA
USA

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 27—32.

EXPLICIT BOUNDS FOR THE DEPARTURE FROM NORMALITY OF SUMS OF DEPENDENT RANDOM VARIABLES

By

R. V. ERICKSON (Lansing), M. P. QUINE (Sydney) and N. C. WEBER (Sydney)

1. Introduction

In this paper, the substitution methods of DVORETZKY [4] are used to bound the departure from normality of a large class of processes which are martingales, or at least close to being martingales. Two main results are given. Theorem 2.1 generalizes Theorem 1 of BASU [1] in the following ways: the function f is not assumed to have a third derivative, the conditional expectations are not assumed to vanish, and the sum of the conditional variances is not assumed to be almost surely unity. As a corollary we are able to give a uniform bound on $\Delta(S_n, N)$ which goes to zero for any array satisfying the conditions of Dvoretzky's main theorem ([4], Theorem 2.2). However, our result cannot give a rate of convergence for all systems satisfying the conditions of Corollary (2.7) of MCLEISH [7]. On the other hand McLeish's result does not contain that of Dvoretzky as he claims. For if we take X_{ni}, $i=1, \dots, n$ independent with

$$X_{ni} = \begin{cases} 6(-1)^{i-1}/(29n)^{1/2} & \text{prob} \frac{1}{2} \\ 0 & \text{prob} \frac{1}{6} \\ 6(-1)^{i}/(29n)^{1/2} & \text{prob} \frac{1}{3}, \end{cases}$$

then McLeish's conditions (2.7 c, d) fail but the array satisfies Dvoretzky's conditions.

Theorem (2.3) may be compared with other results in the literature. HEYDE and BROWN [5] use the Skorokhod imbedding principle to obtain a bound of the same form as ours. However our method is comparatively direct, and yields a numerical bound for the constant; less importantly we remove their requirement that the row sum variance equal 1. BASU [1, Theorem 2] deals with the special case where the row sum of the conditional variances is equal to 1. In this case our Theorem 2.3 gives a better rate, and does not involve a conditional variance term. In the case of independent, identically distributed random variables, with third moments, our bound becomes $O(n^{-1/8})$ whereas that of Basu is $O(n^{-1/12})$; his claim of $O(n^{-1/4})$ when $\delta=1$ and the variables are independent and uniformly bounded does not seem to follow from his Theorem 2. Finally we note that our methods of proof could also be used to give a similar bound to that in Theorem (2.3) when the conditional expectations of the variables do not vanish.

2. Notation and results

Let $X=\{X_{nk},\ k=1,\ldots,n;\ n=1,2,\ldots\}$ be an array of random variables. Put $S_{nk}=\sum_{j=1}^{k}X_{nj}$, $S_n=S_{nn}$. Let $F=\{F_{nk},\ k=0,\ldots,n;\ n=1,2,\ldots\}$ be an array of σ-fields. Assume X_{nk} is F_{nk} measurable and $F_{nk}\subset F_{n,k+1}$. Given (X,F), define the conditional expectation operators $E_{nk}(\cdot)=E(\cdot|F_{nk})$ and write

$$\mu_{nk}=E_{n,k-1}X_{nk},\ X'_{n,k}=X_{nk}-\mu_{nk},\quad S'_n=\sum_{j=1}^{n}X'_{nj},$$

$$\sigma_{nk}^2=E_{n,k-1}X'^2_{nk},\ \mu_n=\sum_{j=1}^{n}\mu_{nj},\quad \sigma_n^2=\sum_{j=1}^{n}\sigma_{nj}^2.$$

Let C^* be the class of real valued functions f with second derivatives everywhere defined which, for some $f^{(3)}$, satisfy, for all $y<x$

$$f''(x)-f''(y)=\int_y^x f^{(3)}(t)\,dt,$$

and let $K_f=\|f''\|_\infty\vee\frac{1}{6}\|f^{(3)}\|_\infty$ $(a\vee b=\max(a,b))$.

Define the stopping time $\tau_n=\max\left(k:\ \sum_{j=1}^{k}\sigma_{nj}^2\leqq 1\right)$ and the random variables

$$U_{nk}=X'_{nk}I(|X'_{nk}|\leqq 1),\ V_{nk}=X'_{nk}-U_{nk},\ \alpha_n^2=\sum_{j=1}^{\tau_n}\sigma_{nj}^2\ \text{ and }\ L_n=\sum_{k=1}^{\tau_n}E_{n,k-1}(|U_{nk}|^3+V_{nk}^2).$$

If the random variables X, Y have distribution functions F, G define $\Delta(X,Y)=\sup_x|F(x)-G(x)|$. Let N be $N(0,1)$.

(2.1) THEOREM. *Given (X,F) and $f\in C^*$ then if $a>1$ and $b>0$,*

$$|Ef(S_n)-Ef(N)|\leqq K_f\{(1+4\pi^{-1/2})EL_n+(1.5+4\pi^{-1/2})E(1-\alpha_n^2)+$$
$$+\frac{1}{2}E[(\sigma_n^2-\alpha_n^2)\wedge a]\}+\|f'\|_\infty E|\mu_n|I(|\mu_n|\leqq b)+2\|f\|_\infty[P(|\mu_n|>b)+P(\sigma_n^2>a)].$$

(2.2) COROLLARY. *Given (X,F), then for any $a>1$, $b>0$*

$$\Delta(S_n,N)\leqq 2\cdot 6([EL_n+E(1-\alpha_n^2)]^{1/4}+E^{1/4}[(\sigma_n^2-\alpha_n^2)\wedge a])+2P(|\mu_n|>b)+$$
$$+2P(\sigma_n^2>a)+1\cdot 8E^{1/2}|\mu_n|I(|\mu_n|\leqq b).$$

(2.3) THEOREM. *Given (X,F) above, if $\mu_{nk}=0$ for all n,k, then for $\delta\in(0,1]$,*

$$\Delta(S_n,N)\leqq\gamma(\delta)\left\{\sum_{k=1}^{n}E|X_{nk}|^{2+\delta}+E|1-\sigma_n^2|^{1+\delta/2}\right\}^{1/(3+\delta)}$$

where $\gamma(\delta)<21\cdot 6$.

3. Proofs of results

(3.1) LEMMA. *Given (X, F) above, suppose $\mu_{nk}=0$ and $\sigma_n=1$ for all n, k. Then for any $f\in C^*$,*

$$|Ef(S_n)-Ef(N)| \leqq K_f(1+4\pi^{-1/2})EL_n.$$

PROOF. We omit the proof, which is a straightforward extension of the methods of DVORETZKY [4] (c.f. the proof of Theorem 1 of BASU [1]), using the fact that

$$|f(x+t)-f(x)-tf'(x)-{}^1/_2t^2f''(x)| \leqq K_f(t^2/|t^3|).$$

PROOF OF THEOREM (2.1). Clearly,

$$|Ef(S_n)-Ef(S_n')| \leqq 2\|f\|_\infty P(|\mu_n|>b)+\|f'\|_\infty E|\mu_n|I(|\mu_n|\leqq b).$$

Define $W_{nk}=X_{nk}'I(k\leqq\tau_n)$, $k=1, \ldots, n$, $W_{n,n+1}=N_n(1-\alpha_n^2)^{1/2}$, N_n being unit normal variables independent of all other random variables. Then $\{W_{n1}, \ldots, W_{n,n+1}; F_{n0}, \ldots, F_{nn}, F_{nn}\times\sigma(N_n)\}$ satisfies the conditions of Lemma (3.1) and so

$$\left|Ef\left(\sum_{k=1}^{n+1} W_{nk}\right)-Ef(N)\right| \leqq K_f(1+4\pi^{-1/2})\big(EL_n+E(1-\alpha_n^2)\big).$$

Now introduce a new stopping time $\nu(a)=\max\left(k\colon \sum_{j=1}^{k}\sigma_{nj}^2\leqq a\right)$, $a>1$. Clearly,

$$\left|Ef(S_n')-Ef\left(\sum_{k=1}^{\nu(a)} X_{nk}'\right)\right| \leqq 2\|f\|_\infty P(\sigma_n^2>a).$$

From second order Taylor expansions,

$$\left|Ef\left(\sum_{k=1}^{\nu(a)} X_{nk}'\right)-Ef\left(\sum_{k=1}^{n} W_{nk}\right)\right| \leqq \frac{1}{2}\|f''\|_\infty E\left(\sum_{k=\tau_n+1}^{\nu(a)} X_{nk}'\right)^2 \leqq$$

$$\leqq \frac{1}{2}\|f''\|_\infty E[(\sigma_n^2-\alpha_n^2)\wedge a],$$

and

$$\left|Ef\left(\sum_{k=1}^{n} W_{nk}\right)-Ef\left(\sum_{k=1}^{n+1} W_{nk}\right)\right| \leqq \frac{1}{2}\|f''\|_\infty E(1-\alpha_n^2).$$

The result now follows.

PROOF OF COROLLARY (2.2). Consider $f(x)=g\big(d^{-1}(x-b+d)\big)$, $d>0$, b real, where

$$g(x)=\begin{cases}0, & x<0\\ x+\dfrac{1}{2\pi}\sin\pi(2x-1), & x\in[0,1]\\ 1, & x>1.\end{cases}$$

Because of the "almost indicator" nature of f,

$$\Delta(S_n, N) \leqq |Ef(S_n)-Ef(N)|+(2\pi)^{-1/2}d$$

and so by using the bound of Theorem (2.1), tedious calculations to minimise this bound with respect to d yield the required result.

PROOF OF THEOREM (2.3). Put $f(x)=e^{izx}(z, x$ real). As noted in DOOB [3, p.38]

$$f(x+t)=f(x)\left(1+izt-\frac{1}{2}z^2t^2+\theta|zt|^{2+\delta}\right),$$

where

$$|\theta|<\frac{2^{1-\delta}}{(2+\delta)(1+\delta)}.$$

Using the methods of DVORETZKY [4] again, it is clear that

$$|Ee^{izS_n}-Ee^{izN}| \leqq |\theta| \sum_{k=1}^{n} \{E|zX_{nk}|^{2+\delta}+E|z\sigma_{nk}Y_{nk}|^{2+\delta}\} \leqq$$

$$\leqq |\theta||z|^{2+\delta}\left(1+2^{1+1/2\delta}\Gamma\left(\frac{3+\delta}{2}\right)\Big/\pi^{1/2}\right)\sum_{k=1}^{n}E|X_{nk}|^{2+\delta}$$

from Jensen's inequality. Thus

$$|Ee^{izS_n}-Ee^{izN}| \leqq |\theta||z|^{2+\delta}u(\delta)L_{n\delta},$$

where $L_{n\delta}=\sum_{k=1}^{n}E|X_{nk}|^{2+\delta}$. Using Esseen's theorem (see e.g. LOÉVE [6], p.285),

$$\Delta(S_n, N) \leqq 2\pi^{-1}|\theta|u(\delta)L_{n\delta}\int_0^u z^{1+\delta}\,dz+24/(\pi^{3/2}2^{1/2}u)$$

whence by choosing u to minimise the bound

(3.2) $$\Delta(S_n, N) \leqq c(\delta)L_{n\delta}^{1/(3+\delta)}$$

where $c(\delta)<4\cdot 7$ and $c(1)\leqq 2\cdot 23$. When $P(\sigma_n^2=1)<1$, we again use the martingale difference sequence $(W_{n1}, \dots, W_{n,n+1})$, for which we know, from (3.2)

(3.3) $$\Delta\left(\sum_{k=1}^{n+1}W_{nk}, N\right) \leqq c(\delta)[L_{n\delta}+E(1-\alpha_n^2)^{1+1/2\delta}]^{1/(3+\delta)}.$$

At this stage it is convenient to introduce the Lévy distance between random variables X and Y defined in terms of their distribution functions F, G as

$$\Lambda(X, Y)=\inf\{h:\ F(x-h)-h \leqq G(x) \leqq F(x+h)+h,\ \forall x\}$$

It is easily show, that

(3.4) $$\Lambda(X, Y) \leqq \varepsilon \vee P(|X-Y|>\varepsilon), \quad \text{for all} \quad \varepsilon>0,$$

(3.5) $$\Lambda(X, N) \leqq \varepsilon \Rightarrow \Delta(X, N) \leqq \varepsilon\left(1+(2\pi)^{-1/2}\right).$$

Using (3.4) and Chebyshev,

$$(3.6)\qquad \Lambda\left(\sum_{k=1}^{n+1} W_{nk},\ \sum_{k=1}^{n} W_{nk}\right) \leqq \varepsilon \vee E|W_{n,n+1}|^{2+\delta}\varepsilon^{-2-\delta} \leqq$$

$$\leqq \left[2^{1+1/2\delta}\pi^{-1/2}\Gamma\left(\frac{3+\delta}{2}\right)E(1-\alpha_n^2)^{1+1/2\delta}\right]^{1/(3+\delta)}.$$

Now note that $\{X_{nk}I(\tau_n<k), k=1, \ldots, n\}$ is a martingale difference sequence with respect to $\{F_{n0}, \ldots, F_{nn}\}$, and that

$$\sum_{k=1}^{n} X_{nk} I(\tau_n < k) = \sum_{k=\tau_n+1}^{n} X_{nk}.$$

Using the inequality (21.5) of BURKHOLDER [2], with $\Phi(\lambda)=\lambda^{2+\delta}$ (whence his $c\leqq 2\cdot 10^{2+\delta}$),

$$E\left|S_n-\sum_{k=1}^{n} W_{nk}\right|^{2+\delta} = E\left|\sum_{k=\tau_n+1}^{n} X_{nk}\right|^{2+\delta} \leqq$$

$$\leqq 2\cdot 10^{2+\delta}\left[E\left(\sum_{k=\tau_n+1}^{n}\sigma_{nk}^2\right)^{1+1/2\delta} + E\sum_{k=\tau_n+1}^{n}|X_{nk}|^{2+\delta}\right].$$

Using (3.4) and Chebyshev again,

$$(3.7)\quad \Lambda\left(\sum_{k=1}^{n} W_{nk},\ S_n\right) \leqq \left\{2\cdot 10^{2+\delta}\left[E\left(\sum_{k=\tau_n+1}^{n}\sigma_{nk}^2\right)^{1+1/2\delta} + E\sum_{k=\tau_n+1}^{n}|X_{nk}|^{2+\delta}\right]\right\}^{1/(3+\delta)}.$$

Now

(3.8)

$$E(1-\alpha_n^2)^{1+1/2\delta} = E(1-\sigma_n^2)_+^{1+1/2\delta} + E(1-\alpha_n^2)^{1+1/2\delta}I(\tau_n<n) \leqq E|1-\sigma_n^2|^{1+1/2\delta}+L_{n\delta}$$

and

$$(3.9)\qquad E(\sigma_n^2-\alpha_n^2)^{1+1/2\delta} \leqq 2^{1/2\delta}\left(E(\sigma_n^2-1)_+^{1+1/2\delta}+E(1-\alpha_n^2)^{1+1/2\delta}I(\tau_n<n)\right) \leqq$$

$$\leqq 2^{1/2\delta}(E|1-\sigma_n^2|^{1+1/2\delta}+L_{n\delta}).$$

Therefore combining (3.6) to (3.9) we get

$$(3.10)\qquad \Lambda\left(\sum_{1}^{n+1} W_{nk},\ S_n\right) \leqq v(\delta)[L_{n\delta}+E|1-\sigma_n^2|^{1+1/2\delta}]^{1/(3+\delta)}$$

where

$$v(\delta) = \left[2^{1+1/2\delta}\pi^{-1/2}\Gamma\left(\frac{3+\delta}{2}\right)\right]^{1/(3+\delta)} + [2\cdot 10^{2+\delta}(1+2^{1/2\delta})]^{1/(3+\delta)}.$$

Using (3.5), (3.3) and (3.10) we see

$$\Delta(S_n, N) \leqq \gamma(\delta)[L_{n\delta}+E|1-\sigma_n^2|^{1+1/2\delta}]^{1/(3+\delta)}$$

where

$$\gamma(\delta) = \left(1+\frac{1}{\sqrt{2\pi}}\right)\left(c(\delta)+v(\delta)\right) < 21.6.$$

References

[1] A. K. Basu, On the rate of convergence to normality for sums of dependent random variables, *Acta Math. Acad. Sci. Hungar.*, **28** (1976), 261—265.

[2] D. L. Burkholder, Distribution function inequalities for martingales, *Ann. Prob.*, **1** (1973), 19—42.

[3] J. L. Doob, *Stochastic Processes*, Wiley (New York, 1953).

[4] A. Dvoretzky, Asymptotic normality for sums of dependent random variables, *Proc. Sixth Berkeley Symp. Math. Statist. Prob.*, **2** (1972), 513—535.

[5] C. C. Heyde and B. M. Brown, On the departure from normality of a certain class of martingales, *Ann. Math. Statist.*, **41** (1970), 2161—2165.

[6] M. Loéve, *Probability Theory*, (2nd Ed.), Van Nostrand (Princeton, 1960).

[7] D. L. McLeish, Dependent central limit theorems and invariance principles, *Ann. Prob.*, **2** (1974), 620—628.

(Received June 6, 1977)

DEPARTMENT OF STATISTICS AND PROBABILITY,
MICHIGAN STATE UNIVERSITY,
EAST LANSING,
MICHIGAN, U.S.A. 48824

AND

DEPARTMENT OF MATHEMATICAL STATISTICS,
UNIVERSITY OF SYDNEY,
N.S.W. 2006, AUSTRALIA.

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 33—41.

INDEPENDENTLY SCATTERED MEASURES[1]

By
S. KHALILI (Cleveland)

1. Introduction

There are numerous practical problems of a probabilistic nature where the classical notion of a stochastic process is not particularly germane. For instance, in the problem of investigating the quantity of rain or crop in different regions, the relevant and useful mathematical concept is the set function which assigns a random variable to each region, briefly a *random measure*[2]. Similar situations arise in the study of spatial distribution of stars or galaxies in space, of bacteria on a slide, of population in a city, etc. Besides their applicability to such physical situations, random measures also perform an important role in the theory of ordinary stochastic processes. Associated with many stochastic processes, there are random measures. For instance, a continuous time strictly or wide sense stationary, purely non-determinisitic stochastic process has an associated "innovation" random measure, and this fact is crucial in the theory. In such situations the process is expressible as a stochastic integral of the random measure. Random measures also have the merit of generalizing more readily than random processes. For instance, when the parameter domain is a general measure space $(\Lambda, \mathscr{A}, m)$ it is much easier to define a Brownian motion random measure over Λ than it is to define a Brownian motion stochastic process over Λ [cf. 9, 10].

In many practical problems, the assumption of independence of the random variables associated with disjoint sets is also quite reasonable. For example, crop yields of different areas greatly separated geographically, or the number of people with certain IQ residing in different sections of a country can be assumed to be independent random variables. For all practical purposes, finiteness of variance of the random variables is also a very reasonable assumption. Thus L_2-valued independently scattered measures are latent in many problems of a stochastic nature.

The theory of independently scattered measures has been studied by CRAMER [1], KINGMAN [8], PRÉKOPA [14, 15, 16], URBANIK [17] and others. The corresponding wide sense concept is that of an orthogonally scattered measure, the theory of which has been presented in detail by MASANI [12]. This paper deals with measures that are independently and orthogonally scattered.

In section 2 we present the basic definitions and elementary properties of L_2-valued independently scattered measures. Section 3 deals with the fundamental question of Hahn-extension of such measures. In section 4 we show that associated

[1] This paper is based on a part of the author's Ph. D. Thesis written under the direction of Professor P. R. Masani at the University of Pittsburgh.

[2] A random measure is a special case of a *random field,* viz., a collection $\{x_\lambda: \lambda \in \Lambda\}$ of random variables, where Λ is any parameter set. For random measures, the parameter set Λ is a collection of subsets of a given set.

with every L_2-valued function with independent increments is a L_2-valued countably additive independently scattered measure. Finally in section 5 we discuss the relationship between countable additivity in the L_2-topology, countable additivity in the "pseudo topology" of a.e. convergence, and countable additivity in the topology of convergence in probability.

The following notation will be used throughout the paper.

NOTATION. (i) **R** and **C** denote the fields of real and complex numbers respectively.

(ii) $\mathbf{N}_{0+}$ and $\mathbf{N}_+$ denote the set of non-negative and positive integers respectively.

(iii) For a topological space Λ, $Bl(\Lambda)$ denotes the Borel family of Λ, i.e., the σ-algebra generated by the open sets of Λ.

(iv) $(\Omega, \mathscr{B}, P)$ is a probability space and $L_2 = L_2(\Omega, \mathscr{B}, P; \mathscr{F})$, $\mathscr{F} = \mathbf{R}$ or $\mathbf{C}$.

2. Definition and elementary properties of c.a.i.s. measures

The most convenient domain on which to define our measures initially is a pre-ring.

2.1. DEFINITION. We say that $\mathscr{P}$ is a *pre-ring* over the set Λ iff $\mathscr{P}$ is a non void collection of subsets of Λ such that for each $A, B \in \mathscr{P}$,

(i) $A \cap B \in \mathscr{P}$;

(ii) there exist $n \geqq 1$ and disjoint sets $C_1, C_2, \ldots, C_n \in \mathscr{P}$ such that $A \setminus B = \bigcup \{C_i\colon i = 1, 2, \ldots, n\}$.

2.2. DEFINITION. Let $\mathscr{P}$ be a pre-ring over Λ. Then we say that ξ is a L_2-valued *countably additive independently scattered (c.a.i.s.) measure* on $\mathscr{P}$ iff

(i) ξ is a function on $\mathscr{P}$ to L_2;

(ii) if $\langle A_k \rangle$ is a sequence of pairwise disjoint sets in $\mathscr{P}$ such that $A = \bigcup_1^\infty A_k \in \mathscr{P}$, then $\sum_{k=1}^\infty \xi(A_k)$ converges unconditionally in L_2 to $\xi(A)$;

(iii) if $n \geqq 1$, $A_1, A_2, \ldots, A_n \in \mathscr{P}$ and A_i are disjoint, then $\{\xi(A_i)\colon i = 1, 2, \ldots, n\}$ is a collection of independent random variables.

2.3. DEFINITION. Let $\mathscr{P}$ be a pre-ring over Λ. Then we say that ξ is a L_2-valued *countably additive orthogonally scattered (c.a.o.s.) measure* on $\mathscr{P}$ iff (i) and (ii) as in 2.2;

(iii) if $A, B \in \mathscr{P}$, and $A \cap B = \emptyset$, then $(\xi(A), \xi(B)) = 0$ where $(\cdot, \cdot)$ denotes the inner product in L_2.

It is trivial to show that not every c.a.o.s. measure is necessarily c.a.i.s. and vice versa.

The following lemma will be needed in the sequel.

2.4. LEMMA. *Let $\mathscr{P}$ be pre-ring over Λ and ξ be a L_2-valued c.a.i.s. measure on $\mathscr{P}$. Then*

(a) $E[\xi(\cdot)]$ *is an $\mathscr{F}$-valued measure on $\mathscr{P}$;*

(b) $\eta(\cdot) = \xi(\cdot) - E[\xi(\cdot)]$ *is a L_2-valued c.a.i.o.s. measure on $\mathscr{P}$.*

Using the above lemma we easily obtain the following necessary and sufficient condition for a c.a.i.s. measure ξ on a ring $\mathscr{R}$ to be orthogonally scattered.

2.5. PROPOSITION. *Let $\mathscr{R}$ be a ring over a set Λ, ξ be a L_2-valued c.a.i.s. measure on $\mathscr{R}$, and $m(\cdot)=E[\xi(\cdot)]$. Then the following are equivalent:*

(a) *ξ is orthogonally scattered on $\mathscr{R}$;*

(b) *$m(\cdot)$ assumes at most one non-zero value on $\mathscr{R}$;*

(c) *either for all $A\in\mathscr{R}$, $m(A)=0$ or m is concentrated on a single atom, i.e., there exists $A\in\mathscr{R}$ such that $m(A)\neq 0$, for all $B\in 2^A\cap\mathscr{R}$, $m(B)=0$ or $m(B)=m(A)$, and for all $C\in\mathscr{R}$ such that $C\cap A=\emptyset$, $m(C)=0$.*

PROOF. (a)$\Rightarrow$(b): Let ξ be orthogonally scattered. Suppose $\exists A, B\in\mathscr{R}$ such that $m(A)\neq 0\neq m(B)$. Let $\alpha=m(A\cap B)$. Then

$$0 = \big(\xi(A), \xi(B\setminus A)\big) = E[\xi(A)]E[\xi(B\setminus A)] = m(A)m(B\setminus A).$$

Hence $m(B\setminus A)=0$, and therefore $m(B)=m(A\cap B)+m(B\setminus A)=\alpha$. Using the same argument we can show that $m(A)=\alpha$.

(b)$\Rightarrow$(a): Suppose m assumes at most one non-zero value. Let $A, B\in\mathscr{R}$ and $A\cap B=\emptyset$. Since $m(A\cup B)=m(A)+m(B)$ at least one of the terms on the right hand side of the above equation must be zero. Hence $\big(\xi(A), \xi(B)\big)=m(A)m(B)=0$.

(b)$\Rightarrow$(c) follows from the fact that m is a countably additive $\mathscr{F}$-valued measure on $\mathscr{R}$.

The following example shows that for a c.a.i.o.s. measure ξ on a ring $\mathscr{R}$, it is not generally true that $E[\xi(\cdot)]$ is identically equal to zero.

2.6. EXAMPLE. Let (i) X_1 and X_2 be independent Gaussian random variables on $(\Omega, \mathscr{B}, P)$ such that $E[X_1]=1$, and $E[X_2]=0$; (ii) $\Lambda=\{1, 2\}$, and $\mathscr{R}=\{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$; (iii) $\xi(\{1\})=X_1$, $\xi(\{2\})=X_2$, and $\xi(\{1, 2\})=X_1+X_2$. It is easy to verify that ξ is a L_2-valued c.a.i.o.s. measure on $\mathscr{R}$. However $E[\xi(\cdot)]$ is not identically equal to zero.

Concerning c.a.o.s. measures, the Pythagorean identity [7, p. 14] at once yields

2.7. LEMMA. *Let ξ be a L_2-valued c.a.o.s. measure on a pre-ring $\mathscr{P}$ over Λ. Then $\|\xi(\cdot)\|^2$ is a finite, non-negative, countably additive measure on $\mathscr{P}$.*

2.8. DEFINITION. Let $\mathscr{P}$ be a pre-ring over Λ and ξ be a L_2-valued c.a.o.s. measure on $\mathscr{P}$. Then $\mu_\xi(\cdot)=\|\xi(\cdot)\|^2$ is called the *control measure* of ξ.

We conclude this section by showing that when the control measure of a c.a.i.o.s. measure ξ has the Darboux property, then expected value of ξ is identically equal to zero.

2.9. THEOREM. *Let $\mathscr{D}$ be a δ-ring over a set Λ and ξ be a L_2-valued c.a.i.o.s. measure on $\mathscr{D}$ such that μ_ξ has the Darboux property, i.e. for all $A\in\mathscr{D}$, $0\leqq\alpha\leqq\mu_\xi(A)\Rightarrow$ there exists $B\in\mathscr{D}\ni B\subseteqq A$ and $\mu_\xi(B)=\alpha$. Then for all $A\in\mathscr{D}$, $E[\xi(A)]=0$.*

PROOF. Let $m(\cdot)=E[\xi(\cdot)]$. By Proposition 2.5, m can assume at most one non-zero value β. Let $A\in\mathscr{D}$ and $m(A)=\beta$. Using countable additivity of μ_ξ and m and the Darboux property of μ_ξ, we can easily construct a sequence C_n in $\mathscr{D}$ such that $A\supseteqq C_1\supseteqq C_2\dots$ and for all $n\geqq 1$, $\mu_\xi(C_n)=\beta/n$ and $m(C_n)=\beta$. Therefore

$\mu_\xi\left(\bigcap_1^\infty C_n\right)=0$ and $m\left(\bigcap_1^\infty C_n\right)=\beta$. But $\mu_\xi\left(\bigcap_1^\infty C_n\right)=\left\|\xi\left(\bigcap_1^\infty C_n\right)\right\|^2$. Hence $\xi\left(\bigcap_1^\infty C_n\right)=0$ a.e. (P) and therefore

$$\beta = m\left(\bigcap_1^\infty C_n\right) = E\left[\xi\left(\bigcap_1^\infty C_n\right)\right] = 0.$$

3. Hahn extension of c.a.i.s. measures

Masani [12, pp. 68—70] has shown that every Hilbert space valued c.a.o.s. measure ξ on a pre-ring $\mathscr{P}$ can be uniquely extended to a c.a.o.s. measure on the δ-ring of sets of finite $\tilde{\mu}$ measure where $\tilde{\mu}$ is the Hahn extension of μ_ξ to the σ-ring generated by $\mathscr{P}$. In this section we will show (Theorem 3.8) that if ξ is a L_2-valued c.a.i.o.s. measure on $\mathscr{P}$, then the L_2-valued c.a.o.s. extension of ξ is in fact a L_2 valued c.a.i.s. measure. As a simple corollary of this result we get the Hahn extension theorem for c.a.i.s. measures (Theorem 3.10). The following assumptions will be understood.

3.1. Assumptions. (i) ξ is a L_2-valued c.a.i.o.s. measure on the pre-ring $\mathscr{P}$ of subsets of Λ.

(ii) $\tilde{\mu}$ is the (unique) non-negative countably additive extension of μ_ξ to the σ-ring Σ generated by $\mathscr{P}$;

(iii) $\mathscr{A}=\{A\colon A\in\Sigma \text{ and } \tilde{\mu}(A)<\infty\}$.

(iv) $\hat{\mathscr{P}}=\{E\colon E=\bigcup_1^n P_k,\ n\in\mathbf{N}_+,\ P_k\in\mathscr{P} \text{ and } P_k \text{ disjoint}\}$.

3.2. Lemma. *Let* $E\in\hat{\mathscr{P}}$ *and* $\bigcup_1^m P_i=E=\bigcup_1^n Q_j$, *where* $P_i, Q_j\in\mathscr{P}$, P_i *are pairwise disjoint and* Q_j *are pairwise disjoint. Then*

$$\sum_{i=1}^m \xi(P_i) = \sum_{i=1}^n \xi(Q_j).$$

Proof. See [12, Lemma A1.2, p. 112].

The last lemma shows that the following definition is unequivocal.

3.3. Definition. Let $E\in\hat{\mathscr{P}}$. We define

$$\tilde{\xi}(E) = \sum_{i=1}^n \xi(P_i)$$

where $\{P_i\colon i=1, 2, \dots, n\}$ is any disjoint collection of sets in $\mathscr{P}\ni E=\bigcup_1^n P_i$.

Using elementary properties of independent random variables one can easily prove the following result.

3.4. LEMMA. *$\tilde{\xi}$ given by* 3.3 *is a c.a.i.o.s. measure on the ring $\hat{\mathcal{P}}$ with control measure $\tilde{\mu}$ restricted to $\hat{\mathcal{P}}$.*

Following MASANI [12, pp. 68—70, 112—115], we use the denseness of $\hat{\mathcal{P}}$ in $\mathcal{A}$ in the metric given by

$$\text{for all} \quad A, B \in \mathcal{A}, \quad \varrho(A, B) = \tilde{\mu}(A \Delta B)$$

to extend $\tilde{\xi}$ to $\mathcal{A}$.

3.5. DEFINITION. Let $A \in \mathcal{A}$. We define

$$\tilde{\xi}(A) = \underset{n \to \infty}{\text{l.i.m.}}\, \tilde{\xi}(E_n)$$

where $\langle E_n \rangle$ is any sequence in $\hat{\mathcal{P}} \ni \varrho(E_n, A) \to 0$ as $n \to \infty$.

3.6. PROPOSITION. *The last definition is unequivocal and $\tilde{\xi}$ is a c.a.o.s. measure on $\mathcal{A}$ with control measure $\tilde{\mu}$ restricted to $\mathcal{A}$.*

PROOF. See [12, Theorem 2.3, p. 69].

In order to prove that the measure $\tilde{\xi}$ defined in 3.5 is independently scattered on $\mathcal{A}$, we need a lemma from probability theory which we shall give in its complete generality for future use.

3.7. LEMMA. *Let $m \in \mathbf{N}_+$ and $\langle {}^1X_n \rangle$, $\langle {}^2X_n \rangle, \dots, \langle {}^mX_n \rangle$ be sequences of random variables such that*

(i) *for all $n \geqq 1$, $\{{}^jX_n;\ j = 1, 2, \dots, m\}$ is a collection of independent random variables;*

(ii) *for all $j \in \{1, 2, \dots, m\}$, $\underset{n \to \infty}{\text{plim}}\, {}^jX_n = {}^jX$.*

Then the collection of (limiting) random variables $\{{}^1X, {}^2X, \dots, {}^mX\}$ is independent.

Proof is straightforward and will not be presented here.

3.8. THEOREM. *$\tilde{\xi}$ given by* 3.5 *is a L_2-valued c.a.i.o.s. measure on $\mathcal{A}$ with control measure $\tilde{\mu}$ restricted to $\mathcal{A}$.*

PROOF. By Proposition 3.6, $\tilde{\xi}$ is a c.a.o.s. measure on $\mathcal{A}$ with control measure $\tilde{\mu}$ restricted to $\mathcal{A}$. Hence it suffices to show that $\tilde{\xi}$ is independently scattered.
Let $n \geqq 1$, and $A_1, A_2, \dots, A_n$ be pairwise disjoint sets in $\mathcal{A}$. Since $\hat{\mathcal{P}}$ is dense in $\mathcal{A}$ in the topology induced by the metric

$$\text{for all} \quad A, B \in \mathcal{A}, \quad \varrho(A, B) = \tilde{\mu}(A \Delta B),$$

for all $i \in \{1, 2, \dots, n\}$, there exists a sequence $\langle {}^iE_k \rangle$ in $\hat{\mathcal{P}}$ such that $\lim_{k \to \infty} \varrho({}^iE_k, A_i) = 0$. Hence given $\varepsilon > 0$, there exist $k_0 \in \mathbf{N}_+$ such that for all $k \geqq k_0$ and $i \in \{1, 2, \dots, n\}$,

$$(1) \qquad \tilde{\mu}({}^iE_k \setminus A_i) < \varepsilon/(n+2) \quad \text{and} \quad \tilde{\mu}(A_i \setminus {}^iE_k) < \varepsilon/(n+2).$$

Let for all $k \geqq 1$, and $i \in \{1, 2, \dots, n\}$, ${}^iF_k = {}^iE_k \setminus \bigcup_{j \neq i} {}^jE_k$.

Clearly for all $k \geqq 1$ and $i \in \{1, 2, \dots, n\}$, ${}^iF_k \in \hat{\mathcal{P}}$. Moreover, for all $k \geqq 1$, ${}^iF_k$ are pairwise disjoint. Since ${}^iF_k \subseteq {}^iE_k$, for all $k \geqq k_0$, $i \in \{1, 2, \dots, n\}$ we have

$$\tilde{\mu}({}^iF_k \setminus A_i) < \varepsilon/(n+2). \tag{2}$$

Using the fact that A_j are pairwise disjoint, one can easily verify that

$$A_i \setminus {}^iF_k \subseteq (A_i \setminus {}^iE_k) \cup [(A_i \cap {}^iE_k) \setminus {}^iF_k] \subseteq (A_i \setminus {}^iE_k) \cup \left[\bigcup_{j=1}^{n} ({}^jE_k \setminus A_j)\right].$$

Therefore

$$\tilde{\mu}(A_i \setminus {}^iF_k) \leqq \tilde{\mu}(A_i \setminus {}^iE_k) + \sum_{j=1}^{n} \tilde{\mu}({}^jE_k \setminus A_j) \leqq \frac{\varepsilon}{n+2} + \frac{n\varepsilon}{n+2} = \frac{(n+1)\varepsilon}{n+2}. \tag{3}$$

Combining (2) and (3) we get

$$\text{for all} \quad k \geqq k_0, \; i \in \{1, 2, \dots, n\}, \; \tilde{\mu}(A_i \Delta {}^iF_k) < \varepsilon.$$

Hence by 3.5 we have for all $i \in \{1, 2, \dots, n\}$

$$\tilde{\xi}(A_i) = \underset{k \to \infty}{\text{l.i.m.}}\, \xi({}^iF_k).$$

Since for all $k \geqq 1$, ${}^iF_k$ are disjoint, $\{\tilde{\xi}({}^iF_k): i = 1, 2, \dots, n\}$ is a collection of independent random variables. Since convergence in L_2 implies convergence in probability, by lemma 3.7, $\{\tilde{\xi}(A_i): i = 1, 2, \dots, n\}$ is a collection of independent random variables.

As an immediate corollary of the above theorem we get:

3.9. Corollary. *If for all* $A \in \mathcal{P}$, $E[\xi(A)] = 0$, *then for all* $A \in \mathcal{A}$, $E[\xi(A)] = 0$.

Using lemma 2.4 and theorem 3.8 one can easily prove the following result on the Hahn extension of L_2-valued countably additive independently scattered measures.

3.10. Theorem. *Let η be a L_2-valued c.a.i.s. measure on the pre-ring $\mathcal{P}$ of subsets of Λ, and $\xi(\cdot) = \eta(\cdot) - E[\eta(\cdot)]$. Then* (a) *$\xi$ is a c.a.i.o.s. measure on $\mathcal{P}$*; (b) *η can be uniquely extended to a c.a.i.s. measure $\tilde{\eta}$ on the δ-ring*

$$\mathcal{A} = \{A\colon A \in \sigma\text{-}ring(\mathcal{P}) \quad and \quad \tilde{\mu}(A) < \infty\}$$

where $\tilde{\mu}$ is the Hahn extension of μ_ξ to the σ-ring $\mathcal{A}$ generated by $\mathcal{P}$.

4. Functions with independent increments

Every **C**-valued function on **R** which is locally of bounded variation gives rise to a **C**-valued measure on a sub δ-ring of $Bl(\mathbf{R})$ and every Hibert space valued function on **R** with orthogonal increments gives rise to a c.a.o.s. measure on a sub δ-ring of $Bl(\mathbf{R})$ [cf. 12, Theorem 8.6, pp. 100—101]. We will now show, in a similar vein, that every L_2-valued function on a subinterval Λ of $\mathbf{R}$ having independent increments gives rise to a L_2-valued c.a.i.s. measure on a sub δ-ring of $Bl(\Lambda)$.

4.1. DEFINITION. Let Λ be a subinterval of $\mathbf{R}$ and $\mathcal{M}$ be a family of $\mathcal{F}$-valued random variables on $(\Omega, \mathcal{B}, P)$ where $\mathcal{F}=\mathbf{R}$ or $\mathbf{C}$. Then we say that $x(\cdot)$ is an *$\mathcal{M}$-valued function on Λ with independent increments* iff (i) $x(\cdot)$ is a function on Λ to $\mathcal{M}$; (ii) for all $n\in\mathbf{N}_+$, $t_1, t_2, \dots, t_n\in\Lambda$, $t_1<t_2\leqq t_3<t_4\dots<t_n\Rightarrow\{x(t_j)-x(t_{j-1})$: $j=2, 3, \dots, n\}$ is a collection of independent random variables.

4.2. THEOREM. *Let* (i) *Λ be a subinterval of* $\mathbf{R}$ *and $x(\cdot)$ be a L_2-valued function on Λ with independet increments $\ni$for all $t\in\Lambda$, $E[x(t)]=0$;* (ii) *$\mathcal{P}=\{(a, b]$: $a\leqq b$ and $(a, b]\subseteqq\Lambda\}$.*

Then there exists a unique L_2-valued c.a.i.s. measure ξ on $\mathcal{P}$ such that for all $(a, b]\in\mathcal{P}$, $\xi(a, b]=x(b+)-x(a+)$.

PROOF. It is known [cf. 12, Theorem 8.6, p. 100] that there exists a unique c.a.o.s. measure ξ on $\mathcal{P}$ such that for all $(a, b]\in\mathcal{P}$, $\xi(a, b]=x(b+)-x(a+)$. Hence it suffices to show that ξ is independently scattered. Let $n\in\mathbf{N}_+$ and for all $i\in\{1, 2, \dots, n\}$, $(a_i, b_i]\in\mathcal{P}$ with $a_1<b_1\leqq a_2<b_2\dots<b_n$. By [5, p. 100] $x(\cdot+)$ is right continuous on L_2 and there exists a countable set $C\subseteqq\Lambda$ such that for all $t\in\Lambda\setminus C$, $x(t-)=x(t)=x(t+)$. Let $1\leqq i\leqq n$. Then there exist sequences $\langle {}^it_m\rangle$, $\langle {}^is_m\rangle$ in $\Lambda\setminus C$ such that ${}^it_m\downarrow a_i$ and ${}^is_m\downarrow b_i$. Clearly these sequences can be chosen such that for all $m\geqq 1$

$$a_i<{}^it_m<b_i<{}^is_m<a_{i+1} \quad \text{if} \quad b_i<a_{i+1}$$

$$a_i<{}^it_m<b_i=a_{i+1}<{}^is_m={}^{i+1}t_m<b_{i+1} \quad \text{if} \quad b_i=a_{i+1}.$$

Hence for all $m\geqq 1$, $\{x({}^is_m)-x({}^it_m)$: $i=1, 2, \dots, n\}$ is a sequence of independent random variables. Using right continuity of $x(\cdot+)$ in L_2, and the fact that convergence in L_2 implies convergence in probability we get for all $i\in\{1, 2, \dots, n\}$

$$\xi(a_i, b_i]=\operatorname*{plim}_{m\to\infty}\{x({}^is_m)-x({}^it_m)\}.$$

Hence by Lemma 3.7, $\{\xi(a_i, b_i]$: $i=1, 2, \dots, n\}$ is independent.

Our Hahn-extension theorem (Theorem 3.8) now guarantees that ξ has a unique c.a.i.s. extension $\tilde{\xi}$ to the δ-ring of sets of finite $\tilde{\mu}$ measure, where $\tilde{\mu}$ is the Hahn-extension of the control measure of ξ to the $Bl(\Lambda)$. We have thus established the following theorem.

4.3. THEOREM. *Let Λ be a subinterval of* $\mathbf{R}$ *and $x(\cdot)$ be a L_2-valed function on Λ with independent increments such that for all $t\in\Lambda$, $E[x(t)]=0$. Then*

(a) *there exists a unique non-negative, σ-finite, countably additive measure μ on $Bl(\Lambda)$ such that for all $(a, b]\subseteqq\Lambda$, $\mu(a, b]=\|x(b+)-x(a+)\|^2$;*

(b) *there exists a unique L_2-valued c.a.i.o.s. measure on $\mathcal{A}=\{A$: $A\in Bl(\Lambda)$ and $\tilde{\mu}(A)<\infty\}$ such that for all $(a, b]\subseteqq\Lambda$, $\xi(a, b]=x(b+)-x(a+)$.*

5. Different types of countable additivity

We now discuss three different types of countable additivity for a random measure depending on whether the "topology" used is that of L_2-convergence, convergence in probability or a.e. convergence.

5.1. THEOREM. *Let $\mathscr{R}$ be a ring of subsets of Λ and ξ be a L_2-valued independently scattered measure on $\mathscr{R}$. Then the following are equivalent:*

(a) *ξ is countably additive in the topology of convergence in probability; more fully, if $\langle A_k\rangle$ is a sequence of pairwise disjoint sets in $\mathscr{R}$ such that*

$$A=\bigcup_1^\infty A_k\in\mathscr{R},\quad \text{then}\quad \operatorname*{plim}_{n\to\infty}\sum_{k=1}^n \xi(A_k)=\xi(A);$$

(b) *ξ is countably additive in the (pseudo) topology of a.e. convergence; more fully, if $\langle A_k\rangle$ is a sequence of pairwise disjoint sets in $\mathscr{R}$ such that $A=\bigcup_1^\infty A_k\in\mathscr{R}$, then there exists $B\in\mathscr{B}$ (depending on the sequence $\langle A_k\rangle$) such that*

$$P(B)=0\quad \text{and for all}\quad \omega\in\Omega\setminus B,\quad \lim_{n\to\infty}\sum_{k=1}^n [\xi(A_k)](\omega)=[\xi(A)](\omega).$$

If in addition for all $A\in\mathscr{R}$, $E[\xi(A)]=0$ and there exists a non-decreasing function Φ on $\mathbf{R}_{0+}$ to itself such that for all $A\in\mathscr{R}$, there exists $N_A\in\mathscr{B}$ such that $P(N_A)=0$ and for all $\omega\notin N_A$, $|\xi(A)(\omega)|\leqq\Phi(\|\xi(A)\|^2)$, then (a) *or* (b) *is equivalent to:*

(c) *ξ is countably additive in the L_2-topology and therefore is a L_2-valued c.a.i.o.s. measure on $\mathscr{R}$.*

PROOF. (a)$\Leftrightarrow$(b)$\Leftarrow$(c) follow from basic properties of series of independent random variables [cf. 11, Theorem 4.2.8, p. 83, and Corollary 2 to Theorem 4.2.4, p. 70].

(b)$\Rightarrow$(c): Let $\langle A_k\rangle$ be a sequence of pairwise disjoint sets in $\mathscr{R}$ such that $A=\bigcup_1^\infty A_k\in\mathscr{R}$. For each $k\geqq1$ let $B_k=A\setminus A_k$. Then for all $k\geqq1$, $\xi(A_k)$, $\xi(B_k)$ are independent and hence orthogonal random variables. Moreover $\xi(A_k\cup B_k)(\omega)=$ $=\xi(A_k)(\omega)+\xi(B_k)(\omega)$ a.e. (P). Hence

$$\|\xi(A_k\cup B_k)\|^2=\|\xi(A_k)+\xi(B_k)\|^2=\|\xi(A_k)\|^2+\|\xi(B_k)\|^2\geqq\|\xi(A_k)\|^2.$$

Since for all $k\geqq1$, $A_k\cup B_k=A$ and Φ is non-decreasing, we have

$$|\xi(A_k)(\omega)|\leqq\Phi\big(\|\xi(A_k)\|^2\big)\leqq\Phi\big(\|\xi(A_k\cup B_k)\|^2\big)\leqq\Phi\big(\|\xi(A)\|^2\big)\quad \text{a.e.}\quad (P).$$

Therefore $\sum_{k=1}^\infty \xi(A_k)$ converges in L_2 to $\xi(A)$ [cf. 11, Corollary to Theorem 4.2.8, p. 83].

6. Concluding remarks

Let Λ be a subinterval of $\mathbf{R}$ and $(x(t)\colon t\in\Lambda)$ be a stochastic process with independent increments such that for all $t\in\Lambda$, $E[x(t)]=0$ and $\mathrm{Var}\,[x(t)]<\infty$. Then using Theorem 4.3 we obtain a c.a.i.o.s. measure ξ on a certain δ-ring of subsets of Λ. Since L_2-convergence implies convergence in probability, by Theorem 5.1, ξ is countably additive in the (pseudo) topology of a.e. convergence. Hence results on a.e. continuity of sample paths can be obtained as simple applications of "continuity" of measures [cf. 3, p. 39 and 6, p. 18]. Also a.e. differentiability of sample paths can be studied within a measure theoretic framework.

c.a.i.s. measures provide a natural generalization of processes with independent increments to the case where the "index" set is a general measure space [cf. 9, 10].

References

[1] H. Cramer, A contribution to the theory of stochastic processes, *Proc. 2nd Berkeley Symp. Math. Stat. Prob.* (1951), pp. 329—339.
[2] K. L. Chung, *A Course in Probability Theory,* Harcourt, Brace & World, Inc. (New York, 1968).
[3] N. Dinculeanu, *Vector Measures,* Pergamon Press (New York, 1967).
[4] N. Dinculeanu and I. Kluvanek, On vector measures, *London Math. Soc. Proc.,* **3** (1967).
[5] J. L. Doob, *Stochastic Processes,* John Wiley and Sons, Inc. (New York, 1953).
[6] P. R. Halmos, *Measure Theory,* van Nostrand (New York, 1950).
[7] P. R. Halmos, *Introduction to Hilbert Space,* Chelsea Publishing Co. (New York, 1951).
[8] J. F. C. Kingman, Completely random measures, *Pacific J. Math.,* **21** (1968), 59—78.
[9] S. Khalili, *Independently scattered measures,* Doctoral Thesis (University of Pittsburgh, 1975).
[10] S. Khalili, On the oscillation of the Brownian motion random measure, *Ann. Prob.,* **5** (1977).
[11] E. Lukacs, *Stochastic Convergence,* D. C. Heath (Lexington, Mass., 1968).
[12] P. R. Masani, Orthogonally scattered measures, *Advances in Math.,* **2** (1968), 61—117.
[13] P. R. Masani, Quasi-isometric measures and their applications, *Bull. Amer. Math. Soc.,* **76** (1970), 427—528.
[14] A. Prékopa, On stochastic set functions. I, *Acta Math. Acad. Sci. Hungar.,* **7** (1956), 215—263.
[15] A. Prékopa, On stochastic set functions. II, *Acta Math. Acad. Sci. Hungar.,* **8** (1957), 337—374.
[16] A. Prékopa, On stochasticset functions. III, *Acta Math. Acad. Sci. Hungar.,* **8** (1957), 375—400.
[17] K. Urbanik, Random measures and harmonizable sequences, *Studia Sci. Math. Hungar.,* **3** (1968), 6—88.

(Received June 8, 1977)

DEPARTMENT OF MATHEMATICS
THE CLEVELAND STATE UNIVERSITY
CLEVELAND, OHIO 44115
USA

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 43—45.

SUR LE PROLONGEMENT DES FONCTIONS

Par

Z. GRANDE (Elblag)

Soit R l'espace des nombres réels. Dans le travail [2] les auteurs ont démontré le théorème suivant:

THÉORÈME 0. *Soit* $A\subset[0, 1]$ *un ensemble. Pour qu'il existe pour toute fonction de première classe de Baire* g: $[0, 1]\to R$ *une fonction* f: $[0, 1]\to R$ *approximativement continue et telle que* $\{x\in[0, 1]: f(x)=g(x)\}\supset A$, *il faut et il suffit que l'ensemble* A *soit de mesure zéro.*

Dans cet article je démontre le théorème suivant:

THÉORÈME 1. *Soit* $A\subset[0, 1]$ *un ensemble. Pour qu'il existe pour toute fonction de première classe de Baire* g: $[0, 1]\to R$ *une fonction* f: $[0, 1]\to R$ *continue presque partout, approximativement continue et telle que* $\{x\in[0, 1]: f(x)=g(x)\}\supset A$, *il faut et il suffit que la fermeture* $\mathrm{Cl}(A)$ *de l'ensemble* A *soit de mesure zéro.*

Dans la démonstration de ce théorème j'appliquerai le lemme suivant:

LEMME 1. *Soit* $A\subset[0, 1]$ *un ensemble fermé de mesure zéro. Étant donnés les nombres réels* $a_1, \dots, a_m$, *s'il existe des ensembles* $B_1, \dots, B_m$ *disjoints deux à deux, étant à la fois du type* F_σ *et* G_δ *et tels que* $A=\bigcup_{i=1}^{m} B_i$, *il existe une fonction* f: $[0, 1]\to R$ *approximativement continue en tout point* $x\in[0, 1]$ *et continue en chaque point* $x\in[0, 1]-A$ *et telle que* $f(x)=a_i$ *pour tout point* $x\in B_i$ $(i=1, \dots, m)$.

DÉMONSTRATION. Il existe pour tout ensemble B_i $(i=1, \dots, m)$ une suite d'ensembles fermés et disjoints deux à deux, que nous désignons par C_k^i $(i=1, \dots, m$ et $k=1, 2, \dots)$ telle que $\bigcup_{k=1}^{\infty} C_k^i=B_i$ (voir [3]). Rangeons tous les ensembles C_k^i $(i=1, \dots, m$ et $k=1, 2, \dots)$ en une suite $D_1, D_2, \dots$ telle que $D_i\neq D_j$ pour $i\neq j$. Désignons par E_n $(n=1, 2, \dots)$ l'ensemble $\{x\in[0, 1]: \varrho(x, A)<1/n\}$, où $\varrho(x, A)= =\inf_{y\in A}\varrho(x, y)$ et $\varrho(x, y)$ désigne la distance des points x et y. Il existe pour tout ensemble D_i $(i=1, 2, \dots)$ un ensemble ouvert $G_i\subset E_i-A-\bigcup_{j=1}^{i-1}\mathrm{Cl}(G_j)$ tel que tout point de l'ensemble D_i est un point de densité de l'ensemble G_i et pour lequel $\mathrm{Cl}(G_i)\subset\left[E_i-A-\bigcup_{j=1}^{i-1}\mathrm{Cl}(G_j)\right]\cup D_i$. Soit f_i: $[0, 1]\to R$ une fonction approximativement continue en tout point $x\in[0, 1]$ et continue en tout point $x\in[0, 1]-D_i$ pour laquelle on ait $f_i(x)=a_k$ pour tout $x\in D_i$, où k désigne l'indice tel que $D_i\subset B_k$; et $f_i(x)=0$ pour tout point $x\in[0, 1]-\mathrm{Cl}(G_i)$.

Posons

$$f(x)=\begin{cases} f_i(x) & \text{pour} \quad x\in \mathrm{Cl}(G_i), \quad i=1,2,\dots \\ 0 & \text{pour} \quad x\in[0,1]-\bigcup_{n=1}^{\infty}\mathrm{Cl}(G_i). \end{cases}$$

Remarquons que $f(x)=a_k$ pour $x\in B_k$ $(k=1, \dots, m)$ et que la fonction f est continue en chaque point $x\in[0, 1]-A$. La fonction f est également approximativement continue en tout point $x\in A$. En effet, si $x\in A$, il existe l'indice naturel k tel que $x\in D_k\subset \mathrm{Cl}(G_k)$. Les fonctions f et f_k sont égales sur l'ensemble $\mathrm{Cl}(G_k)$, x est un point de densité de l'ensemble G_k et la fonction f_k est approximativement continue au point x, la fonction f est donc également approximativement continue en ce point.

DÉMONSTRATION DU THÉORÈME 1. *Suffisance.* Soit $A\subset[0, 1]$ un ensemble tel que sa fermeture soit de mesure zéro. Étant donnée une fonction bornée $g\colon [0, 1]\to R$ de première classe de Baire, il existe une suite de fonctions $\{g_n\}$ uniformément convergente vers la fonction g et telle que toutes les fonctions g_n sont de première classe de Baire et tous les ensembles des valeurs g_n ([0, 1]) sont finis (voir [1], pp. 294—295). D'après le lemme 1 il existe pour toute fonction g_n une fonction $f_n\colon [0, 1]\to R$ approximativement continue en tout point $x\in[0, 1]$ et continue en tout point $x\in[0, 1]-\mathrm{Cl}(A)$ et telle que $\{x\in[0, 1]\colon f_n(x)=g_n(x)\}\supset \mathrm{Cl}(A)$. Soit $\{g_{n_k}\}$ la sous-suite de la suite $\{g_n\}$ pour laquelle on ait $|g_{n_{k+1}}(x)-g_{n_k}(x)|<1/2^{k+1}$ pour tout $x\in[0, 1]$. Posons $\bar{h}_{n_1}(x)=f_{n_1}(x)$ pour $x\in[0, 1]$ et posons pour $i>1$,

$$h_{n_i}(x)=\begin{cases} 1/2^i & \text{lorsque} \quad f_{n_i}(x)-f_{n_{i-1}}(x)>1/2^i \\ f_{n_i}(x)-f_{n_{i-1}}(x) & \text{lorsque} \quad -1/2^i\leqq f_{n_i}(x)-f_{n_{i-1}}(x)\leqq 1/2^i \\ -1/2^i & \text{lorsque} \quad f_{n_i}(x)-f_{n_{i-1}}(x)<-1/2^i. \end{cases}$$

Soit $\bar{h}_{n_i}(x)=\bar{h}_{n_{i-1}}(x)+h_{n_i}(x)$ pour $x\in[0, 1]$. Toutes les fonctions $\bar{h}_{n_i}$ $(i=1, 2, \dots)$ sont approximativement continues partout et continues presque partout et la suite $\{\bar{h}_{n_i}\}$ est uniformément convergente, la fonction $f(x)=\lim_{i\to\infty}\bar{h}_{n_i}(x)$ est donc continue presque partout et approximativement continue. Comme, de plus, $\bar{h}_{n_i}(x)=f_{n_i}(x)=g_{n_i}(x)$ pour tout point $x\in \mathrm{Cl}(A)$, on a donc $f(x)=g(x)$ pour tout point $x\in A$.

Nécessité. Soit $A\subset[0, 1]$ un ensemble tel que sa fermeture $\mathrm{Cl}(A)$ soit de mesure positive. Il existe un ensemble fermé $B\subset \mathrm{Cl}(A)$ qui est non dense dans l'ensemble $\mathrm{Cl}(A)$, qui est dense en soi et qui est de mesure positive. Le complémentaire $[0, 1]-B$ de l'ensemble B est la somme de ses composantes J_n. Fixons un point x_n de l'ensemble A dans toute composante J_n qui coupe l'ensemble A. Il existe deux sous-suites $\{x_{n_k}\}$ et $\{x_{m_k}\}$ de la suite $\{x_n\}$ qui sont disjointes et denses dans B. Posons

$$g(x)=\begin{cases} 1 & \text{lorsque} \quad x=x_{n_k} \quad \text{et} \quad k=1,2,\dots \\ 0 & \text{lorsque} \quad x\notin\{x_{n_k}\}. \end{cases}$$

La fonction g est de première classe de Baire et toute fonction f: $[0, 1] \to R$ telle que $g(x)=f(x)$ pour tout point $x \in A$ n'est continue en aucun point $x \in B$.

REMARQUE. Dans le théorème 1, si la fonction g est bornée, la fonction f est bornée aussi.

Bibliographie

[1] K. KURATOWSKI, *Topologie I* (Warszawa, 1958).
[2] G. PETRUSKA, M. LACZKOVICH, Baire 1 functions, approximately continuous functions and derivatives, *Acta Math. Acad. Sci. Hungar.*, **25** (1974), 189—212.
[3] W. SIERPIŃSKI, Sur une propriété des ensembles F_σ linéaires, *Fund. Math.*, **14** (1929), 216—220.

(Reçu le 20 juin 1977.)

INSTYTUT MATEMATYKI
POLITECHNIKI GDAŃSKIEJ
GDANSK, POLSKA

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 47—57.

A FUNCTORIAL APPROACH TO NEAR-RINGS

By
S. FEIGELSTOCK and A. KLEIN (Ramat Gan)

Algebraic structures arise naturally as sets of homomorphisms of a given structure. For instance, rings arise as sets of endomorphisms of abelian groups. It seems that a natural algebraic structure carried by sets of mappings of groups is the structure of near-rings. They were introduced in various contexts and consequently they became a subject of independent interest. The first important investigations were performed by FRÖHLICH [6] who employed them in constructing non-abelian homology [7] and by H. NEUMANN [15] who employed them in studying varieties of groups.

In this paper a functorial approach to the subject is initiated. Near-$\mathcal{C}$-rings are introduced as structures on objects in fairly general categories $\mathcal{C}$, by generalizing an idea of R. BAER and FUCHS [8] who introduced groups of multiplications. This generalization follows the introduction of a tensor-like functor. Examples show that such structures are of interest not only in the category of groups.

The categories $\mathcal{C}$ are embedded in a natural way into categories of near-$\mathcal{C}$-rings. Various aspects of these embeddings are considered, including facts about subobjects, projectives and injectives, and extensions of mappings. In this sense the category Mod R over a semisimple ring is shown to be well-behaved. As a benefit we obtain results on the structure of classical near-rings; one of them is shown to be of geometrical significance. Another benefit is a generalization of a theorem of R. BAER [2] and H. NEUMANN [15] concerning hopfian groups.

Associativity conditions are formulated in the general setting. A representation theorem holds for associative near-$\mathcal{C}$-rings in the common categories. In certain cases, where zero-divisors are excluded, the associativity problem is related to questions about groups of fixed-point-free automorphisms, whose importance in group or ring theory is well-known.

The categorical approach leads naturally to the definition of the dual structure which is briefly sketched in the last section.

1. The functor $\bar{\otimes}$

Every object A in a category $\mathcal{C}$ determines a functor $\hom_{\mathcal{C}}(A, -)\colon \mathcal{C}\to\mathrm{Sets}$. If this functor has a left adjoint $F_A\colon \mathrm{Sets}\to\mathcal{C}$, namely if there is a natural equivalence $\hom_{\mathcal{C}}(F_A -, -)\cong \mathrm{Map}\,(-, \hom_{\mathcal{C}}(A, -))$, we denote $F_A S = S\bar{\otimes} A$ ($S\in \mathrm{ob\,Sets}$). As a left-adjoint functor, F_A preserves epimorphisms and colimits (=direct limits). It is easy to verify that a left adjoint to $\hom_{\mathcal{C}}(A, -)$ exists, iff there is a coproduct in $\mathcal{C}$ for every family of copies of A, and we may write

$S\tilde{\otimes} A=\coprod_{s\in S} A_s$, $A_s=A$ for all $s\in S$. For $\sigma\in\mathrm{Map}(S, T)$ the arrow $F_A\sigma\colon S\tilde{\otimes} A\to T\tilde{\otimes} A$ is uniquely determined by the family of arrows $A_s \xrightarrow{1} A_{\sigma s} \xrightarrow{\iota_{\sigma s}} \coprod_{t\in T} A_t$ (all $s\in S$), where the ι's are the canonical arrows into the coproduct. The covariant hom-functors take coproducts into products, hence for objects S in Sets and Y in $\mathscr{C}$ we obtain

$$\text{(1)}\qquad \hom_{\mathscr{C}}(S\tilde{\otimes}A, Y)\cong \hom_{\mathscr{C}}\big(\coprod_s A_s, Y\big)=\prod_s \hom_{\mathscr{C}}(A_s, Y)= \\ =\mathrm{Map}\big(S, \hom_{\mathscr{C}}(A, Y)\big).$$

The following is implied.

PROPOSITION 1.1. *If F_A exists for all objects A in $\mathscr{C}$ then $\tilde{\otimes}$ is a bifunctor* $\mathrm{Sets}\times\mathscr{C}\to\mathscr{C}$.

For $\sigma\in\mathrm{Map}(S, T)$, $\alpha\in\hom_{\mathscr{C}}(A, B)$ we denote the determined arrow $S\tilde{\otimes} A\to T\tilde{\otimes} B$ in $\mathscr{C}$ by $\sigma\tilde{\otimes}\alpha$.

2. $\tilde{\otimes}$ and $\otimes$

Consider the categories Mod R. Here the adjoint pair $\mathrm{Hom}_R(A, -)$; Mod $R\to\mathscr{A}b$, $-\underset{\mathbf{Z}}{\otimes} A\colon \mathscr{A}b\to$ Mod R, is well-known, with the natural isomorphism of abelian groups $\mathrm{Hom}_R(S\underset{\mathbf{Z}}{\otimes} A, Y)=\mathrm{Hom}_{\mathbf{Z}}\big(S, \mathrm{Hom}_R(A, Y)\big)$ for ${}_{\mathbf{Z}}S$, A_R, Y_R. The "near-tensor" $\tilde{\otimes}$ constructed above is an adjoint for Hom_R, when the latter is viewed as a functor into Sets. Hence $\mathrm{Hom}_{\mathbf{Z}}$ has to be replaced by Map and we obtain the equivalence indicated in (1). Strictly speaking, we have a composite functor, namely

$$\text{(2)}\qquad \mathrm{Mod}\,R\xrightarrow{\mathrm{Hom}_R(A,-)}\mathscr{A}b\xrightarrow{\Phi}\mathrm{Sets},$$

Φ the forgetful functor. A left-adjoint to Φ is the "free-functor" $\Psi\colon \mathrm{Sets}\to\mathscr{A}b$, $\Psi S=\sum_{|S|}\mathbf{Z}$, and we obtain a left-adjoint to (2) by composing the two left-adjoints

$$\text{(3)}\qquad \mathrm{Sets}\xrightarrow{\Psi}\mathscr{A}b\xrightarrow{-\underset{\mathbf{Z}}{\otimes} A}\mathrm{Mod}\,R.$$

(Recall $(F_1F_2)^*=F_2^*F_1^*$.) Indeed

$$\big(\sum_{|S|}\mathbf{Z}\big)\underset{\mathbf{Z}}{\otimes} A=\sum_{|S|}\mathbf{Z}\underset{\mathbf{Z}}{\otimes} A\cong\sum_{|S|} A = S\tilde{\otimes} A.$$

Thus "near-tensor" $\tilde{\otimes}$ may be viewed as the composite of tensor $\otimes$ with the free-functor.

3. Near-$\mathscr{C}$-rings

For an abelian group A, the group of multiplications on A, i.e. binary operations inducing ring structures, as introduced by BAER and FUCHS [8], is

$$\mathrm{Mult}\,A\cong\mathrm{Hom}_{\mathbf{Z}}(A\underset{\mathbf{Z}}{\otimes} A, A).$$

CLAY [3] has replaced multiplications by "left-multiplications" for which only one-sided (left) distrbiutivity is required.

Let $\mathscr{C}$ be a concrete category and A an object in $\mathscr{C}$ for which a left-adjoint F_A of $\hom_{\mathscr{C}}(A, -)$ exists. We denote

$$\text{(4)} \qquad \operatorname{Mult}^L A = \hom_{\mathscr{C}}(A \tilde{\otimes} A, A) = \operatorname{Map}(A, \hom_{\mathscr{C}}(A, A)),$$

and we call every $\tilde{\mu} \in \operatorname{Mult}^L A$ a (left-)near-multiplication on A. The pair $\langle A, \tilde{\mu} \rangle$ will be called a *near-$\mathscr{C}$-ring*. Every near-multiplication $\tilde{\mu}$ on A induces a binary operation μ on the underlying set

$$\text{(5)} \qquad a \cdot b = \mu(a, b) = \tilde{\mu}\iota_a(b),\ a, b \in A; \quad A_a \xrightarrow{\iota_a} \coprod_{s \in A} A_s \xrightarrow{\tilde{\mu}} A.$$

Conversely, for every binary operation μ on A with $\mu(s, -) \in \hom_{\mathscr{C}}(A, A)$ for all $s \in A$, the family of arrows $\mu(s, -)\colon A_s \to A$ $(s \in A)$ determines a unique arrow $\tilde{\mu}\colon \big(\coprod_{s \in A} A_s = \big) A \tilde{\otimes} A \to A$, satisfying (5).

Let $\langle A, \tilde{\mu} \rangle$ and $\langle B, \tilde{\nu} \rangle$ be near-$\mathscr{C}$-rings. The arrow $\varphi \in \hom_{\mathscr{C}}(A, B)$ is a *near-ring homomorphism* if the diagram

$$\text{(6)} \qquad \begin{array}{ccc} A \tilde{\otimes} A & \xrightarrow{\varphi \tilde{\otimes} \varphi} & B \tilde{\otimes} B \\ \downarrow{\scriptstyle \tilde{\mu}} & & \downarrow{\scriptstyle \tilde{\nu}} \\ A & \xrightarrow[\varphi]{} & B \end{array}$$

is commutative. This holds iff $\varphi(a \cdot a') = (\varphi a) \cdot (\varphi a)'$ for all $a, a' \in A$.

We obtain a category of near-$\mathscr{C}$-rings as objects and near-ring-homomorphisms as arrows. We denote it by $\operatorname{nr} \mathscr{C}$.

It might be observed that, although we employed the adjoint functor F_A, in the last definitions we only made use of $F_A A$. Therefore, if we disregard the functorial aspects, we can extend the definitions to objects A, in a concrete category $\mathscr{G}$, by skipping the middle in (4), namely $\operatorname{Mult}^L A = \operatorname{Map}(A, \hom_{\mathscr{C}}(A, A))$.

In (4) the only map $A \to \hom_{\mathscr{C}}(A, A)$ always present is the constant map $a \mapsto 1_A$ (for all $a \in A$), that is the codiagonal $\nabla\colon A \tilde{\otimes} A \to A$. The induced binary operation is $a \cdot b = b$ for all $a, b \in A$. We shall refer to it as the trivial operation.

If there is a zero-arrow $\zeta_A\colon x \mapsto o_A$ (all $x \in A$) in $\hom_{\mathscr{C}}(A, A)$, then the map $a \mapsto \zeta_A$ of A into $\hom_{\mathscr{C}}(A, A)$ induces the zero operation, $a \cdot b = o_A$ for all $a, b \in A$. In this case the "characteristic maps" χ_B of subsets $B \subset A$, namely $\chi_B\colon b \mapsto 1_A (b \in B)$, $x \mapsto \zeta_A (x \notin B)$, induce near-multiplications $a_1 \cdot a_2 = a_2 (a_1 \in B)$, $a_1 \cdot a_2 = o_A (a_1 \notin B)$.

There is no reason for considering "left" and not "right" multiplications. For those preferring to write mappings on the "left", right near-multiplications might be suitable. (See $\operatorname{Map}(X, X)$, next section.)

4. Examples

The category nrSets is clearly the category of ordinary groupoids.

The near-Mod R-rings are $\langle A, \tilde{\mu}\rangle$, $\tilde{\mu}\colon \sum_{s\in A} A_s \to A$ an R-homomorphism. The induced binary operation is linear in its second argument, namely $a\cdot(b+b')=a\cdot b+a\cdot b'$, $a\cdot(br)=(a\cdot b)r$ (for all $a, b, b'\in A, r\in R$), since $\mu(a, -)$ is an R-homomorphism.

The near-Gr-rings turn out to be the "classical" near-rings, not necessarily associative. $A\otimes A$ is the free product of $|A|$ copies of A, μ is induced as in (5) and it is left-distributive with respect to the group operation, written $+$ for convenience. A natural example of a near-Gr-ring is the following. For any group $X, +$ consider the group $A, +$ with $A=\mathrm{Map}\,(X, X)$, mappings written on the right, and viewed as a product $\prod_{|X|} X, +$ in Gr. The usual composition of maps induces a near-ring structure on $A, +$.

The above are special cases of near-Ωalg-rings in the category Ω alg of algebras with a fixed set of operations Ω [5]. A near-Ωalg-ring is an Ω-algebra endowed with a binary operation $\notin\Omega$, $a\cdot b$, left distributive with respect to all $\omega\in\Omega$, i.e. $a\cdot(b_1\dots b_k\omega)=(a\cdot b_1)\dots(a\cdot b_k)\omega$ for any k-ary $\omega\in\Omega$ and any $a, b_1, \dots, b_k\in A$. Again $\mathrm{Map}\,(X, X)$, as a product $\prod_{|X|} X$ in Ωalg, (namely, for k-ary $\omega\in\Omega$ and $f_1, \dots, f_k\in\mathrm{Map}\,(X, X)$, $f_1, \dots f_k\omega\colon x\mapsto(xf_1)\dots(xf_k)\omega)$, and with the composition of maps is a near-Ωalg-ring. The Ω-algebra X is embedded into $\mathrm{Map}\,(X, X)$ by the diagonal $\Delta\colon x\to\hat{x}=$the constant map $X\to X$ taking everything into x. The only near-ring structure on X for which $\Delta\colon X\to\mathrm{Map}\,(X, X)$ is an embedding is the trivial structure mentioned above, since in this case $\widehat{a\cdot b}$ must be $\hat{a}\circ\hat{b}=\hat{b}$.

A special near-Gr-ring, introduced by H. NEUMANN [15] in studying varieties of groups, was widely generalized and investigated by FRÖHLICH [6]. We describe a generalization to Ωalg. For an arbitrary Ω-algebra X, $\hom_{\Omega\,\mathrm{alg}}(X, X)$ is obviously not, in general, an algebra if the operations are defined as above in $\mathrm{Map}\,(X, X)$. In the following case, however, we modify the definitions to make it an Ω-algebra. Let $\mathscr{K}$ be a subcategory of Ωalg with free algebras and let X be $\mathscr{K}$-free on a set G [5]. We write mappings on the left, and for k-ary ω and endomorphisms $f_1, \dots, f_k\colon X\to X$ we denote by $\omega f_1\dots f_k$ the unique endomorphism $X\to X$ determined by $(\omega f_1\dots f_k)a=\omega(f_1 a)\dots(f_k a)$ for all $a\in G$. Then $A=\hom_{\Omega\,\mathrm{alg}}(X, X)$ is an Ω-algebra and the composition of maps induces a near-Ωalg-ring structure on A. Following a well-known definition for groups, we call X hopfian if X is not isomorphic to a proper quotient of itself, or equivalently — if every surjective endomorphism of X is an automorphism. Then the following is obtained similarly to theorems of H. NEUMANN [15] and R. BAER [2]. *A $\mathscr{K}$-free algebra X is hopfian iff the relation $\alpha\cdot\beta=1$ in the near-Ωalg-ring $\hom_{\Omega\,\mathrm{alg}}(X, X)$ implies that α (hence β) is invertible. If X is not hopfian then X is isomorphic to one of its proper subalgebras.*

Returning to the general case, we observe that (5) "clearly" implies that μ inherits properties of the arrows of $\mathscr{C}$, in its second argument. One example was "linearity" in the algebraic case. An additional example of this occurrence is continuity, in the category Tp of topological spaces and continuous mappings. A near-Tp-ring is a space A with a binary operation on A continuous in its second argument. (Here $A\otimes A$ is the union-space of $|A|$ disjoint copies of A and $\tilde{\mu}\iota_a=\mu(a, -)$

is a continuous map on A.) The following is similar to the above algebraic examples. For a space X, take $A=\text{Map}(X, X)=X^X$ with the product topology. The usual composition of maps is continuous in its second argument, i.e. for fixed f, the map $\bar{f}: A \to A$, $g \mapsto f \circ g$ is continuous. Indeed, for all $t \in X$, $A \xrightarrow{\bar{f}} A \xrightarrow{\pi_t} X$ is the evaluation map $g \mapsto (tf)g$ and this is π_{tf}, hence all $\bar{f}\pi_t$ are continuous. It follows that $A, \circ$ is a near-Tp-ring. Again the diagonal is an embedding of near-Tp-rings iff the operation on X is trivial. These facts remain true for instance in Tp_2(T_2-spaces).

Combining the algebraic and topological structures may give rise to interesting near-$\mathscr{C}$-rings in the various categories of topological algebras.

5. Associativity

Let $\langle A, \tilde{\mu}\rangle$ be a near-$\mathscr{C}$-ring. The binary operation μ is associative iff the selected map $A \to \hom_{\mathscr{C}}(A, A)$ turns out to be a homomorphism of $A, \cdot$ into $\hom_{\mathscr{C}}(A, A)$ with the composition in $\mathscr{C}$. This is the case in the examples $A=\text{Map}(X, X)$ of section 4. The following representation theorem reveals once more that $\text{Map}(X, X), \circ$ assumes in certain cases the role of End X in ring theory. We denote λ_a, resp. ϱ_a, the map $x \mapsto a \cdot x$, resp. $x \mapsto x \cdot a$, of X into X.

THEOREM 5.1. *Let $\mathscr{C}=\Omega\text{alg}$ or* Tp *and $\langle X, \tilde{\mu}\rangle$ a near-$\mathscr{C}$-ring. The operation μ is associative iff the map $a \mapsto \varrho_a$ is a near-ring-homomorphism $X \mapsto \text{Map}(X, X)$. If there is a left cancellable element for the operation μ then $a \mapsto \varrho_a$ is an embedding of near-$\mathscr{C}$-rings.*

In the general case a near-multiplication on A is determined by a set $M \subset \hom_{\mathscr{C}}(A, A)$, $0<|M|<|A|$, and a partition $A=\bigcup_{\varphi \in M} A_\varphi$ with $\mu_a=\varphi$ iff $a \in A_\varphi$, and $a \cdot b=\mu_a(b)$ for all $a, b \in A$. Evidently this operation is associative iff $\mu_a \circ \mu_b = \mu_{\mu_a(b)}$ for all $a, b \in A$. In particular $M, \circ$ has to be a subsemigroup of $\hom(A, A), \circ$. It follows

PROPOSITION 5.2. *Let M be a subsemigroup of* $\hom_{\mathscr{C}}(A, A), \circ$ *and $A=\bigcup_{\varphi \in M} A_\varphi$ a partition of A. The operation determined by μ, with $\mu_a=\varphi \Leftrightarrow a \in A_\varphi$, is associative iff for all $\varphi, \psi \in M$, φ maps A_ψ into $A_{\varphi \circ \psi}$. An operation is induced on the quotient set of A making it into a semigroup isomorphic to M.*

Call a semigroup M of endomorphisms *regular* if there is a partition of A as in the proposition. Obviously $\{\varphi\}$ with $\varphi^2=\varphi$, is always regular.

A semigroup $\{1_A, \varphi\}$ with $\varphi^2=\varphi$ is regular iff φ is not surjective: with any A_φ, $A \neq A_\varphi \supset \operatorname{im} \varphi$ and with $A_1=A-A_\varphi$ we obtain a suitable partition. In particular in the cases with zeros $\{1_A, 0_A\}$ goes with $A_0 \neq A$, $0 \in A_0$, $A_1=A-A_0$ (a fact mentioned for groups in [13]).

If A has a fixed point for all its endomorphisms, then a regular semigroup M must possess a right zero, namely that endomorphism ζ for which A contains the fixed point. Thus in this case $M \neq \{1\}$ is never a group, but $M-\{\zeta\}$ can be a group, as shown by the following theorem, which seems to be related to questions of fixed-point-free automorphisms in various categories.

THEOREM 5.3. *Let G be a group of automorphisms of an object A in the category $\mathcal{C}$.*

(i) *G is regular iff all its elements except 1_A are fixed-point-free. In this case A with the induced operation is a semigroup with left-hand cancellation and left identities.*

(ii) *In the cases with zeros, $G\cup\{0_A\}$ is regular iff $x\in A$ is either fixed by all $\varphi\neq 1_A$ of G or not fixed by any element of G except 1_A. If $0\in A$ is the only fixed point it will be a two-sided zero for the induced operation and there will not be zero divisors.*

PROOF. Suppose $A=\bigcup_{\varphi\in G} A$ is a suitable partition. If $\varphi(c)=c\in A_\psi$ then, since $\varphi(c)$ has to be in $A_{\varphi\circ\psi}$, it follows $\varphi\circ\psi=\psi$, so $\varphi=1_A$. Conversely assume G satisfies the conditions of (i). Then A decomposes into disjoint orbits under the action of G, each orbit containing $|G|$ elements. Collect one element from each orbit into a set, say A_1, and for every $\varphi\in G$ put $A_\varphi=\{\varphi(x)|x\in A_1\}$. It follows that $A=\cup A_\varphi$ is a suitable partition. All the elements of A_1 will be left identities. In the cases with zeros take A_0=set of all fixed points and apply the above argument to $A-A_0$.

In Gr the term regular group of automorphisms is well established [9]. By the theorem, if G is a regular group of automorphisms on A then $\{0_A\}\cup G$ is regular in the sense of our definition. The near-Gr-rings produced by Theorem 5.3 are near-integral-domains in the sense of [4], with left identities, namely associative near-rings with no zero-divisors and with $0\cdot x=0$ for all $x\in A$. Moreover there are unique solutions to $a\cdot x=c$ for $a\neq 0$ and any c. Conversely a near-integral-domain with (not necessarily unique) left identity e and right inverses for e is induced by a regular group of automorphism.

Observe that the choice of A_1 in Theorem 5.3 ($|G|^{|A|/|G|}$ possibilities in the finite case, $|G|^{(|A|-1)/|G|}$ where there are zeros), determines the near-$\mathcal{C}$-ring structure on the object A.

An element $b\in A$ is *distributive* if $\varrho_b\in \hom_{\mathcal{C}}(A,A)$. Any element $b\in A$ commuting with all $x\in A$ is clearly distributive. We denote by $D(A)$ the set of distributive elements of A. FRÖHLICH [6] considered associative near-Gr-rings A for which $D(A)$ generates the group A. The following is shown for Gr in [6].

PROPOSITION 5.4. *For $X\in$* ob Ωalg

$$D(\operatorname{Map}(X,X))=\hom(X,X). \tag{7}$$

PROOF. For $g\in\hom_{\Omega\,\mathrm{alg}}(X,X)$, $f\mapsto f\circ g$ is the Ω-endomorphism of $\prod_{|X|} X$ determined by the family of homomorphisms $\pi_t\circ g\colon \prod X\to X_t(=X)$, for all $t\in X$. Conversely, for $g\in D(\operatorname{Map}(X,X))$ and any k-ary $\omega\in\Omega$ and $t_1,\dots,t_k\in X$ take $f_1,\dots,f_k\in\operatorname{Map}(X,X)$ with $t_1f_i=t_i$ $(i=1,\dots,k)$. Hence

$$(t_1\dots t_k\omega)g=((t_1f_1)\dots(t_1f_k)\,\omega)g=t_1(f_1\dots f_k\,\omega)\circ g=t_1((f_1\circ g)\dots(f_k\circ g)\,\omega)=$$

$$=(t_1f_1g)\dots(t_1f_kg)\,\omega=(t_1g)\dots(t_kg)\,\omega$$

proving that $g\in\hom_{\Omega\,\mathrm{alg}}(X,X)$.

PROPOSITION 5.5. (7) *holds in* Tp.

PROOF. Here we have to show that for fixed g, the map $f \mapsto f \circ g$ is continuous on X^X iff g is continuous on X. The map $f \mapsto f \circ g$ on the product is determined by the family of maps $\pi_t g$ as above hence continuous, if g is. Conversely if $f \mapsto f \circ g$ is continuous and U is open in X, then $Ug^{-1} = (Ug^{-1}\pi_t^{-1})\pi_t$ is open since π_t is open; hence g is continuous.

We call $\langle X, \tilde{\mu} \rangle \in \mathrm{ob}\,(\mathrm{nr}\ \Omega\mathrm{alg})$ *distributively generated* if there is no proper sub-Ω-algebra of X which contains $D(X)$. This definition generalizes the definition of FRÖHLICH [6].

THEOREM 5.6. *If* $\hom_{\Omega\mathrm{alg}}(X, X), \circ$ *is commutative then any distributively generated* $\langle X, \tilde{\mu} \rangle$ *is associative.*

PROOF. The subset $S = \{x \mid x \in X,\ a \cdot (b \cdot x) = (a \cdot b) \cdot x \text{ for all } a, b \in X\}$ is a sub-Ω-algebra; for k-ary ω and any $x_1, \dots, x_k \in S$ write y for $x_1 \dots x_k \omega$ and

$$a \cdot (b \cdot y) = a \cdot \big((b \cdot x_1) \dots (b \cdot x_k)\,\omega\big) = \big((a \cdot b) \cdot x_1\big) \dots \big((a \cdot b) \cdot x_k\big)\,\omega =$$
$$= (a \cdot b) \cdot (x_1 \dots x_k\, \omega) = (a \cdot b) \cdot y,$$

hence $y \in S$. If $x \in D(X)$ then

$$(a \cdot b) \cdot x = \varrho_x\big(\lambda_a(b)\big) = \lambda_a\big(\varrho_x(b)\big) = a \cdot (b \cdot x),$$

so $x \in S$, hence $S \supset D(X)$. It follows $X = D(X)$.

REMARK. A known problem in Abelian group theory is to determine for which A the set of associative multiplications on A is a subgroup of Mult A, + [11]. For $\mathrm{Mult}^L A$ the answer is trivial: unless $|A| = 1$, the set of associative near-multiplications on A is not a groupoid with respect to +. (Proof. Since ∇ is associative, 2∇ must be associative and $a2\nabla(b2\nabla c) = (a2\nabla b)2\nabla c$ implies $2c = 4\boldsymbol{c}$ for all c. Thus A is 2-elementary $\cong \sum \mathbf{Z}_2$. For more then one copy of $\mathbf{Z}_2$ a non-associative near-multiplication is given in [11]. For $\mathbf{Z}_2$ three out of the four elements of $\mathrm{Mult}^L \mathbf{Z}_2 = \mathrm{Map}\,(\mathbf{Z}_2, \mathbf{Z}_2)$, namely $0, \nabla, 1_{\mathbf{Z}_2}$ induce associative near-multiplications, while $\nabla + 1_{\mathbf{Z}_2}$ does not.)

6. Subobjects in nr$\mathscr{G}$

Let $\langle A, \tilde{\mu} \rangle$ be a near-$\mathscr{C}$-ring and $K \subset A$. If the immersion $\iota\colon K \hookrightarrow A$ is in $\mathscr{C}$ and if $x \cdot x' \in K$ for all $x, x' \in K$ then we call K a sub-near-$\mathscr{C}$-ring and denote $K < A$. Following [6] we call $K < A$ a *left-* (resp. *right-*) *module* in A if $a \cdot x \in K$ (resp. $x \cdot a \in K$) for all $a \in A$, $x \in K$. Given $\langle A, \tilde{\mu} \rangle$, an element of A is a right-zero for μ iff it is a fixed point of $\tilde{\mu}\iota_a$ for all $a \in A$. This is the case with the zeros in Gr, Mod R, $\dot{\mathrm{T}}\mathrm{p}$(=topological spaces with base points). If $\varphi \in \hom_{nr\mathscr{C}}(\langle A, \tilde{\mu} \rangle, \langle B, \tilde{\nu} \rangle)$ and if there are right-zeros in B (for ν), then $K = \{a \mid \varphi a \text{ is a right-zero}\}$ is a left-module in A, provided $K \hookrightarrow A$ is in $\mathscr{C}$. In Ωalg, $\{\hat{x} \mid x \in X\} = \Delta X \subset \mathrm{Map}\,(X, X)$ (Section 4) is both a left and right-module since $\alpha \circ \hat{x} = \hat{x}$, $\hat{x} \cdot \alpha = \widehat{x\alpha}$ for any $\alpha \in \mathrm{Map}\,(X, X)$. In $\dot{\mathrm{T}}\mathrm{p}$, $\Delta X \subset \mathrm{Map}\,(X, X)$ is a left and right-module. (It is even normal as a subobject in $\dot{\mathrm{T}}\mathrm{p}$ or $\dot{\mathrm{T}}\mathrm{p}_2$).

We call $K \subset A$ fully-invariant in A if $K \hookrightarrow A$ is in $\mathscr{C}$ and $\varphi K \subset K$ for all $\varphi \in \hom_{\mathscr{C}}(A, A)$.

THEOREM 6.1. *K is a left-module in every near-$\mathscr{C}$-ring on A iff K is fully-invariant in A.*

PROOF. We observe that K is a left-module with respect to $\tilde{\mu}$ iff $\tilde{\mu}\iota_a K \subset K$ for all $a \in A$. Some of the μ's are $\tilde{\alpha}$ with a fixed $\alpha \in \hom_{\mathscr{C}}(A, A)$ and $\tilde{\alpha}\iota_s = \alpha$ for all $s \in A$.

The following is an example in $\dot{\mathrm{T}}\mathrm{p}$. The connected component K of the base point o_A in a space A is fully-invariant (and normal, even in $\dot{\mathrm{T}}\mathrm{p}_2$) since $\varphi \in \hom_{\dot{\mathrm{T}}\mathrm{p}}(A, A)$ takes a connected subspace into a connected subspace and o_A into o_A.

The following facts are consequences of the preceding theorem, applied to Gr. We omit proofs.

COROLLARY 6.2. *Let $A, +$ be a group.*

(i) *If there is a near-*Gr*-ring on A without proper left module, then $A, +$ is abelian or perfect.*

(ii) *If $A, +$ is abelian then either A is p-elementary ($\sum \mathbf{Z}_p$) or divisible torsion-free ($\sum \mathbf{Q}$) or reduced homogeneous torsion-free.*

(iii) *No near-*Gr*-ring on A possesses a proper left-module iff A is cyclic of prime order.*

(iv) *If every subgroup of A is normal and a left module in every near-*Gr*-ring on A, then A is abelian, and if in addition A is finite then A is cyclic. Conversely if A is a locally cyclic torsion group then every subgroup is normal and a left module in every near-*Gr*-ring on A.*

The consequence (ii) 6.2 is of geometrical significance. A Veblen—Wedderburn system [12], [10] is a near-Ab-ring R with $0 \cdot x = 0$ for all x and whose non-zero elements form a loop under multiplication satisfying the following condition: for all $a, b, c \in R$, if $a \neq b$ there is a unique solution in R to $a \cdot x - b \cdot x = c$. This condition clearly implies that in this case R does not possess a proper left-module, hence (ii) of the corollary yields considerable information about the group $R, +$. In particular one obtains that its non-zero elements are either all of a fixed prime order p or all of infinite order. This fact is important in geometry and it is part of a theorem of BAER [1], [10]. It may be easily seen that a Veblen—Wedderburn system of prime order p must be a field, and this means that every Veblen—Wedderburn plane of order $p(p^2+p+1)$ must be Desarguesian [12].

7. Embedding $\mathscr{C}$ into nr$\mathscr{C}$

THEOREM 7.1. *$\mathscr{C}$ is fully embedded into* nr$\mathscr{C}$. *If $\mathscr{C}$ has (finite) products, resp. difference kernels, then* nr$\mathscr{C}$ *has (finite) products, resp. difference kernels. The embedding $\mathscr{C} \to$* nr$\mathscr{C}$ *is continuous.*

PROOF. An embedding $\mathscr{C} \to \mathrm{nr}\mathscr{C}$ is $A \mapsto \langle A, \nabla \rangle$ for objects, $\varphi \mapsto \varphi$ for arrows. Now assume $A = \prod A_k$, with projections π_k, is a product in $\mathscr{C}$. Given near-multiplications $\tilde{\mu}_k \colon A_k \otimes A_k \to A_k$, we define $\tilde{\mu} \colon A \otimes A \to A$ as the unique arrow determined by $\pi_k \tilde{\mu} = \tilde{\mu}_k(\pi_k \otimes \pi_k) \colon A \otimes A \to A_k \otimes A_k \to A_k$. Hence the π_k's are in nr$\mathscr{C}$ and $\langle A, \tilde{\mu} \rangle$ with the π_k's is a product of $\langle A_k, \tilde{\mu}_k \rangle$ in nr$\mathscr{C}$. Indeed, with $\alpha_k \in \hom_{\mathrm{nr}\mathscr{C}}(\langle X, \xi \rangle, \langle A_k, \tilde{\mu}_k \rangle)$, the unique α with $\pi_k \alpha = \alpha_k$ satisfies $\pi_k \tilde{\mu}(\alpha \otimes \alpha) = \pi_k \alpha \xi$ for all k. In particular with the trivial structures $\tilde{\mu}_k = \nabla$ on A_k we get the trivial $\tilde{\mu} = \nabla$ on $\prod A_k$.

Let $\varkappa\colon K\to A$ be a difference kernel in $\mathscr{C}$ of $\varphi, \psi\in \hom_{nr\mathscr{C}}(\langle A, \tilde{\mu}\rangle, \langle B, \tilde{\nu}\rangle)$. Then $\varphi\tilde{\mu}(\varkappa\tilde{\otimes}\varkappa)=\psi\tilde{\mu}(\varkappa\tilde{\otimes}\varkappa)$ hence there is a unique $\tilde{\lambda}\colon K\tilde{\otimes}K\to K$ such that $\varkappa\tilde{\lambda}=\tilde{\mu}(\varkappa\tilde{\otimes}\varkappa)$, so $\varkappa$ is in nr$\mathscr{C}$. If $\varkappa'\in\hom_{nr\mathscr{C}}(\langle K', \lambda'\rangle, \langle A, \tilde{\mu}\rangle)$ and $\varphi\varkappa'=\psi\varkappa'$ then the arrow δ satisfying $\varkappa\delta=\varkappa'$ is in nr$\mathscr{C}$ (since $\varkappa$ is monic and $\varkappa\delta\lambda'=\varkappa\tilde{\lambda}(\delta\tilde{\otimes}\delta)$). Obviously $\tilde{\mu}=\nabla$ on A implies $\tilde{\lambda}=\nabla$ on K.

The α's with α fixed $\in\hom_{\mathscr{C}}(A, A)$, yield a full embedding of the following subcategory of $\mathscr{C}^2$ (=the arrow category for $\mathscr{C}$) into nr$\mathscr{C}$. Denote by e$\mathscr{C}$ the category whose objects are all endomorphisms of $\mathscr{C}$ and whose arrows from $A\xrightarrow{\alpha}A$ to $B\xrightarrow{\beta}B$ are all $\varphi\in\hom_{\mathscr{C}}(A, B)$ satisfying $\varphi\alpha=\beta\varphi$. The embedding e$\mathscr{C}\to$nr$\mathscr{C}$ is $\alpha\mapsto\langle A, \alpha\rangle$ for objects, $\varphi\mapsto\varphi$ for arrows. $\mathscr{C}$ is fully embedded into nr$\mathscr{C}$ via e$\mathscr{C}$. The embedding $\mathscr{C}\to$e$\mathscr{C}$ is $A\mapsto 1_A$, $\varphi\mapsto\varphi$.

8. Extensions

PROPOSITION 8.1. *Let $\varphi\in\hom_{\mathscr{C}}(A, B)$ and assume that for every $\beta\in\hom_{\mathscr{C}}(B, B)$ there is an $\alpha\in\hom_{\mathscr{C}}(A, A)$ satisfying $\varphi\alpha=\beta\varphi$ (that is $\varphi\in\hom_{e\mathscr{C}}(\alpha, \beta)$). Then for every near-multiplication $\tilde{\nu}$ on B there is a near-multiplication $\tilde{\mu}$ on A such that $\varphi\in\hom_{nr\mathscr{C}}(\langle A, \tilde{\mu}\rangle, \langle B, \tilde{\nu}\rangle)$.*

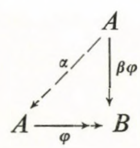
PROOF. For $a\in A$, $\tilde{\nu}\iota_{\varphi a}\in\hom_{\mathscr{C}}(B, B)$, hence there is an arrow $\mu_a\in\hom_{\mathscr{C}}(A, A)$ such that $\varphi\tilde{\mu}_a=\tilde{\nu}\iota_{\varphi a}\varphi$. Let $\tilde{\mu}\colon A\tilde{\otimes}A\to A$ be the unique arrow with $\tilde{\mu}\iota_a=\tilde{\mu}_a$ for all $a\in A$. Then

$$\varphi\tilde{\mu}\iota_a=\varphi\tilde{\mu}_a=\tilde{\nu}\iota_{\varphi a}\varphi=\tilde{\nu}(\varphi\tilde{\otimes}\varphi)\iota_a$$

implying $\varphi\tilde{\mu}=\tilde{\nu}(\varphi\tilde{\otimes}\varphi)$.

COROLLARY 8.2. *If A is projective in $\mathscr{C}$ and $\varphi\in\hom_{\mathscr{C}}(A, B)$ is an epimorphism then every near-multiplication on B can be "lifted by φ" to a near-multiplication on A, i.e. for every $\tilde{\nu}\colon B\tilde{\otimes}B\to B$ there is a $\tilde{\mu}\colon A\tilde{\otimes}A\to A$ such that $\varphi\in\hom_{nr\mathscr{C}}(\langle A, \tilde{\mu}\rangle, \langle B, \tilde{\nu}\rangle)$.*

PROOF. For every $\beta\in\hom_{\mathscr{C}}(B, B)$ there is an α for which



is commutative.

COROLLARY 8.3. *If a product $A'\times A''=A$ exists in $\mathscr{C}$, then every near-multiplication on A' can be lifted by $\pi\colon A\to A'$ (projection of the product) to a near-multiplication on A.*

PROOF. For every $\beta\in\hom_{\mathscr{C}}(A, A')$ we have $\pi(\beta\times 1_{A''})=\beta\pi$.

COROLLARY 8.4. *Let $\mathscr{C}$ be abelian. If $0\to A''\to A\xrightarrow{\eta}A'\to 0$ is split exact, in particular if A'' is injective or if A' is projective, then every near-multiplication on A' can be lifted by η to a near-multiplication on A.*

PROOF. A' is a direct summand of A.

PROPOSITION 8.5. *Let A be an injective object in $\mathscr{C}$ and assume that $B < A$ and that the arrow $\coprod B \to \coprod A$ determined by $B \hookrightarrow A$ is a monomorphism. (This holds in particular in C_1-categories* [14].*) Then every $\varphi \in \hom_{\mathscr{C}}(B \tilde{\otimes} B, A)$ extends to a near-multiplication on A.*

PROOF. Given $B \hookrightarrow A$, the assumption implies that the arrow $\coprod_{|B|} B \to \coprod_{|A|} A$ is a monomorphism. So there is a $\tilde{\varphi}$ such that

$$\begin{array}{ccc} \coprod_{|B|} & \longrightarrow & \coprod_{|A|} \\ {\scriptstyle\varphi}\big\downarrow & \swarrow{\scriptstyle\tilde{\varphi}} & \\ A & & \end{array}$$

is commutative.

COROLLARY 8.6. *If A is injective in $\mathscr{C}$, $\mathscr{C}$ a $\mathscr{C}_1$-category, and if $\iota \in \hom_{\mathscr{C}}(B, A)$ is a monomorphism then every near-multiplication on B can be "extended by ι" to a near-multiplication on A, i.e. for every $\tilde{\nu}\colon B \tilde{\otimes} B \to B$ there is a $\tilde{\mu}\colon A \tilde{\otimes} A \to A$ such that $\iota \in \hom_{\mathrm{nr}\mathscr{C}}(\langle B, \tilde{\nu}\rangle, \langle A, \tilde{\mu}\rangle)$.*

COROLLARY 8.7. *Let R be a semisimple ring. Every near-multiplication on a submodule (resp. quotient module) of a module M_R can be extended (resp. lifted) to M_R.*

PROOF. M_R is both projective and injective and $\mathrm{Mod}\, R$ is C_1.

9. The dual

We sketch the definition of the dual concept. For an object A in $\mathscr{C}$ there is the (covariant) functor $\hom_{\mathscr{C}^0}(A, -)\colon \mathscr{C}^0 \to \mathrm{Sets}$ ($\mathscr{C}^0$ the dual of $\mathscr{C}$). If this functor has a left adjoint $F^A\colon \mathrm{Sets} \to \mathscr{C}^0$, we denote $F^A S = S \underset{\sim}{\otimes} A$ and we get a (natural) equivalence $\hom_{\mathscr{C}}(Y, S \underset{\sim}{\otimes} A) = \mathrm{Map}\,(S, \hom_{\mathscr{C}}(Y, A))$. The left adjoint F^A exists iff every family of copies of A has a product in $\mathscr{C}$ and it preserves epimorphisms and colimits (hence $\mathrm{Sets} \xrightarrow{F^A} \mathscr{C}^0 \to \mathscr{C}$ takes epimorphisms, resp. colimits, into monomorphisms, resp. limits, in $\mathscr{C}$). We may write $S \underset{\sim}{\otimes} A = \prod_{s \in S} A_s$ ($A_s = A$ for all $s \in S$). Consequently $\hom_{\mathscr{C}}(Y, S \underset{\sim}{\otimes} A) = \prod_{s \in S} \hom_{\mathscr{C}}(Y, A_s)$, since the contravariant hom-functors preserve products. The co-near-multiplications on A are the arrows $\mu \in \hom_{\mathscr{C}}(A, A \underset{\sim}{\otimes} A) = \text{co-Mult}^L A$. It follows that if both $A \underset{\sim}{\otimes} A$ and $A \tilde{\otimes} A$ are defined then $\text{co-Mult}^L A = \mathrm{Mult}^L A$.

References

[1] R. BAER, Homogeneity of projective planes, *Amer. J. Math.*, **64** (1942), 137—152.

[2] R. BAER, Groups with proper isomorphic quotient groups, *Bull. AMS*, **50** (1967), 267—278.

[3] J. R. CLAY, The group of left distributive multiplications of an abelian group, *Acta Math., Acad. Sci. Hungar.*, **19** (1968), 221—227.

[4] J. R. CLAY, Generating balanced incomplete block design for planar near-rings, *J. Algebra*, **22** (1972), 319—331.

[5] P. M. COHN, *Universal Algebra,* Harper & Row (N. Y., 1965).
[6] A. FRÖHLICH, Distributively generated near-rings I, *Proc. London Math. Soc.,* **8** (1958), 76—94.
[7] A. FRÖHLICH, Non-abelian homological algebra I, *Proc. London Math. Soc.,* **(3) 11** (1961), 239—275.
[8] L. FUCHS, *Abelian groups,* Akadémiai Kiadó (Budapest, 1966).
[9] D. GORENSTEIN, *Finite groups,* Harper & Row (N. Y., 1968).
[10] M. HALL, Jr., *The theory of groups,* Macmillan (N. Y., 1959).
[11] F. L. HARDY, On groups of ring multiplications, *Acta Math. Acad. Sci. Hungar.,* **14** (1963), 283—294.
[12] D. R. HUGHES, Review of some results in collineation groups, *First symp. in Pure Math.,* A.M.S., 1959.
[13] J. J. MALONE, Near-rings with trivial multiplication, *Amer. Math. Monthly,* **74** (1967), 1111—1112.
[14] B. MITCHELL, *Theory of categories,* Academic Press (N. Y., 1965).
[15] H. NEUMANN, On varieties of groups and their associated near-rings, *Math. Z.,* **65** (1956), 36—69.

(Received June 24, 1977)

DEPARTMENT OF MATHEMATICS
BAR-ILAN UNIVERSITY
RAMAT GAN
ISRAEL

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 59—62.

TWO THEOREMS ON ABSOLUTELY CONTINUOUS SET FUNCTIONS

By

W. C. BELL (Murray)

1. Introduction

Suppose S is a set, F a field of subsets of S, $r(F)$ the set of functions from F into R, $b(F)$ the set of elements of $r(F)$ which are bounded and $ba(F)$ the set of elements of $b(F)$ which are (finitely) additive on F. Let μ and λ be in $ba(F)^+$ where, for $G \subseteq r(F)$, G^+ is the set of non-negative valued elements of G.

In section three we will prove that λ is μ-continuous iff the subset, $\mathcal{S}_\mu^+$, of functions summable with respect to μ [4] is contained in the set, M_λ, of functions measurable with respect to λ [7].

In section four we will show that the λ-continuous part of μ is the limit of successive applications of a certain nonlinear function.

2. Preliminary theorems

All integrals in this paper are refinement limits of sums over finite subdivisions of S by elements of F. If $\alpha \in r(F)$ and $\int_S \alpha(I)$ exists, then the function $\left\{\left(V, \int_V \alpha(I)\right) \middle| V \in F\right\}$ will be denoted by $\int \alpha$. For further details concerning the integral and the theorems of this section we refer the reader to [4].

Denote the μ-continuous elements of $ba(F)$ by A_μ and the μ-continuous part of λ by $\alpha_\mu(\lambda)$ which can be written $\sup_K \int \min\{\lambda, K\mu\}$.

The following theorem is due to KOLMOGOROFF [10].

THEOREM 2.K. *If* $\alpha \in r(F)$ *and* $\int_S \alpha(I)$ *exists, then* $\int_S \left|\alpha(V) - \int_V \alpha(I)\right|$ *exists and is* 0.

COROLLARY 2.K.1. *If* $\beta \in b(F)$, $\alpha \in r(F)$, $V \in F$ *and* $\int_S \alpha(I)$ *exists, then* $\int_V \beta(I)\alpha(I)$ *exists iff* $\int_V \beta(I) \int_I \alpha(J)$ *exists in which case they are equal.*

COROLLARY 2.K.2. *Suppose each of* α *and* β *is in* $r(F)$, *each of* $\int_S \alpha(I)$ *and* $\int_S \beta(I)$ *exists and* M *is either the function* max *or* min. *Then for each* $V \in F$ *we have* $\int_V M\{\alpha(I), \beta(I)\}$ *exists iff* $\int_V M\left\{\int_I \alpha(J), \int_I \beta(J)\right\}$ *exists in which case they are equal.*

Combining parts of Theorem 4.1 of [3] and Theorem 1.A.3 of [5] we have

THEOREM 2.A. *Suppose* $\eta\in A_\mu^+$ *and* $\beta\in b(F)^+$. *If* $\int\beta\mu$ *exists (and is zero), then* $\int\beta\eta$ *exists (and is zero).*

We close this section by noting that, in what follows, unless otherwise stated convergence of sequences of elements of $ba(F)$ will be with respect to the total variation norm.

3. The summability-measurability characterization

For $\alpha\in r(F)$ and $p\leqq 0\leqq q$ denote $\max\{\min\{\alpha(I), q\}, p\}$ by $\alpha(p, q)(I)$ for each $I\in F$. In [4] APPLING defined the class, $\mathcal{S}\mu$, of functions summable with respect to μ and the (linear) summability operator, s_μ, from $\mathcal{S}_\mu$ onto A_μ. Subsequently [6] (Theorem 3.1) he proved the following theorem.

THEOREM 3.A. *If* $\alpha\in r(F)$, *then these are equivalent:*
1) $\alpha\in\mathcal{S}_\mu$
2) *For each* $p\leqq 0\leqq q$, $\int\alpha(p, q)\mu$ *exists and* $\lim\limits_{p,q\to\infty}\int\alpha(p, q)\mu=s_\mu(\alpha)$.

Adopting the convention $a/0=0$ we have that if $\eta\in A_\mu$, then $\eta/\mu\in\mathcal{S}_\mu$ and $s_\mu(\eta/\mu)=\eta$. In particular, if $\alpha\in\mathcal{S}_\mu$, then $s_\mu(s_\mu(\alpha)/\mu)=s_\mu(\alpha)$.

The set $\mathcal{S}_\mu$ is clearly contained in the set, M_μ, of functions measurable with respect to μ defined by: $\alpha\in r(F)$ is in M_μ iff 1) $\int\alpha(p, q)\mu$ exists for each $p\leqq 0\leqq q$ and 2) if $c>0$, then there exists a $T>0$ and a subdivision D of $S(D\ll\{S\})$ such that for each E which refines D $(E\ll D)$ we have $\sum\limits_{E'}\mu(I)<c$ where $E'=\{I\in E|\,|\alpha(I)|>T\}$.

THEOREM 3.1. *If* $\mu(V)=0$ *implies* $\lambda(V)=0$, *then these are equivalent:*
1) $\mathcal{S}_\mu^+\subseteqq M_\lambda$
2) $\lambda/\mu\in M_\lambda$
3) $\lambda\in A_\mu$.

PROOF. Since one implies two is obvious we begin by supposing two is true. Let $c>0$. Let $D_1\ll\{S\}$ and $K>0$ be such that if $E\ll D_1$ and $E'=\{I\in E|\,\lambda(I)/\mu(I)>K\}$, then $\sum\limits_{E'}\lambda(I)<c/2$. Let $D_2\ll\{S\}$ be such that if $E\ll D_2$, then $\sum\limits_{E}\min\{\lambda(I), K\mu(I)\}-\int\limits_{S}\min\{\lambda(I), K\mu(I)\}<c/2$. Now let $D\ll\{S\}$ be such that $D\ll D_1$ and $D\ll D_2$ and let $E\ll D$ with $E'=\{I\in E|\,K<\lambda(I)/\mu(I)\}$. Then

$$0\leqq\lambda(S)-\int\limits_{S}\min\{\lambda(I), K\mu(I)\}\leqq\lambda(S)-\sum_{E}\min\{\lambda(I), K\mu(I)\}+c/2=$$

$$=c/2+\sum_{E}\lambda(I)+\sum_{E}\max\{-\lambda(I), -K\mu(I)\}=c/2+\sum_{E}\max\{0, \lambda(I)-K\mu(I)\}=$$

$$=c/2+\sum_{E'}\max\{0, \lambda(I)-K\mu(I)\}\leqq c/2+\sum_{E'}\lambda(I)<c/2+c/2=c.$$

Therefore $\lambda=\sup\limits_{K}\int\min\{\lambda, K\mu\}=\alpha_\mu(\lambda)\in A_\mu$.

Before proving that three implies one we note that if $\alpha\in\mathcal{S}_\mu$ and $s_\mu(\alpha)=0$, then $\lim\limits_{-p,q\to\infty}\int|\alpha(p, q)|\mu=0$. Since for $r\leqq p\leqq 0\leqq q\leqq s$ we have $|\alpha(p, q)|\leqq|\alpha(r, s)|$

it follows that $\int |\alpha(p,q)|\mu=0$. Consequently, by 2.A we have that if $\eta\in A_\mu^+$, then $\int |\alpha(p,q)|\eta$ exists and is zero, hence $\alpha\in\mathcal{S}_\eta\subseteqq M_\eta$.

Now suppose three is true and $\alpha\in\mathcal{S}_\mu^+$. We will first consider te case where $\int \alpha\mu$ exists and is in A_μ. Let $c>0$ and $d>0$ be such that $\mu(V)<d$ implies $\lambda(V)<c$. Since $\int \alpha\mu\in A_\mu$ there exists a $K>0$ such that

$$\int_S \alpha(I)\mu(I)-\int_S \min\Big\{\int_V \alpha(I)\mu(V), K\mu(I)\Big\} = \int_S \alpha(I)\mu(I)-\int_S \min\{\alpha(I), K\}\mu(I)<d/2$$

and therefore a $D\ll\{S\}$ such that if $E\ll D$, then

$$\sum_E (\alpha(I)-\min\{\alpha(I), K\})\mu(I)<d.$$

Now let $E\ll D$ and $E'=\{I\in E\,|\,\alpha(I)>K+1\}$. If $I\in E'$, then $\alpha(I)-\min\{\alpha(I),K\}=$ $=\alpha(I)-K>1$ so that

$$\mu(\cup\{I|I\in E'\}) = \sum_{E'}\mu(I)\leqq\sum_{E'}(\alpha(I)-\min\{\alpha(I), K\})\mu(I)\leqq$$
$$\leqq\sum_E(\alpha(I)-\min\{\alpha(I), K\})\mu(I)<d.$$

Consequently $\sum_{E'}\lambda(I)=\lambda(\cup\{I|I\in E'\})<c$ and therefore $\alpha\in M_\lambda$.

In general we have, for $\alpha\in\mathcal{S}_\mu^+$, $\alpha=s_\mu(\alpha)/\mu+\alpha-s_\mu(\alpha)/\mu$ with $s_\mu(\alpha)/\mu$ in M_λ since $\int (s_\mu(\alpha)/\mu)\mu\in A_\mu$ and $\alpha-s_\mu(\alpha)/\mu$ in M_λ since $s_\mu(\alpha-s_\mu(\alpha)/\mu)=0$. Therefore since M_λ is closed under addition we have $\alpha\in M_\lambda$.

4. The limit theorem

We begin this section with certain theorems and observations concerning the function R_μ and refer the reader to [8] for further details.

Let $H_\mu=\Big\{\eta\in A_\mu\,\Big|\,\int_S \eta(I)^2/\mu(I)$ exists$\Big\}$. HELLINGER [9] has sown that $\xi\in A_\mu^+$ iff there exists an $\eta\in H_\mu$ such that $\xi=\int \eta^2/\mu$. Consequently the function defined on H_μ by $T_\mu(\eta)=\int\eta^2/\mu$ is onto A_μ^+. The restriction of T_μ to H_μ^+ is one to one and the inverse coincides, on A_μ^+, with the function $R_\mu(\delta)=\int(\delta\mu)^{1/2}$ which is defined on all of $ba(F)^+$. For $\delta\in ba(F)^+$ we have $R_\mu(\delta)\in H_\mu$ and $\alpha_\mu(\delta)=T_\mu(R_\mu(\delta))$. R_μ is continuous with respect to the variation norm on $ba(F)^+$ and the stronger μ norm $(\|\delta\|^2)=\int_S \dfrac{\delta(I)^2}{\mu(I)}$ on H^+. Also, if M is either the function max or min and each of η and δ is in $ba(F)^+$, then

$$R_\mu\Big(\int M\{\eta,\delta\}\Big)=\int M\{R_\mu(\eta), R_\mu(\delta)\}.$$

For interval functions theorems similar to 4.1 can be found in [1] (Theorem 3) and [2] (Theorem 2).

THEOREM 4.1. $R_\mu^n(\lambda)\to\alpha_\lambda(\mu)$ $(n\to\infty)$.

PROOF. If $\mu \leqq \lambda$, then $R_\mu^n(\lambda)$ is nonincreasing with n. From induction and Theorem 2.2 of [4] we have

$$R_\mu^n(\lambda) = \int \mu^{1-1/2^n} \lambda^{1/2^n}$$

and since, for $0<p<1$, the expression $\sum_E \lambda(I)^p \mu(I)^{1-p}$ is nonincreasing with refinements of E it follows that:

$$\mu(S) \leqq R_\mu^n(\lambda)(S) = \int_S \lambda(I)^{1/2^n} \mu(I)^{1-1/2^n} \leqq \lambda(S)^{1/2^n} \mu(S)^{1-1/2^n} =$$
$$= (\lambda(S)/\mu(S))^{1/2^n} \mu(S) \to \mu(S).$$

Consequently $R_\mu^n(\lambda) \to \mu$.

If $\lambda \leqq \mu$, then $R_\mu^n(\lambda)$ is nondecreasing with n and $R_\mu^n(\lambda) \leqq \mu$ for each n so that there is a $\delta \in ba(F)^+$ such that $R_\mu^n(\lambda) \to \delta$ and $\lambda \leqq \delta \leqq \mu$. Therefore $A_\lambda \subseteqq A_\delta$ and since $R_\mu^n(\lambda) \in A_\lambda$ for each n and A_λ is closed (variation norm) we have $\delta \in A_\lambda$ so that $A_\delta = A_\lambda$ and $\alpha_\delta = \alpha_\lambda$. By the continuity of R_μ we have $R_\mu^n(\lambda) \to R_\mu(\delta)$ (H_μ norm) and it follows by the comparability of norms on H_μ that $\delta = R_\mu(\delta)$. Consequently $R_\mu^n(\lambda) \to \delta = T_\delta(\delta) = T_\delta(R_\mu(\delta)) = T_\delta(R_\delta(\mu)) = \alpha_\delta(\mu) = \alpha_\lambda(\mu)$.

In general we have (denoting $\int \min\{\mu, \lambda\}$ by ξ)

$$\lim R_\mu^n(\lambda) = \lim[-\mu + R_\mu^n(\lambda) + \mu] = -\mu + \lim\left[\int \max\{\mu, R_\mu^n(\lambda)\} + \int \min\{\mu, R_\mu^n(\lambda)\}\right] =$$
$$= -\mu + \lim R_\mu^n\left(\int \max\{\mu, \lambda\}\right) + \lim R_\mu^n\left(\int \min\{\mu, \lambda\}\right) = -\mu + \mu + \lim R_\mu^n(\xi) =$$
$$= \alpha_\xi(\mu) = \sup_K \int \min\{\mu, K\xi\} = \sup_K \int \min\left\{\mu, K \int \min\{\mu, \lambda\}\right\} =$$
$$= \sup_K \int \min\{\mu, K\mu, K\lambda\} = \sup_K \int \min\{\mu, K\lambda\} = \alpha_\lambda(\mu).$$

References

[1] W. D. L. APPLING, Interval functions and absolute continuity, *Illinois J. Math.*, **8** (1964), 439—445.

[2] W. D. L. APPLING, A convergence theorem concerning real valued interval functions, *Math. Zeitschr.*, **85** (1964), 314—318.

[3] W. D. L. APPLING, Some integral characterizations of absolute continuity, *Proc. Amer. Math. Soc.*, **18** (1967), 94—99.

[4] W. D. L. APPLING, Summability of real-valued set functions, *Riv. Mat. Parma*, **(2), 8** (1967), 77—100.

[5] W. D. L. APPLING, Addendum to: "Some integral characterizations of absolute continuity", *Proc. Amer. Math. Soc.*, **24** (1970), 788—793.

[6] W. D. L. APPLING, Continuity and set function summability, *Ann. di Mat. pur. ed Applic.*, **XXXVII** (1970), 357—374.

[7] W. D. L. APPLING, A notion of measurability for real-valued set functions, *Boll. U.M.I.*, **(4), 7** (1973), 42—49.

[8] WAYNE C. BELL, A characterization theorem for the nearest point mappings of linear C-sets, *Ann. Mat. Pura Appl.*, to appear (accepted).

[9] E. HELLINGER, *Die Orthogonalinvarianten quadratischer Formen von unendlichnielen Variablen*, dissertation (Göttingen, 1907).

[10] A. KOLMOGOROFF, Untersuchen über den Integralbegriff, *Math. Ann.*, **103** (1930), 654—696.

(Received June 27, 1977)

MURRAY STATE UNIVERSITY
DEPARTMENT OF MATHEMATICS
MURRAY, KENTUCKY 42071
USA

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 63—66.

DYADIC MATRICES AND WALLMAN COMPACTIFICATION

By

A. ABIAN[1] (Ames) and S. SALBANY[2] (Oxford)

Let m and n be any finite or infinite cardinal numbers and let M be an m by n *dyadic* matrix (i.e., whose entries are restricted solely to the real numbers 0 and 1). Thus, for every $i<m$ the i-th row r_i of M is a dyadic sequence $(r_{ij})_{j<n}$ of type n.

In what follows all matrices are dyadic and any statement concerning the multiplication (or product) of the rows of a matrix, refers to the coordinate-wise multiplication of these rows.

A row of a matrix is called the *zero row* if and only if every entry of that row is 0. We say that a set S of rows of a matrix *generates* the zero row of that matrix if and only if the product of some finite number of rows of S is equal to the zero row of that matrix (in this connection we observe that S is not required to be closed under multiplication).

Let E' be a set of rows of an m by n' (where n' like m is any finite or infinite cardinal) matrix M' and let $\bar{0}'$ be the zero row of M'. We say that $\bar{0}'$ is *covered* by the rows of E' if and only if for every $j<n'$ it is the case that

$$r'_{ij}=0 \quad \textit{for some} \quad r'_i \in E'. \tag{1}$$

From (1) it follows that if F' is a finite set of rows of a matrix M' then the zero row $\bar{0}'$ of M' is covered by the rows of F' if and only if $\bar{0}'$ is the product of the rows of F'. From this we have immediately:

LEMMA. *Let M' be a matrix and E' be a set of rows of M'. Then the zero row $\bar{0}'$ of M' is covered by no finite number of rows of E' if and only if $\bar{0}'$ is not generated by E'.*

The *sum* of rows r_i and r_h of a matrix is defined to be the row r_k of that matrix where:

$$r_{kj}=1 \quad \textit{if and only if} \quad r_{ij}=1 \quad \textit{or} \quad r_{hj}=1.$$

The sum of r_i and r_h is denoted by $r_i \dotplus r_h$.

The *complement* of a row r_i of a matrix is defined to be the dyadic sequence which is obtained by exchanging the 0's and 1's which appear in r_i. The complement of r_i is denoted by $C(r_i)$.

The columns c_t of a matrix are defined to be distinct in the *strict sense* if and only if $j\neq k$ implies:

$$r_{ij}=r_{hk}=0 \quad \textit{and} \quad r_{ik}=r_{hj}=1 \quad \textit{for some} \quad i \quad \textit{and} \quad h.$$

[1] The first author acknowledges an F.I.L. grant from Iowa State University.
[2] The second author acknowledges a grant from C.S.I.R., South Africa.

Based on the above notions and definitions we prove:

THEOREM. *Let M be an m by n matrix whose rows are pairwise distinct and include the zero row $\bar{0}$. Then M can be extended to an m by $n'>n$ matrix M' via extending every row r_i of M to a row r' of M' such that:*

(i) *The rows of M' include the zero row $\bar{0}'$.*

(ii) *The correspondence $r_i \to r_i'$ is one to one and preserves the zero, the products, the sums and the complements, i.e., $(\bar{0})'=\bar{0}'$, and*

$$(r_i \cdot \ldots \cdot r_h)' = r_i' \cdot \ldots \cdot r_h' \quad \text{for every} \quad i, \ldots, h < m \quad \text{and}$$

$$(r_i \dotplus \ldots \dotplus r_h)' = r_i' \dotplus \ldots \dotplus r_h' \quad \text{for every} \quad i, \ldots, h < m \quad \text{and}$$

$$(C(r_i))' = C(r_i') \quad \text{for every} \quad i < m,$$

provided $r_i \cdot \ldots \cdot r_h \in M$, or $r_i \dotplus \ldots \dotplus r_h \in M$, or $C(r_i) \in M$, respectively.

(iii) *The columns of the matrix $M'-M$ are distinct in the strict sense.*

(iv) *If $\bar{0}'$ is covered by the rows of a set E' of rows of M' then $\bar{0}'$ is already covered by finitely many rows of E'.*

PROOF. Let us consider the set N_j of rows r_i of M given by:

$$N_j = \{r_i | i < m \quad \text{and} \quad r_{ij} = 1\} \quad \text{for every} \quad j < n. \tag{2}$$

Clearly, N_j does not generate the zero row $\bar{0}$ of M.

By Zorn's lemma, every set (such as N_j given in (2)) of rows of M which does not generate $\bar{0}$ is contained (as a subset) in a set of rows of M which is maximal with respect to the property of not generating $\bar{0}$. Let

$$\{N_h^* | h \in H\} \quad \text{with} \quad N_h^* \neq N_k^* \quad \text{if} \quad h \neq k \quad \text{for every} \quad h, k \in H \tag{3}$$

be the set of *all* the sets N_h^* of rows of M such that each N_h^* is maximal w.r.t. not generating $\bar{0}$ and such that $N_h^* \neq N_j$ for every $h \in H$ and $j<n$.

Thus, the elements of the set given in (3) are pairwise distinct and different from the elements of the set given in (2). Let the cardinal number n' be given by:

$$n' = n + \overline{\overline{H}}. \tag{4}$$

Clearly, $n' \geqq n$ and from (2), (3), (4) with an obvious renaming of the elements of H it follows that:

$$\{N_j | j < n\} \cup \{N_j^* | n \leqq j < n'\} \tag{5}$$

contains among its elements *all* the sets of rows of M which are maximal w.r.t. not generating $\bar{0}$.

Recalling that every row r_i of M is a dyadic sequence r_{ij} of type n, to every r_i we correspond a dyadic sequence r_i' of type n' according to the following rule:

$$r_{ij}' = r_{ij} \quad \text{for} \quad j < n, \tag{6}$$

$$r_{ij}' = 1 \quad \text{for} \quad n \leqq j < n' \quad \text{if and only if} \quad r_i \in N_j^*.$$

Let M' be a matrix hose rows are r_i' with $i<m$. Obviously, M' is an m by n' dyadic matrix and is an extension of M obtained by extending r_i to r_i'.

From (3), (4), (5) it follows that $\bar{0} \notin N_j^*$ for every $n \leqq j < n'$ which by (6) implies that the extension $\bar{0}'$ of $\bar{0}$ is the zero row of matrix M'. Moreover, since the rows of M are pairwise distinct, again (6) implies that the correspondence $r_i \to r_i'$ is one to one. Thus, (i) and the first part of (ii) are established.

To prove the second part of (ii), it suffices to prove the statement for the product of two rows (since the proof of the product of any finite number of rows can be given in a similar way). To this end, in view of (6), it is enough to show that for every $i, h<m$ and $n\leqq j<n'$, $r_i r_h \in M$ implies

$$r_i r_h \in N_j^* \quad \textit{if and only if} \quad r_i \in N_j^* \quad \textit{and} \quad r_h \in N_j^*.$$

To prove the "if" part, let $r_i \in N_j^*$ and $r_h \in N_j^*$ and assume to the contrary that $r_i r_h \notin N_j^*$. However, since N_j^* is maximal w.r.t. not generating $\bar{0}$, we see that our assumption implies that $(r_i r_h) r_p \dots r_q = \bar{0}$ for some finitely many $r_p, \dots, r_q \in N_j^*$ and hence for some finitely many $r_i, r_h, r_p, \dots, r_q \in N_j^*$ which contradicts the fact that N_j^* does not generate $\bar{0}$. To prove the "only if" part, let $r_i r_h \in N_j^*$ and assume to the contrary that, say, $r_i \notin N_j^*$. Again, since N_j^* is maximal w.r.t. not generating $\bar{0}$, we see that our assumption implies that $r_i r_u \dots r_v = \bar{0}$ for some finitely many $r_u, \dots, r_v \in N_j^*$ and hence $(r_i r_h) r_u \dots r_v = \bar{0}$ for some finitely many $(r_i r_h), r_u, \dots, r_v \in N_j^*$ which contradicts the fact that N_j^* does not generate $\bar{0}$.

To prove the third part of (ii), again, it suffices to prove the statement for the sum of two rows, i.e., it is enough to show that for every $i, h<m$ and $n\leqq j<n'$, $r_i \dotplus r_h \in M$ implies:

$$(r_i \dotplus r_h) \in N_j^* \quad \textit{if and only if} \quad r_i \in N_j^* \quad \textit{or} \quad r_h \in N_j^*.$$

To prove the "if" part, let, say, $r_i \in N_j^*$ and assume to the contrary that $(r_i \dotplus r_h) \notin N_j^*$. However, since N_j^* is maximal w.r.t. not generating $\bar{0}$, we see that our assumption implies that $(r_i \dotplus r_h) r_p \dots r_q = r_i r_p \dots r_q \dotplus r_h r_p \dots r_q = \bar{0}$, and therefore, $r_h r_p \dots r_q = r_i r_p \dots r_q = \bar{0}$ for some finitely many $r_p, \dots, r_q \in N_j^*$ and hence for some finitely many $r_i, r_p, \dots, r_q \in N_j^*$ which contradicts the fact that N_j^* does not generate $\bar{0}$. To prove the "only if" part, let $(r_i \dotplus r_h) \in N_j^*$ and assume to the contrary that $r_i \notin N_j^*$ and $r_h \notin N_j^*$. Again, since N_j^* is maximal w.r.t. not generating $\bar{0}$, see that our assumption implies that $r_i r_u \dots r_v = \bar{0}$ and $r_h r_s \dots r_t = \bar{0}$ for some finitely many $r_u, \dots, r_v, r_s, \dots, r_t \in N_j^*$. Consequently, $(r_i \dotplus r_h) r_u \dots r_v r_s \dots r_t = \bar{0}$ for some finitely many $(r_i \dotplus r_h), r_u, \dots, r_v, r_s \dots, r_t \in N_j^*$ which constradicts the fact that N_j^* does not generate $\bar{0}$.

To prove the fourth part of (ii), in view of (6), it is enough to show that for every $i<m$ and $n\leqq j<n'$, $C(r_i) \in M$ implies

$$r_i \in N_j^* \quad \textit{if and only if} \quad C(r_i) \notin N_j^*.$$

Clearly, $r_i \in N_j^*$ and $C(r_i) \in N_j^*$ is impossible since N_j^* does not generate $\bar{0}$ and $r_i C(r_i) = \bar{0}$. We show that $r_i \notin N_j^*$ and $C(r_i) \notin N_j^*$ is also impossible. To this end, assume on the contrary that $r_i \notin N_j^*$ and $C(r_i) \notin N_j^*$. However, since N_j^* is maximal w.r.t. not generating $\bar{0}$, we see that our assumption implies that $r_i r_p \dots r_q = \bar{0}$ and $C(r_i) r_u \dots r_v = \bar{0}$ and therefore, $(r_i \dotplus C(r_i)) r_p \dots r_q r_u \dots r_v = \bar{0}$ which implies $r_p \dots r_q r_u \dots r_v = \bar{0}$ for some finitely many $r_p, \dots, r_q, r_u, \dots, r_v \in N_j^*$ which contradicts the fact that N_j^* does not generate $\bar{0}$.

Thus, (ii) is proved.

The proof of (iii) follows immediately from (6) and the fact that $(N_j^* - N_k^*) \neq \emptyset \neq (N_k^* - N_j^*)$ since N_j^* and N_k^* are maximal for $n\leqq j, k<n'$, $j\neq k$.

To prove (iv), let the zero row $\bar{0}$, of M' be covered by the rows of a set E' of rows of M', where

$$E' = \{r_i' \mid i \in I\}. \tag{7}$$

Thus, from (1) it follows that for every $j<n'$ we have:

(8) $$r'_{ij}=0 \quad for\ some \quad i\in I.$$

We must show that $\bar{0}'$ is already covered by finitely many rows of E'. Assume to the contrary that $\bar{0}'$ is covered by no finite number of rows of E'. But then, by the Lemma, $\bar{0}'$ is not generated by E'. From this, in view of (7), it follows that the set E of rows of M given by:

(9) $$E=\{r_i|i\in I\}$$

does not generate the zero row $\bar{0}$ of M. Indeed, if $r_u\dots r_v=\bar{0}$ for some finitely many $r_u,\dots,r_v\in E$ then by (ii), (7), (9) it would follow that $r'_u\dots r'_v=\bar{0}'$ which would contradict that $\bar{0}'$ is not generated by E'. Thus, E as given in (9) is a subset of a set of rows of M which is maximal w.r.t. not generating $\bar{0}$. Consequently, from (5) it follows that:

$$E\subseteqq N_j \quad for\ some \quad j<n \quad or \quad E\subseteqq N_j^* \quad for\ some \quad n\leqq j<n'.$$

But then, from (9), (7), (6) we derive $r'_{ij}=1$ *for every* $i\in I$ contradicting (8).

Thus, the Theorem is proved.

REMARK. Let n be a topological space. The characteristic functions of all the closed sets of n can be arranged to form the rows of an m by n dyadic matrix M which has all the properties mentioned in the Theorem. In addition, the rows of M are closed under addition and multiplication. Let us topologize the cardinal n' given in (4) such that the rows r'_i of matrix M' are the characteristic functions of the basic closed sets of n' (this is possible since the rows of M are closed under multiplication and by (ii) the correspondence $r_i\to r'_i$ preserves the products which implies that the rows of M' are also closed under multiplication). But then (ii) implies that n is dense in n' and therefore (iv) implies that n' is a compactification of n. If the topology on n is T_1 then by (iii) it follows that n' is a T_1 compactification of n which coincides with the Wallman compactification [2, p. 139] of n.

For the case where compactification is defined via the finite intersection property, see [1].

The authors thank the Referee, Á. CSÁSZÁR for valuable suggestions.

References

[1] A. ABIAN and S. SALBANY, Wallman compactifications and matrices of zeros and ones, *J. Math. Anal. and Appl.*, **69** (1979), 432—442.
[2] W. J. THRON, *Topological Structures,* Holt, Rinehart and Winston (New York, 1966).

(Received August 8, 1977; revised June 5, 1978)

DEPARTMENT OF MATHEMATICS
IOWA STATE UNIVERSITY
AMES, IOWA 50011, U.S.A.

MATHEMATICAL INSTITUTE
UNIVERSITY OF OXFORD
OXFORD, ENGLAND

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 67—84.

HERMITE—FEJÉR TYPE INTERPOLATIONS. III

By
P. VÉRTESI (Budapest)

1. Introduction

Let for a fixed $N=n+m$ $(n\geqq 1, m\geqq 0)$

(1.1) $$-\infty \leqq A \leqq x_{1,n}, x_{2,n}, \ldots, x_{n,n}; \xi_{1,m}, \xi_{2,m}, \ldots, \xi_{m,m} \leqq B \leqq \infty$$

be finite distinct points on the real axis and let

(1.2) $$d_{1,n}, d_{2,n}, \ldots, d_{n,n}$$

be real numbers. For $f\in C$ ($=f$ is continuous on $[A, B]$) we consider the uniquely determined polynomial $S(x)$ of degree $\leqq N_1=2n+m-1$

(1.3) $$\begin{cases} S(x_k)=f(x_k),\ S'(x_k)=d_k & (k=1,2,\ldots,n);\\ S(\xi_i)=f(\xi_i) & (i=1,2,\ldots,m). \end{cases}$$

(Sometimes we omit the superfluous notations.) Following P. SZÁSZ [2], we call $S(x)$ *extended Hermite—Fejér interpolatory polynomial.* We state theorems of Grünwald-type for $S(x)$ $(n\to\infty)$, further we give general convergence-divergence theorems for the special cases when some nodes in (1.1) are the roots of the Jacobi polynomial $P_n^{(\alpha,\beta)}(x)$. Equiconvergence theorems will be established, too. We intend to investigate the case $\max(|A|, |B|)=\infty$ in another paper.

2. Grünwald type theorems

2.1. Let us introduce the following notations.

(2.1) $$\omega_n(x)=C_n(x-x_1)(x-x_2)\ldots(x-x_n) \quad (C_n\neq 0),$$

(2.2) $$\Omega_m(x)=K_m(x-\xi_1)(x-\xi_2)\ldots(x-\xi_m) \quad (K_m\neq 0),$$

(2.3) $$c_k=\sum_{i=1}^{m}\frac{1}{\xi_i-x_k}-\frac{\omega''(x_k)}{\omega'(x_k)} \quad (k=1,2,\ldots,n).$$

Then $S(x)=S_N(f;d;x)$ has the form

(2.4) $$\begin{cases} S_N(f;d;x)=\sum_{k=1}^{n} f(x_k)\frac{\Omega(x)}{\Omega(x_k)}[1+c_k(x-x_k)]\,l_k^2(x)+\\ +\sum_{i=1}^{m} f(\xi_i)\frac{\omega^2(x)}{\omega^2(\xi_i)}L_i(x)+\sum_{k=1}^{n} d_k\frac{\Omega(x)}{\Omega(x_k)}(x-x_k)\,l_k^2(x) \overset{\text{def}}{=}\\ \overset{\text{def}}{=}\sum_{k=1}^{n} f(x_k)h_k(x)+\sum_{i=1}^{m} f(\xi_i)t_i(x)+\sum_{k=1}^{n} d_k\mathfrak{H}_k(x), \end{cases}$$

where, as usual, $l_k(x)=\omega(x)[\omega'(x_k)(x-x_k)]^{-1}$ $(k=1, 2, \dots, n)$ and $L_i(x)=$ $=\Omega(x)[\Omega'(\xi_i)(x-\xi_i)]^{-1}$ $(i=1, 2, \dots, m)$ (see [2], (24**)).
If $d_{k,n}=0$ $(k=1, 2, \dots, n)$ we write $S_N(f; x)$, further $S_N(f, f'; x)$ stands for the polynomial $S_N(f, d, x)$ for which $d_{k,n}=f'(x_{k,n})$ $(k=1, 2, \dots, n)$. For $m=0$ let $\sum_{i=1}^{m} \overset{\text{def}}{=} 0$ and $\Omega(x)/\Omega(x_k) \overset{\text{def}}{=} 1$; so from (2.4) we get the well known Hermite—Fejér interpolatory polynomial $H_n(f; x)$ of degree $\leqq 2n-1$. As it is well known, the functions $v_k(x) \overset{\text{def}}{=} 1+c_k(x-x_k)$ play an important role in the investigation of the expression $|H_n(f; x)-f(x)|$. We shall see that they are rather important from the point of $S_N(f; x)$, too.

2.2. Remark that $S_N(P, P'; x)\equiv P(x)$ whenever $P(x)$ is a polynomial of degree $\leqq 2n+m-1$. So we obtain with $f_1(x)=1$, $S_N(f_1, f_1'; x)\equiv 1$, i.e. by (2.4)

$$\sum_{k=1}^{n} h_k(x)+\sum_{i=1}^{m} t_i(x) \equiv 1 \quad (N=1, 2, \dots). \tag{2.5}$$

If in the *finite* $[a, b]\subseteqq[A, B]$ the nodes (1.1) form a *ϱ-normal pointsystem,* i.e. for a suitable $\varrho>0$

$$\text{(D)} \qquad \left.\begin{cases} v_k(x) \geqq \varrho > 0 & (k=1, 2, \dots, n), \\ \dfrac{\Omega(x)}{\Omega(x_k)} \geqq 0 & (k=1, 2, \dots, n), \\ L_i(x) \geqq 0 & (i=1, 2, \dots, m) \end{cases}\right\} \quad \text{if} \quad x\in[a, b],\ N=1, 2, \dots;^{1}$$

then we have the important relations

$$\begin{cases} 0 \leqq \displaystyle\sum_{k=1}^{n} \frac{\Omega(x)}{\Omega(x_k)} l_k^2(x) \leqq \frac{1}{\varrho}, \\ 0 \leqq \displaystyle\sum_{i=1}^{m} t_i(x) \leqq 1 \quad (x\in[a, b]). \end{cases} \tag{2.6}$$

From now on we suppose A and B are finite, e.g. $-A=B=1$.

2.3. We state theorems using S_N which are analogous to theorems proved for H_n by G. GRÜNWALD [3]. In fact the proofs are similar, too, so sometimes we sketch or omit them.

Let $\|g\|_{[a,b]}=\max_{a\leqq x\leqq b} |g(x)|$, $\|g\|=\|g\|_{[-1,1]}$. We state

THEOREM 2.1. *Supposing that the nodes form a ϱ-normal system in* $[a, b]$,

$$\|S_N(f, f'; x)-f(x)\|_{[a,b]} = (6+2/\varrho)E_{N_1}(f'), \tag{2.7}$$

whenever $f, f'\in C$.

(Here $E_n(f; [a, b])=\min_{P_n} \{\|f-P_n\|_{[a,b]}\}$, P_n are polynomials of degree $\leqq n$; $E_n(f; [-1, 1])=E_n(f)$.)

[1] This definition is a generalization of the usual one (see further 2.4.1).

Dropping the ϱ-normality and $f' \in C$ we can often use the following convergence theorem (see further 2.4.3).

THEOREM 2.2. *Let us suppose that for each* N

$$\sum_{k=1}^{n} |h_k(x)| + \sum_{i=1}^{m} |t_i(x)| = O(1) \quad \textit{uniformly in} \quad x \in [a, b], \tag{2.8}$$

moreover, with $M_n = \max\left(1, \max_{1 \le k \le n} |d_{k,n}|\right)$

$$\lim_{N \to \infty} M_n \sum_{k=1}^{n} |\mathfrak{H}_k(x)| = 0 \quad \textit{uniformly in} \quad x \in [a, b], \quad \textit{too.} \tag{2.9}$$

Then

$$\lim_{N \to \infty} \|S_N(f, d, x) - f(x)\|_{[a,b]} = 0 \quad \textit{if} \quad f \in C.$$

This theorem will be applied in the proof of the following assertion which generalizes Grünwald's important statement proved for H_n (see further 2.4.4).

THEOREM 2.3. *Let us suppose that the system of nodes is ϱ-normal in* $[a, b]$, *moreover with certain fixed* $\varepsilon > 0$

$$M_n = O(N_1^{\varrho - \varepsilon}) \quad (N = 1, 2, \ldots). \tag{2.10}$$

Then

$$\lim_{N \to \infty} \|S_N(f, d; x) - f(x)\|_{[a,b]} = 0 \tag{2.11}$$

whenever $f \in C$.

2.4. REMARKS **2.4.1.** Theorems 2.1—2.3 for H_n and $[a, b] = [-1, 1]$ were proved in [3]. For $m = 2$, $\xi_1 = 1$, and $\xi_2 = -1$ (the so called quasi-Hermite—Fejér interpolation), Theorem 2.3 was proved with $[a, b] = [-1, 1]$ and $M_n = O(1)$ by J. SÁNTHA [12], using [3], 4. §. In these cases (D) has the form $v_{k,n}(x) \ge \varrho > 0$ (for each k and n) in $[-1, 1]$, which was the original definition ([3], [10]).

2.4.2. If in (D) we know only $v_k(x) \ge 0$ (without any change in the further requirements), we say that the pointsystem is *normal* in $[a, b]$. We can establish *pointwise convergence theorems if* $x \in [a, b]$ and *uniform convergence in* $[a + \varepsilon, b - \varepsilon]$. We omit the details (see the corresponding part of [3]).

2.4.3. In theorem 2.2, the condition (2.9) can be replaced by

$$\lim_{n \to \infty} \left\| \sum_{k=1}^{n} (1 + |d_k|) \frac{|\mathfrak{H}_k(x)|}{\sqrt{1 - x_k^2}} \right\|_{[a,b]} = 0 \quad (|x_k| < 1), \tag{2.9*}$$

involving $w(f; t)$ in the estimation of the error (see (2.13)).

Indeed, there exists a polynomial $P_Q(x)$ of degree $\le Q$ such that

$$\begin{cases} \|f(x) - P_Q(x)\| = E_Q(f), \\ |P_Q'(x)| = \dfrac{O(Q)}{\sqrt{1 - x^2}} w\left(f; \dfrac{1}{Q}\right) \quad \text{if} \quad |x| < 1, \end{cases} \tag{2.12}$$

where w is the modulus of continuity of f (S. B. STECKIN [21]).

I.e., for suitable $Q=Q(N)$ we get for $N\to\infty$, as in 4.2,

$$\|S_N(f;\, d;\, x)-f(x)\|_{[a,b]} = O(1)\, w\left(f;\, \frac{1}{Q}\right)\left[1+Q\sum_{k=1}^{n}(1+|d_k|)\frac{|\mathfrak{H}_k(x)|}{\sqrt{1-x_k^2}}\right] = o(1). \tag{2.13}$$

For $|x_k|=1$, we use

$$|P_Q'(x)| = O(Q^2)\, w\left(f;\, \frac{1}{Q}\right) \tag{2.14}$$

(see [21]). If, e.g. $x_1=1$, the first term of the sum (2.9*) will be $(1+|d_1|)\,|\mathfrak{H}_1(x)|$. For the process H_n similar arguments where applied by M. SALLAY [22].

2.4.4. We can get a convergence order for the ϱ-normal system. Indeed, we can prove that if the nodes form a ϱ-normal system in $[a, b]$ then

$$\sum_{k=1}^{n} |\mathfrak{H}_k(x)| = O(N^{-\varrho+\varepsilon}) \quad (\varepsilon > 0 \text{ is arbitrary}) \tag{2.15}$$

(see (4.5)). So we get by (2.14), as in 4.2 as follows.

Supposing that the system of nodes is ϱ-normal in $[a, b]$, we have with arbitrary fixed $\varepsilon>0$

$$\|S_n(f;\, x)-f(x)\|_{[a,b]} = O(1)\, w\left(f;\, \frac{1}{N^{(\varrho-\varepsilon)/2}}\right). \tag{2.16}$$

One can get estimation using $S_N(f;\, d,\, x)$.

2.4.5. For a ϱ-normal in $[a, b]$ system we can obtain

$$\lim_{n\to\infty}\left\|\sum_{i=1}^{m}|\xi_i - x|\,\frac{\omega^2(x)}{\omega^2(\xi_i)}\,L_i(x)\right\|_{[a,b]} = 0,$$

if we use the relation (4.4).

3. Some special extended Hermite—Fejér interpolations

3.1. Let

$$-1 \leqq x_{n,n}^{(\alpha,\beta)} < x_{n-1,n}^{(\alpha,\beta)} < \ldots < x_{2,n}^{(\alpha,\beta)} < x_{1,n}^{(\alpha,\beta)} \leqq 1 \tag{3.1}$$

be the roots of the Jacobi polynomial $P_n^{(\alpha,\beta)}(x)$ $(\alpha, \beta \geqq -1)$ of degree n. As we know, if $\alpha>-1$ then $x_{1,n}<1$, and similarly, $x_{n,n}>-1$ whenever $\beta>-1$. We often use the normalization

$$P_n^{(\alpha,\beta)}(1) = \binom{n+\alpha}{\alpha}, \quad P_n^{(\alpha,\beta)}(x) = (-1)^n P_n^{(\beta,\alpha)}(-x) \tag{3.2}$$

(see [1], chapter 4). Remark that in the sequel a certain knowledge of the Jacobi polynomials is supposed. Namely, we shall apply (mainly in the proofs) the formulae and notations (4.1.1), (4.1.3), (4.2.1), (4.21.7), (4.22.2), (7.32.5), 8.9 and 14.5 from [1], sometimes without any further references. Further, we use Lemmas 1 and 2 from [17] which state that

$$0 < c_1(\alpha, \beta) \leqq n(\vartheta_{k+1}^{(\alpha,\beta)} - \vartheta_k^{(\alpha,\beta)}) \leqq c_2(\alpha, \beta) \quad (k = 0, 1, \dots, n),$$

where

$$x_k^{(\alpha,\beta)} = \cos \vartheta_k^{(\alpha,\beta)},\ x_0^{(\alpha,\beta)} \equiv 1,\ x_{n+1}^{(\alpha,\beta)} \equiv -1 \quad (\alpha, \beta > -1).$$

Using the nodes (3.1) and the points ± 1, the following special $S_N(f; x)$ processes have been investigated.

Notation, name	Informations on the nodes	Some references
$H_n^{(\alpha,\beta)}$ Hermite—Fejér	$m = 0$	[1], [3], [4], [5], [6], [7], [8], [9], [15], [16], [18], [19], [20], [22], [24], [26]
$R_{n+1}^{(\alpha,\beta)}$, $\alpha > -1$	$m = 1$, $\xi_1 = 1$	[7], [9], [14]
$Q_{n+2}^{(\alpha,\beta)}$, $\alpha, \beta > -1$ quasi-Hermite—Fejér	$m = 2$, $\xi_1 = 1$, $\xi_2 = -1$	[7], [9], [10], [11], [12], [13], [16], [19], [20], [25], [28], [29], [30]
$U_{n+1}^{(\alpha,\beta)}$, $\alpha > -1$	$m = 0$ $x_{0,n} = 1$	[7], [9], [15], [27], [30]
$V_{n+2}^{(\alpha,\beta)}$, $\alpha, \beta > -1$	$m = 1$, $\xi_1 = 1$ $x_{0,n} = 1$	[7], [9], [14]
$W_{n+2}^{(\alpha,\beta)}$, $\alpha, \beta > -1$	$m = 0$ $x_{0,n} = 1$, $x_{n+1,n} = -1$	[6], [7], [9], [16]

(Here, e.g., $V_{n+2}^{(\alpha,\beta)}$ is based on the nodes $x_{0,n}, x_{1,n}^{(\alpha,\beta)}, \dots, x_{n,n}^{(\alpha,\beta)}$ and ξ_1. The value of the derivative of $V_{n+2}^{(\alpha,\beta)}$ is not prescribed at ξ_1.)

3.2. The convergence (or divergence) results in $[-1, 1]$ generally were established only for the roots x_k of the Tchebyshev polynomials of first or second kind. Now we intend to investigate *all Jacobi matrices* and which is more important, we establish close connection among the above processes; we shall consider *all the processes together* to clear up the regularity of their convergence-divergence behaviour.

We say that a pointsystem is ϱ-normal for H_n (or R_n, Q_n, ...) in $[a, b]$ if (D) is true for the corresponding process. We state the following

THEOREM 3.1. *The fundamental nodes of the extended Hermite—Fejer interpolation form a ϱ-normal pointsystem in* $[-1, 1]$ *for the following special processes,* *parameters and values of* ϱ:

$$(3.3)\quad \begin{cases} H_n^{(\alpha,\beta)} & -1\leqq\alpha,\beta<0 & \min(-\alpha,-\beta),\\ R_{n+1}^{(\alpha,\beta)} & 0\leqq\alpha<1;\ -1\leqq\beta<0 & \min[-(\alpha-1),-\beta],\\ Q_{n+2}^{(\alpha,\beta)} & 0\leqq\alpha,\beta<1 & \min[-(\alpha-1),-(\beta-1)],\\ U_{n+1}^{(\alpha,\beta)} & 1\leqq\alpha<2;\ -1\leqq\beta<0 & \min[-(\alpha-2),-\beta];\\ V_{n+2}^{(\alpha,\beta)} & 1\leqq\alpha<2;\ 0\leqq\beta<1 & \min[-(\alpha-2),-(\beta-1)],\\ W_{n+2}^{(\alpha,\beta)} & 1\leqq\alpha,\beta<2 & \min[-(\alpha-2),-(\beta-2)]. \end{cases}$$

Moreover, by Theorem 2.3

$$(3.4)\qquad \lim_{n\to\infty}\|S_N(f;\,x)-f(x)\|=0 \quad \textit{if}\quad f\in C,$$

where S_N *should be replaced by* $H_n^{(\alpha,\beta)}, R_{n+1}^{(\alpha,\beta)}, \dots, W_{n+2}^{(\alpha,\beta)}$ *having the parameters specified in* (3.3).

3.3. In the sequel denote by $S_N(\alpha,\beta)(f;x)$ $(\alpha,\beta>-3)$ *any meaningful* extended Hermite—Fejér operator from the following ones:

$$H_n^{(\alpha,\beta)}(f;\,x),\ R_{n+1}^{(\alpha+1,\beta)}(f;\,x),\ Q_{n+2}^{(\alpha+1,\beta+1)}(f;\,x),\ U_{n+1}^{(\alpha+2,\beta)}(f;\,x),$$
$$V_{n+2}^{(\alpha+2,\beta+1)}(f;\,x)\quad\text{and}\quad W_{n+2}^{(\alpha+2,\beta+2)}(f;\,x).$$

By this notation Theorem 3.1 states: *The fundamental nodes of* $S_N(\alpha,\beta)(f;x)$ *form a* $\varrho=\min(-\alpha,-\beta)$*-normal system in* $[-1,1]$ *whenever* $-1\leqq\alpha,\beta<0$. *Further, for these* α*'s and* β*'s.*

$$(3.5)\qquad \lim_{N\to\infty}\|S_N(\alpha,\beta)(f;\,x)-f(x)\|=0 \quad \textit{for any}\quad f\in C.$$

Let

$$C(w)=\{f(x);\ f\in C\ \text{ and }\ w(f;\,t)\leqq a(f)w(t)\}$$

where $w(f;t)$ is the modulus of continuity of $f(x)$ and $w(t)$ is any modulus of continuity. If $a(f)\leqq M$ we shall write $C_M(w)$.

With these notations we state the following equiconvergence theorem.

THEOREM 3.2. *If* $\alpha,\beta\geqq-1$ *then we have*

$$(3.6)\qquad \frac{\sup\limits_{f\in C_1(w)}\|S_N(\alpha,\beta)(f;\,x)-f(x)\|}{\sup\limits_{f\in C_1(w)}\|H_n^{(\alpha,\beta)}(f;\,x)-f(x)\|}\sim 1.$$

The statement will be generalized, in a certain sense, in Section 3.5. (Remark that Theorem 3.2 becomes an "equidivergence" relation for $\min(\alpha,\beta)\geqq0$; as for the meaning of the symbol "$\sim$", see [1], 1.1.)

3.4. If $\min(\alpha,\beta)<-1$, $S_N(\alpha,\beta)(f;x)$ is not a positive operator anymore. Of course, e.g. $H_n^{(\alpha,\beta)}$ has no meaning, but for the meaningful processes we settle all the cases. First we prove a convergence statement which includes (3.5), too.

THEOREM 3.3. *Supposing* $-1.5 \leqq \alpha,\ \beta < 0$ *and* $|\alpha-\beta| \leqq 1$, *we have*

$$\lim_{n\to\infty} \|S_N(\alpha,\beta)(f;\ x)-f(x)\| = 0 \quad \textit{whenever} \quad f\in C. \tag{3.7}$$

(Of course, we suppose $S_N(\alpha,\beta)(f;\ x)$ has meaning. E.g., if $S_N(\alpha,\beta)=U_{n+1}^{(\alpha+2,\beta)}$ we must require $\beta \geqq -1$.)

This theorem is the best possible in the following sense

THEOREM 3.4. *Suppose*

$$\begin{cases} \max(\alpha,\beta) \geqq 0, & \textit{or} \\ \min(\alpha,\beta) < -1.5, & \textit{or} \\ |\alpha-\beta| > 1. \end{cases} \tag{3.8}$$

Then for certain $f\in C$, (3.7) *is not valid.*

3.5. Using much stronger tools than for Theorem 3.2, we can prove the following general equiconvergence statement.

THEOREM 3.5. *Let* $\alpha,\beta > -3$ *and* $\{q_i\}$ *be an arbitrary sequence. Then one can choose a subsequence* $\{r_i\}\subset\{q_i\}$ *such that*

$$\frac{\sup\limits_{f\in C_1(w)} \|S_N(\alpha,\beta)(f;\ x)-f(x)\|}{\sup\limits_{f\in C_1(w)} \|W_{n+2}^{(\alpha+2,\beta+2)}(f;\ x)-f(x)\|} \sim 1 \quad (n = r_1, r_2, \ldots). \tag{3.9}$$

3.6. REMARKS. **3.6.1.** It is a natural question to investigate the order of convergence. This has been done for different processes (see, e.g., [8], [13], [18]—[20], [22] and [26]—[30]). By the method used in [18] we get for $\alpha,\beta \geqq -1$.

$$|S_N(\alpha,\beta)(f;\ x)-f(x)| = O(1) \sum_{i=1}^{n} \left[w\left(f; \frac{i\sqrt{1-x^2}}{n}\right) + w\left(f; \frac{i^2|x|}{n^2}\right)\right] i^{2\gamma-1} \quad (x\in[-1,1]$$

where $\gamma=\max(\alpha,\beta,-0.5)$ and the sign O depends only on α and β.

Near the endpoints one can prove better estimations if we use the method applied in [19].

Finally we mention the following uniform estimation:

$$\|S_N(\alpha,\beta)(f;\ x)-f(x)\|_{[a,b]} = O(1) \sum_{i=1}^{n} w\left(f; \frac{i}{n}\right)\frac{1}{i^2} \quad ([a,b]\subset(-1,1),\ \alpha,\beta \geqq -1)$$

(see [19]).

3.6.2. Now we list some results proved in previous papers. (3.4) was established for $H_n^{(\alpha,\beta)}$ $(-1 \leqq \alpha,\beta < 0)$, $R_{n+1}^{(0.5,-0.5)}$, $Q_{n+2}^{(\alpha,\beta)}$ $(0<\alpha,\beta<1)$ and $V_{n+1}^{(1,0)}$ in [3], [2], [12] and [14], respectively.

The estimation (3.7) was proved for $\alpha=\beta=-0.5$ and $\alpha=\beta=0.5$ for different $S_N^{(\alpha,\beta)}$ ([7] and [9]). The processes $U_{n+1}^{(\alpha,\beta)}$ and $W_{n+2}^{(\alpha,\beta)}$ were mainly used to prove theorems of divergence-type ([6], [9] and [15]).

Theorem 3.2 for Q_{n+2} was proved in [19].

3.6.3. Theorems 3.3 and 3.4 serve answers for the convergence-divergence behaviour of $Q_{n+2}^{(\alpha,\beta)}$ (see P. TURÁN [23], Problem XXVII).

3.6.4. We have seen that the ϱ-normality plays an important role in the convergence behaviour of $S_N(f, d; x)$ (see, e.g. Theorem 2.3). *P. Turán raised the problem whether there exists such a system of nodes for which the conjugate-points* $\{s_{k,n}\}$ *are dense in* $[-1, 1]$ *but* $\|H_n(f; x)-f(x)\| \to 0$ *whenever* $f\in C$. J. BALÁZS [24] and P. VÉRTESI [20] gave positive solutions for this problem using the roots of the orthogonal polynomials corresponding to the weight $p(x)=|x|^{2\alpha+1}(1-x^2)^\beta$ $(\alpha, \beta > -1)$. Now we give answer for a generalization of the question raised above using the roots of the Jacobi polynomials.

In accordance with the usual definition (see [5]), the root $s_{k,n}$ of the linear function $v_{k,n}(S; x)$ will be called *conjugate point to* $x_{k,n}$ $(k=1, 2, \dots, n;\ n=1, 2, \dots)$. First we state

3.6.4.1. *Let* $\alpha=-1-\varepsilon$ *and* $\beta=\alpha+2\delta$ $(0<\varepsilon;\ 0\leqq\delta<0.5)$. *Then the conjugate points* $\{s_{k,n}\}$, *corresponding to the process* $\{S_N(\alpha, \beta)(f; x)\}$ *are dense in the interval* $[-1, 1]$.

Indeed, it is easy to see that

$$v_{k,n}(1) = (1+\beta)\frac{1-x_k}{1+x_k}-\alpha \quad (|x_k| < 1). \tag{3.10}$$

So $v_k(1)\to-\alpha=1+\varepsilon$ if $n\to\infty$ and $2\leqq k\leqq M$. Similarly $v_k(1)\to1+\beta-\alpha= =1+2\delta<2$ if $n\to\infty$ and $0<n/2-k\leqq M$. By these and $v_k(x_k)=1$ one can obtain $s_{k,n}\to1$ $(n\to\infty,\ 2\leqq k\leqq M)$ and $s_{k,n}<-1$ $(0<n/2-k\leqq M,\ n\geqq n_0)$, from where the statement follows (see further [20]).

Let us consider two special cases. By 3.6.4.1 and Teorem 3.3 we get

3.6.4.1.1. *Let* $0.5\leqq\alpha<1$. *Then the conjugate points* $\{s_{k,n}\}$ *corresponding to* $W_{n+2}^{(\alpha,\alpha)}$ *are dense in* $[-1, 1]$, *and at the same time,* $\lim_{n\to\infty}\|W_{n+2}^{(\alpha,\alpha)}(f; x)-f(x)\|=0$ *if* $f\in C$. *The analogous statement holds for* $Q_{n+2}^{(\alpha,\alpha)}(f; x)$ *if* $-0.5\leqq\alpha<0$.

Remarking that $W_{n+2}^{(\alpha,\alpha)}$ is an Hermite—Fejér interpolatory process, we gave answer for the Turán's problem.

3.6.4.2. We can investigate the conjugate points for the remaining α's and β's, too (see, e.g. [5], [20] and [24]).

3.7. The investigation of $\overline{R}_{n+1}^{(\alpha,\beta)}$ (where $\beta>-1$, $m=1$, $\xi_1=-1$) can be reduced to $R_{n+1}^{(\beta,\alpha)}$ (see (3.2)). Similar remark holds for $\overline{U}_{n+1}^{(\alpha,\beta)}$ and $\overline{V}_{n+2}^{(\alpha,\beta)}$, too.

3.8. By usual methods one can investigate the $S_N(\alpha, \beta)(f; d; x)$ processes, too.

4. Proofs

4.1. PROOF OF THEOREM 2.1. By (2.6) we obtain

$$\sum_{k=1}^{n} |\mathfrak{H}_k(x)| \leqq \frac{2}{\varrho} \quad (x\in[a, b]). \tag{4.1}$$

Let $P_r'(x)$ be the polynomial such that $|f'(x)-P_r'(f; x)|\leqq E_r(f')$. Let $P_r(x)=$

$=f(-1)+\int_{-1}^{x} P_r'(t)\,dt$. Obviously $|f(x)-P_r(x)|\leqq 2E_r(f')$. So we obtain with $S_N(P_{N_1}, P_{N_1}'; x)\equiv P_{N_1}(x)$, (4.1) and (2.1)

$$|S_N(f,f';x)-f(x)| = |S_N(f-P_{N_1},f';x)+P_{N_1}(x)-f(x)| \leqq$$

$$\leqq \sum_{k=1}^{n} |f(x_k)-P_{N_1}(x_k)|\, h_k(x)+\sum_{i=1}^{m} |f(\xi_i)-P_{N_1}(\xi_i)|\, t_i(x)+$$

$$+\sum_{k=1}^{n} |f'(x_k)-P_{N_1}'(x_k)|\, |\mathfrak{H}_k(x)|+2E_{N_1}(f') = (6+2/\varrho)\big(E_{N_1}(f')\big)$$

in $[a, b]$, which was stated.

4.2. Proof of Theorem 2.2. Using a suitable $P_r(f; x)$ we have, by (2.8) and (2.9) as above

$$|S_N(f, d;\ x)-f(x)|\leqq \sum_{k=1}^{n} |f(x_k)-P(x_k)|\, |h_k(x)|+\sum_{i=1}^{m} |f(\xi_i)-P(\xi_i)|\, |t_i(x)|+$$

$$+\sum_{k=1}^{n} \big|M_n+|P'(x_k)|\big|\, |\mathfrak{H}_k(x)|+|P(x)-f(x)|<\varepsilon.$$

4.3. Proof of Theorem 2.3. **4.3.1.** Because of (D) and (2.5), (2.8) is valid. So we have to prove that from (D) we obtain (2.9), too (see Theorem 2.2).

4.3.2. In the following proof we use ideas analogous to [3], 5. §. Let α be any fixed point in $[a, a_1]$ (where $a_1=(b+a)/2$). Define the function $g(x)\in C$ as follows:

$$g(x)=\begin{cases}0 & \text{if} \quad -1\leqq x\leqq \alpha,\\ (x-\alpha)^{\varrho_1} & \text{if} \quad \alpha\leqq x\leqq 1 \quad \text{where} \quad 0<\varrho_1<\varrho\leqq 1.\end{cases} \tag{4.2}$$

Using $g(x)$, we shall verify for any fixed $0<\varrho_2<\varrho_1$

$$\sum_{\alpha\leqq x_k} \frac{\Omega(\alpha)}{\Omega(x_k)}(x_k-\alpha)^{\varrho_1} l_k^2(\alpha)\leqq \frac{c}{N_1^{\varrho_2}} \quad \text{if} \quad \alpha\in[a, a_1]. \tag{4.3}$$

If $\alpha=x_k$ for a certain k, then obviously $\sum_{\alpha\leqq x_k}=0$. So we consider N's such that $\alpha\neq x_k$. Remark that for these N's we can form $S_N(g, g'; x)$. By (4.2) we obtain

$$S_N(g, g';\ \alpha) = \sum_{\alpha<x_k} (x_k-\alpha)^{\varrho_1}\frac{\Omega(\alpha)}{\Omega(x_k)}[v_k(\alpha)-\varrho_1]\, l_k^2(\alpha)+\sum_{\alpha\leqq\xi_i} (\xi_i-\alpha)^{\varrho_1}\frac{\omega^2(\alpha)}{\omega^2(\xi_i)} L_i(\alpha)\geqq 0. \tag{4.4}$$

4.3.3. To estimate $S_N(g, g';\alpha)$ in another way we quote the following

LEMMA 4.1. *For the function* $g(x)$, *there exist polynomials* $P_n(x)$ *of degree* $\leqq n$ *such that in* $-1 \leqq x \leqq 1$

$$|g(x)-P_n(x)| \leqq c\frac{\ln n}{n^{\varrho_1}},$$

$$|(x-\alpha)(g'(x)-P_n'(x))| \leqq c\frac{\ln n}{n^{\varrho_1}} \quad (x \neq \alpha,\ n = 2, 3, \ldots).$$

(See [3], 6.§.)

So by (2.6) and the lemma

$$|S_N(g, g'; \alpha)-P_{N_1}(\alpha)| \leqq \sum_{k=1}^{n} |g(x_k)-P_{N_1}(x_k)|\, h_k(\alpha)+$$

$$+\sum_{i=1}^{m} |g(\xi_i)-P_{N_1}(\xi_i)|\, t_i(\alpha)+\sum_{k=1}^{n} |g'(x_k)-P'_{N_1}(x_k)|\, |x_k-\alpha| \frac{\Omega(\alpha)}{\Omega(x_k)}\, l_k^2(\alpha) \leqq$$

$$\leqq c\frac{\ln N}{N_1^{\varrho_1}} \leqq \frac{c}{N_1^{\varrho_2}}$$

with $0<\varrho_2<\varrho_1$. Again by the lemma, $|P_{N_1}(\alpha)|=O(N_1^{-\varrho_2})$, i.e. by (4.4)

$$0 \leqq S_N(g, g'; \alpha) \leqq \frac{c}{N_1^{\varrho_2}}. \tag{4.5}$$

By (2.6), (4.4), (4.5) and $v_k(\alpha)-\varrho_1 \geqq \varrho-\varrho_1>0$ we obtain (4.3).

4.3.4. Using $(x_k-\alpha)^{\varrho_1}=2^{\varrho_1}\left(\frac{x_k-\alpha}{2}\right)^{\varrho_1} \geqq 2^{\varrho_1}\frac{x_k-\alpha}{2}$ and similar arguments for $\alpha \in [a_1, b]$ we obtain

$$\sum_{k=1}^{n} |\alpha-x_k| \frac{\Omega(\alpha)}{\Omega(x_k)}\, l_k^2(\alpha) = O(N_1^{-\varrho_2}) \quad \text{if} \quad \alpha \in [a, b] \tag{4.5}$$

uniformly in α, which, by (2.10), gives (2.9), the desired result.

4.4. PROOF OF THEOREM 3.1. **4.4.1.** If $m=1$, $\xi_1=\pm 1$, we get by (2.2)

$$\Omega(x)[\Omega(x_k)]^{-1} = (1\mp x)(1\mp x_k)^{-1} \geqq 0$$

and $L_1(x)\equiv 1$. For $m=2$, $\xi_1=1$, $\xi_2=-1$ we obtain

$$\Omega(x)[\Omega(x_k)]^{-1} = (1-x^2)(1-x_k^2)^{-1} \geqq 0$$

and

$$L_1(x) = \frac{1+x}{2} \geqq 0, \quad L_2(x) = \frac{1-x}{2} \geqq 0.$$

So from (D) we have to investigate only $v_k(x)$.

4.4.2. As it is well-known for $H_n^{(\alpha,\beta)}$

$$(4.6)\quad \begin{cases} v_k(H;\ x) = 1-\dfrac{(\alpha+\beta+2)x_k-\beta+\alpha}{1-x_k^2}(x-x_k) & (|x_k|<1;\ \alpha,\beta\geqq -1), \\ v_1(H;\ x) = 1-\dfrac{(n+\beta-1)(n-1)}{2}(x-1) & (\alpha=-1), \\ v_n(H;\ x) = 1+\dfrac{(n+\alpha-1)(n-1)}{2}(x+1) & (\beta=-1), \end{cases}$$

further, using the values $v_k(\pm 1)$, we have

$$(4.7)\quad v_k(H;\ x)\geqq \min(-\alpha,\ -\beta)\quad (k=1,2,\dots,n)$$

(see [1], 14.5).

4.4.3. Considering $R_{n+1}^{(\alpha,\beta)}$ we get by (2.3) and (4.6)

$$c_k = \frac{1}{1-x_k}-\frac{(\alpha+\beta+2)x_k-\beta+\alpha}{1-x_k^2} = -\frac{(\alpha+\beta+1)x_k-\beta+\alpha-1}{1-x_k^2}\quad (|x_k|<1),$$

$$c_n = \frac{1}{2}+\frac{(n+\alpha-1)(n-1)}{2} = \frac{(n+\alpha-1)(n-1)+1}{2}\quad (\beta=-1)$$

from where we obtain

$$v_k(R;\ x) = 1-\frac{(\alpha+\beta+1)x_k-\beta+\alpha-1}{1-x_k^2}(x-x_k)\quad (|x_k|<1,\ \beta\geqq -1),$$

$$v_n(R;\ x) = 1+\frac{(n+\alpha-1)(n-1)+1}{2}(1+x)\quad (\beta=-1).$$

Using these, it is easy to get the desired result (see [1], 14.5).

4.4.4. For $Q_{n+2}^{(\alpha,\beta)}$

$$c_k = \frac{1}{1-x_k}-\frac{1}{1+x_k}-\frac{(\alpha+\beta+2)x_k-\beta+\alpha}{1-x_k^2} = -\frac{(\alpha+\beta)x_k-\beta+\alpha}{1-x_k^2}\quad (k=1,2,\dots,n).$$

The remaining part is the same as above.

4.4.5. Now let us see $U_{n+1}^{(\alpha,\beta)}$. Obviously

$$(4.8)\quad \omega_{n+1}(U^{(\alpha,\beta)};\ x) = (1-x)P_n^{(\alpha,\beta)}(x) = (1-x)\omega_n(H^{(\alpha,\beta)};\ x),$$

which means

$$(4.9)\quad \omega'(U;\ x) = -P_n(x)+(1-x)P_n'(x);\quad \omega''(U;\ x) = -2P_n'(x)+(1-x)P_n''(x).$$

So by (2.3), (4.6), (4.8) and (4.9)

$$c_k = -\frac{\omega''(U;\, x_k)}{\omega'(U;\, x_k)} = \frac{2}{1-x_k} - \frac{(\alpha+\beta+2)x_k-\beta+\alpha}{1-x_k^2} = -\frac{(\alpha+\beta)x_k-\beta+\alpha-2}{1-x_k^2} \quad (|x_k|<1),$$

$$c_0 = -\frac{\omega''(U;\, 1)}{\omega'(U;\, 1)} = -\frac{2P_n'(1)}{P_n(1)} = -\frac{n(n+\alpha+\beta+1)}{1+\alpha},$$

$$c_n = -\frac{\omega''(U;\, -1)}{\omega'(U;\, -1)} = 1+\frac{(n+\alpha-1)(n-1)}{2} \quad (\beta=-1),$$

where at c_0 we used (3.2) and [1] (4.3.1), which gives

$$P_n'^{(\alpha,\beta)}(1) = \frac{n+\alpha+\beta+1}{2}\binom{n+\alpha}{n-1}. \tag{4.10}$$

4.4.6. For $V_{n+2}^{(\alpha,\beta)}$, using 4.4.5 and (2.3), we have

$$c_k = -\frac{1}{1+x_k} - \frac{\omega''(U;\, x_k)}{\omega'(U;\, x_k)} = -\frac{(\alpha+\beta-1)x_k-\beta+\alpha-1}{1-x_k^2} \quad (k=1, 2, \ldots, n),$$

$$c_0 = -\frac{1}{2} - \frac{\omega''(U;\, 1)}{\omega'(U;\, 1)} = -\left[\frac{n(n+\alpha+\beta+1)}{1+\alpha}+\frac{1}{2}\right].$$

4.4.7. Finally for $W_{n+2}^{(\alpha,\beta)}$

$$\omega_{n+2}(W;\, x) = (1-x^2)P_n^{(\alpha,\beta)}(x),$$

i.e.

$$\omega'(W;\, x) = -2xP_n(x)+(1-x^2)P_n'(x),$$

$$\omega''(W;\, x) = -2P_n(x)-4xP_n'(x)+(1-x^2)P_n''(x),$$

so

$$c_k = -\frac{\omega''(W;\, x_k)}{\omega'(W;\, x_k)} = \frac{4x_k}{1-x_k^2} - \frac{(\alpha+\beta+2)x_k-\beta+\alpha}{1-x_k^2} = -\frac{(\alpha+\beta-2)x_k-\beta+\alpha}{1-x_k^2}$$

$$(k=1, 2, \ldots, n),$$

$$c_0 = -\frac{\omega''(W;\, 1)}{\omega'(W;\, 1)} = -\frac{2P_n(1)+4P_n'(1)}{2P_n(1)} = -\left[\frac{n(n+\alpha+\beta+1)}{1+\alpha}+1\right],$$

$$c_{n+1} = -\frac{\omega''(W;\, -1)}{\omega'(W;\, -1)} = \frac{n(n+\alpha+\beta+1)}{1+\beta}+1.$$

4.4.8. By 4.4.1—4.4.7 we proved the ϱ-normality. For proving (3.4) we have to apply Theorem 2.3.

4.5 PROOF OF THEOREM 3.2. **4.5.1.** As we mentioned we have proved the statement if $S_N(\alpha, \beta) = Q_{n+2}^{(\alpha+1, \beta+1)}$ ([19]). Using similar considerations we sketch the proof for another characteristic case, e.g. when $S_N(\alpha, \beta) = V_{n+2}^{(\alpha+2, \beta+1)}$.

4.5.2. By (2.4), 4.4.4, (4.8) and 4.4.6 we get

(4.11)

$$V_{n+2}^{(\alpha+2,\beta+1)}(f;\,x)-f(x)=\sum_{k=1}^{n}[f(x_k)-f(x)]\frac{1+x}{1+x_k}\left[1-\frac{(\alpha+\beta+2)x_k-\beta+\alpha}{1-x_k^2}(x-x_k)\right]\cdot$$
$$\cdot\left[\frac{1-x}{1-x_k}\frac{P_n^{(\alpha+2,\beta+1)}(x)}{P_n'^{(\alpha+2,\beta+1)}(x_k)(x-x_k)}\right]^2+[f(1)-f(x)]\frac{1+x}{2}\cdot$$
$$\cdot\left\{1+\left[\frac{n(n+\alpha+\beta+4)}{\alpha+3}+\frac{1}{2}\right](1-x)\right\}\left[\frac{P_n^{(\alpha+2,\beta+1)}(x)}{P_n^{(\alpha+2,\beta+1)}(1)}\right]^2+$$
$$+[f(-1)-f(x)]\left[\frac{(1-x)P_n^{(\alpha+2,\beta+1)}(x)}{2P_n^{(\alpha+2,\beta+1)}(-1)}\right]^2=\sum_{k=1}^{n}\ldots+\sum_{k=0}\ldots+\sum_{k=n+1}\ldots=$$
$$=I_{1,n}(x)+I_{2,n}(x)+I_{3,n}(x).$$

4.5.3. First let $-1\leqq\alpha,\,\beta\leqq 0$. One can see that now $V_{n+2}^{(\alpha+2,\beta+1)}(f;\,x)$ is a positive operator so

$$\sup_{f\in C_1(w)}\|V_{n+2}^{(\alpha+2,\beta+1)}(f;\,x)-f(x)\|=\sup_{x\in[-1,1]}V_{n+2}^{(\alpha+2,\beta+1)}\big(w(\cdot\,;\,x);\,x\big)\stackrel{\text{def}}{=\!=}F_{n+2}^{(\alpha+2,\beta+1)}(w),$$

where $w(t;x)=w(|t-x|)$ (for the details see [19], 4.8.) By (4.11) it is easy to see that for certain $c>0$

(4.12) $$V_{n+2}^{(\alpha+2,\beta+1)}\big(w(\cdot\,;\,y);\,y\big)=O(1)\sup_{x\in[a_n,b_n]}V_{n+2}^{(\alpha+2,\beta+1)}\big(w(\cdot\,;\,x);\,x\big),$$

where $|y|\geqq 1-cn^{-2}$ and $[a_n,b_n]=[-1+cn^{-2};\,1-cn^{-2}]$; so one can choose the sequence $p_n=\cos\psi_n$ such that for the fixed $c>0$, $cn^{-1}\leqq\psi_n\leqq\pi-cn^{-1}$, moreover

$$V_{n+2}^{(\alpha+2,\beta+1)}\big(w(\cdot\,;\,p_n);\,p_n\big)\sim F_{n+2}.$$

4.5.4. One can prove

LEMMA 4.2. *We have for* $-1\leqq\alpha,\,\beta\leqq 0$

(4.13) $$V_{n+2}^{(\alpha+2,\beta+1)}\big(w(\cdot\,;\,p_n);\,p_n\big)\sim I_1(p_n).$$

Indeed, let $cn^{-1}\leqq\psi_n\leqq\pi/2$. Using the formulae (4.11); [1], (7.32.5), (8.9.2) and [17] Lemmas 1, 2 we get as in [19], 4.8.2,

$$I_1\geqq c_1\frac{\psi_n^{-2\alpha-5+4}}{n}\sum_{k=1}^{[n/2]}{}'\frac{w(|x_k-p_n|)}{(x_k-p_n)^2}\frac{k^{2\alpha+7-4}}{n^{2\alpha+8-4}}\geqq$$
$$\geqq c_1\frac{\psi_n^{-2\alpha-1}}{n}\sum_{k=1}^{[j/2]}\frac{w(|x_k-p_n|)}{(k+j)^2(k-j)^2}\frac{k^{2\alpha+3}}{n^{2\alpha}}\sim\frac{w(\sin^2\psi_n)}{(n\psi_n)^{2\alpha+1}}\frac{j^{2\alpha+4}}{j^4}\sim$$
$$\sim\frac{w(\sin^2\psi_n)}{n\psi_n}\geqq c_2\frac{w(\sin^2\psi_n)}{(n\psi_n)^{2\alpha+3}}\sim I_2(p_n),$$

where x_{jn} is the nearest root to p_n and $\sum'$ stands for $\sum\limits_{k\neq j}$.

The remaining parts can be handled similarly, i.e. we proved (4.13) (for the details see [19]).

4.5.5. Denote $x_k^{(\alpha,\beta)}$ by y_k. Then by (4.11) and (4.13)

$$|V_{n+2}^{(\alpha+2,\beta+1)}(f;\, p_n)-f(p_n)| = O(1)I_1(p_n) =$$

$$= O(1)\sum_{k=1}^{n}{}' w(|x_k-p_n|)\left[\frac{P_n^{(\alpha,\beta)}(z_n)}{P_n'^{(\alpha,\beta)}(y_k)(x_k-p_n)}\right]^2 v_{kn}(H^{(\alpha,\beta)};\, z_n) \sim$$

$$\sim \sum_{k=1}^{n} w(|y_k-z_n|)v_{k,n}(H^{(\alpha,\beta)};\, z_n)l_{k,n}^2(H^{(\alpha,\beta)};\, z_n)+w\left(\frac{1}{n}\right) =$$

$$= O(1)\left[\sup_{f\in C_1(w)}\|H_n^{(\alpha,\beta)}(f;\, x)-f(x)\|+w\left(\frac{1}{n}\right)\right] =$$

$$= O(1)\sup_{f\in C_1(w)}\|H_n^{(\alpha,\beta)}(f;\, x)-f(x)\| \overset{\text{def}}{=\!=} O(1)G_n^{(\alpha,\beta)}$$

for a certain $\{z_n=\cos\xi_n\}$ where $|\psi_n-\xi_n|=O(n^{-1})$. I.e., $F_{n+2}=O(G_n)$. With similar argument $G_n=O(F_{n+2})$ which proves our assertion if $-1\leqq\alpha,\ \beta\leqq 0$.

4.5.6. To complete our proof we state

LEMMA 4.3. *If* $\alpha\geqq\beta\geqq-1$, $\alpha>0$ *we have for arbitrary* $w(t)$

$$\sup_{f\in C_1(w)}\|S_N(\alpha,\beta)(f;\, x)-f(x)\| \sim n^{2\alpha}. \tag{4.14}$$

We proved (4.14) for $H_n^{(\alpha,\beta)}$ and $Q_{n+2}^{(\alpha+1,\beta+1)}$ ([19], 4.8.5). The remaining cases can be treated similarly.

4.6. PROOF OF THEOREM 3.3. **4.6.1.** Let

$$\sum_{k=1}^{n}|f(x_k)-f(x)||h_k(x)|+\sum_{i=1}^{m}|f(\xi_i)-f(x)||t_i(x)| \overset{\text{def}}{=\!=} D_N(f;\, x).$$

As in 4.5.3,

$$\|D_N(f;\, x)\| = O(1)\|D_N(f;\, x)\|_{[a_n,b_n]} \quad (n\geqq n_0).$$

4.6.2. First let $S_N(\alpha,\beta)=U_{n+1}$ or W_{n+2}.

4.6.3. Now using (2.4), 4.4.5 and 4.4.7, we obtain for $f\in C_1(\omega)$, as in [18] and [19]

$$D_n(f;\, x) = O(1)\left\{\sum_{i=1}^{n}\left[w\left(\frac{i\sin\vartheta}{n}\right)+w\left(\frac{i^2|\cos\vartheta|}{n^2}\right)\right]i^{2\gamma-1}+\right. \tag{4.15}$$

$$+w(\sin^2\vartheta)[(n\sin\vartheta)^{-2\alpha-3}+(n\sin\vartheta)^{-2\beta-3}]+(\sin\vartheta)^{-2\alpha-1}n^{-2\beta-3}\sum_{l=1}^{n}l^{2\beta+1}+$$

$$\left.+(\sin\vartheta)^{-2\beta-1}n^{-2\alpha-3}\sum_{l=1}^{n}l^{2\alpha+1}\right\} = \overset{\text{def}}{=\!=} O(1)B_N(w;\, x),$$

where $x=\cos\vartheta$ $(cn^{-1}\leqq\vartheta\leqq\pi-cn^{-1})$, $\gamma=\max(\alpha,\beta,-0.5)$.

By (4.15) we obtain

(4.16) $\|S_N(\alpha,\beta)(f;x)-f(x)\|_{[a_n,b_n]}=o(1)$ if $-1.5<\alpha,\ \beta<0,\quad |\alpha-\beta|<1.$

4.6.4. For the remaining cases we shall apply Theorem 2.2. Again by (4.15) we get

(4.17) $\|S_N(\alpha,\beta)(f;x)-f(x)\|_{[a_n,b_n]}=O(1)$ if $-1.5\leqq\alpha,\ \beta<0,\quad |\alpha-\beta|\leqq 1,$

from where one can get (2.8). So we have to prove (2.9). To get this, we suppose only $-1.5\leqq\alpha, \beta<0$. First remark that for U_{n+1}

$$\mathfrak{H}_k(x)=(x-x_k)\left[\frac{1-x}{1-x_k}\frac{P_n^{(\alpha+2,\beta)}(x)}{P_n'^{(\alpha+2,\beta)}(x_k)(x-x_k)}\right]^2\quad(1\leqq k\leqq n)$$

and

$$\mathfrak{H}_0(x)=(x-1)\left[\frac{P_n^{(\alpha+2,\beta)}(x)}{P_n^{(\alpha+2,\beta)}(1)}\right]^2.$$

As above, one can suppose $cn^{-1}\leqq\vartheta\leqq\pi-cn^{-1}$. By the formulae of [1] we get

$$|\mathfrak{H}_0(x)|=\begin{cases}O(1)\dfrac{\vartheta^2\vartheta^{-2\alpha-5}}{n\cdot n^{2\alpha+4}}=O(n^{-2}) & \text{if}\quad cn^{-1}\leqq\vartheta\leqq\pi/2,\\[2ex] O(1)\dfrac{(\pi-\vartheta)^{-2\beta-1}}{n\cdot n^{2\alpha+4}}=o(n^{-1}) & \text{if}\quad \pi/2\leqq\vartheta\leqq\pi-cn^{-1}.\end{cases}$$

For the remaining terms we get, using also [1] and the previous arguments, if $cn^{-1}\leqq\vartheta\leqq\pi/2$

$$\sum_{k=1}^{n}|\mathfrak{H}_k(x)|=O(1)\frac{\vartheta^{-2\alpha-1}}{n}\left[\sum_{k=1}^{n}{}'\frac{k^{2\alpha+3}}{n^{2\alpha+4}}\frac{n^2}{|k+j||k-j|}+\sum_{k=1}^{n}\frac{k^{2\beta+3}}{n^{2\beta+4}}\right]=$$

$$=O(1)\left[\frac{n^{2\alpha+1}j^{-2\alpha-1}}{n^{2\alpha+3}}\left(\sum_{k+1}^{[j/2]}+\sum_{k=[j/2]}^{2j}{}'+\sum_{k=2j}^{n}\right)\frac{k^{2\alpha+3}}{|k+j||k-j|}+\frac{n^{2\alpha+1}j^{-2\alpha-1}}{n^{2\beta+5}}\sum_{k=1}^{n}k^{2\beta+3}\right]=$$

$$=O(1)\left[\frac{j}{n^2}+\frac{j\ln j}{n^2}+\frac{j^{-2\alpha-1}}{n^2}|j^{2\alpha+2}-n^{2\alpha+2}|+j^{-2\alpha-1}n^{2\alpha}\right]=o(1),$$

where, as usual, $x\approx x_{j,n}$, so $\vartheta\approx jn^{-1}$. On the other hand, for $\pi/2\leqq\vartheta\leqq\pi-cn^{-1}$, as above,

$$\sum_{k=1}^{n}|\mathfrak{H}_k(x)|=O(1)\frac{\vartheta^{-2\beta-1}}{n}\left[\sum_{k=1}^{n}{}'\frac{k^{2\beta+3}}{n^{2\beta+4}}\frac{n^2}{|k+j||k-j|}+\sum_{k=1}^{n}\frac{k^{2\alpha+3}}{n^{2\alpha+4}}\right]=o(1),$$

which can be obtained as above. So we proved (2.9) for $S_N=U_{n+1}$. Similar arguments hold for W_{n+2}.

4.6.5. Now let e.g., $S_N=V_{n+1}$. By (4.11) we obtain, as at (4.16) and (4.17)

$$(4.18)\qquad \|I_{1,n}(x)\|+\|I_{2,n}(x)\|=\begin{cases}o(1) & \text{if}\quad -1.5<\alpha,\beta<0,\quad |\alpha-\beta|<1,\\ O(1) & \text{if}\quad -1.5\leqq\alpha,\beta<0,\quad |\alpha-\beta|\leqq 1.\end{cases}$$

If $-1.5<\alpha, \beta<0$, $|\alpha-\beta|<1$ we obtain

$$(4.19)\qquad |I_{3,n}(x)| = \begin{cases} O(1)\dfrac{\vartheta^4\vartheta^{-2\alpha-5}}{n\cdot n^{2\beta+2}} = o(1) & \text{if } x \geqq 0, \\ O(1)w(\sin^2\vartheta)\dfrac{(\sin\vartheta)^{-2\beta-3}}{n\cdot n^{2\beta+2}} = o(1) & \text{if } x \leqq 0, \end{cases}$$

which, by (4.18) gives the desired result. On the other hand, by (4.19)

$$(4.20)\qquad \|I_{3,n}(x)\| = O(1) \quad \text{if} \quad -1.5 \leqq \alpha,\ \beta < 0, \quad |\alpha-\beta| \leqq 1.$$

Remarking that (2.9) can be verified as in 4.6.4, we obtain our statement. (The "exceptional cases", e.g. $\alpha=-1$, $\beta=0$, can be handled as in [19]. We omit the details.)

4.7. Proof of Theorem 3.4. **4.7.1.** First let $\alpha=\beta=0$. With $f(x)=1-x$ and $2z_n=x_1+x_0$ we have, as in [1], (14.6)

$$S_N(0,0)(f; z_n) \sim 1 > f(z_n) \sim n^{-2}.$$

4.7.2. By usual argument we get

$$\sum_k |h_k(z_n)| \geqq cn^{2\alpha} \quad (\alpha \geqq 0)$$

(see, e.g., [1], 14.6), i.e., by 4.7.1 and 4.7.2 we settled the cases $\max(\alpha,\beta)\geqq 0$ (see the arguments of [19], 4.7).

4.7.3. Suppose e.g. $\alpha<-1.5$. Then by usual argument we have with $2s_n= =x_{[n/2]}+x_{[n/2]+1}$ that

$$|t_1(s_n)| \sim n^{-2\alpha-3} \quad \text{for} \quad R_{n+1} \quad \text{and} \quad Q_{n+2},$$

$$|h_0(s_n)| \sim n^{-2\alpha-3} \quad \text{for} \quad U_{n+1}, V_{n+2} \quad \text{and} \quad W_{n+2},$$

which tend to infinity with n.

4.7.4. Finally let $-1.5\leqq\alpha, \beta<0$ and, e.g., $\alpha-\beta>1$. We obtain

$$\sum_{\vartheta_k \geqq \pi-\varepsilon} |h_k(z_n)| \sim n^{2(\alpha-\beta-1)},$$

which completes the proof.

4.8. Proof of theorem 3.5. **4.8.1.** Let, e.g. $S_N=V_{n+2}$. By (4.15) and (4.19)

$$(4.21)\qquad B_N(W;\ w;\ x) = B_N(V;\ w;\ x) \overset{\text{def}}{=} C_N(x) \quad (cn^{-1} \leqq \vartheta \leqq \pi-cn^{-1}),$$

further, as above

$$\|D_N(f;\ x)\| = O(1)\,a(f)\,\|B_N(w;\ x)\|_{[a_n,b_n]}$$

whenever $f\in C(w)$ (for any S_N), i.e., by (4.21)

$$(4.22)\qquad \|S_N(\alpha,\beta)(f;\ x)-f(x)\| = O(1)\,a(f)\,\|C_N(x)\|_{[a_n,b_n]} \quad \text{if} \quad f\in C(w).$$

4.8.2. By the argument used in [19] we can prove that for any fixed sequence e_n $(0<e_{n+1}<e_n,\ \lim_{n\to\infty} e_n=0)$ there exist a function $f_1(x)\in C_1(w)$ and a sequence $\{n_i\}\subset\{q_i\}$ such that

$$(4.23)\quad \|V_{n+2}^{(\alpha+2,\beta+1)}(f_1;\,x)-f_1(x)\| \geqq c(f_1)e_n\|B_N(V;\,w;\,x\|_{[a_n,b_n]} \quad (n=n_1,n_2,\ldots).$$

Similarly, for any fixed $\tilde{e}_n$ $(0<\tilde{e}_{n+1}<\tilde{e}_n,\ \lim_{n\to\infty}\tilde{e}_n=0)$ there exist a function $f_2(x)\in$ $\in C_1(w)$ and a sequence $\{\tilde{n}_i\}\subset\{n_i\}$ such that

$$(4.24)\quad \|W_{n+2}^{(\alpha+2,\beta+2)}(f_2;\,x)-f_2(x)\| \geqq c(f_2)\tilde{e}_n\|B_N(W;\,w;\,x)\|_{[a_n,b_n]} \quad (n=\tilde{n}_1,\tilde{n}_2,\ldots).$$

(Indeed, obviously one can choose the term $F_N(x)$ from $B_N(w;\,x)$ such that $\|F_N(x)\|_{[a_n,b_n]}\sim\|C_N(x)\|_{[a_n,b_n]}$ (see (4.15)). Now, again by (4.15), we can find a fixed point $x^*\in[a_n,b_n]$ such that $F_N(x^*)\sim\|F_N(x)\|_{[a_n,b_n]}$. For this x^* we apply the arguments of [19], 4.3 and 4.7 to obtain (4.23) and (4.24)).

4.8.3. By (4.22)—(4.24) we obtain (3.9).

References

[1] G. Szegő, *Orthogonal polynomials,* Amer. Math. Soc. (New York, 1974)

[2] P. Szász, The extended Hermite—Fejér interpolation formula with application to the theory of generalized almost-step parabolas, *Publ. Math., Debrecen,* **11** (1964), 85—100.

[3] G. Grünwald, On the theory of interpolation, *Acta Math.,* **75** (1942), 219—245.

[4] L. Fejér, Über Interpolation, *Göttinger Nachrichten,* (1916), 66—91.

[5] L. Fejér, Lagrangesche Interpolation und die zugehörigen konjugierten Punkte, *Math. Ann.,* **106** (1932), 1—55.

[6] D. L. Berman, On everywhere divergent interpolation process of Hermite—Fejér type, *Izv. Vyss. Ucebn. Zaved., Mathematica,* **92** (1970), 3—8 (Russian).

[7] D. L. Berman, On interpolatory processes of Hermite—Fejér-type based on the Tchebycheff roots of second type, *Izv. Vyss. Ucebn. Zaved., Mathematica,* **1** (1977), 128—131 (Russian).

[8] R. B. Saxena, The Hermite—Fejér process on the Tchebycheff matrix of second kind, *Studia Sci. Math. Hungar.,* **9** (1974), 223—232.

[9] D. L. Berman, On the convergence of some interpolatory processes of Hermite—Fejér-type, *Izv. Vyss. Ucebn. Zaved., Mathematica,* **8** (1975), 97—101 (Russian).

[10] P. Szász, On quasi-Hermite—Fejér interpolation, *Acta Math. Acad. Sci. Hungar.,* **10** (1959), 413—439.

[11] L. P. Povchun, A. A. Privalov, On an interpolatory process defined by Egerváry and Turán, *Izv. Vyss. Ucebn. Zaved., Mathematica,* **8** (1974), 82—87 (Russian).

[12] J. Sántha, Convergence theorems of quasi-Hermite—Fejér interpolation, *Publ. Math., Debrecen,* **22** (1975), 23—29.

[13] R. B. Saxena, K. K. Mathur, The rapidity of convergence of quasi-Hermite—Fejér interpolation polynomials, *Acta Math. Acad. Sci. Hungar.,* **28** (1976), 343—347.

[14] A. Sharma, Remarks on quasi-Hermite—Fejér interpolation, *Canad. Math. Bull.,* **7** (1964), 101—119.

[15] R. B. Saxena, A note on D. L. Berman's theorem on the divergence of Hermite—Fejér interpolation, *Studia Sci. Math. Hungar.,* **7** (1972), 417—421.

[16] D. L. Berman, Interpolatory processes based on the roots of Jacobi polynomials, *Izv. Vyss. Ucebn. Zaved., Mathematica,* **10** (1973), 14—22 (Russian).

[17] G. I. Natanson, Two-sided estimate for the Lebesgue function of the Lagrange interpolation with Jacobi nodes, *Izv. Vyss. Ucebn. Zaved., Mathematica,* **11** (1967), 67—74 (Russian).

[18] P. Vértesi, Notes on the Hermite—Fejér interpolation based on the Jacobi abscissas, *Acta Math. Acad. Sci. Hungar.,* **24** (1973), 233—239.

[19] P. Vértesi, Hermite—Fejér-type interpolations. I, *Acta Math. Acad. Sci. Hungar.,* **32** (1978), 349—369.

[20] P. VÉRTESI, Hermite—Fejér-type interpolations. II, *Acta Math. Acad. Sci. Hungar.*, **33** (1979),
[21] S. B. STECKIN, Generalization of some inequalites of Bernstein, *DAN SSSR,* **60** (1948), 1511—1514 (Russian).
[22] M. SALLAY, Über Hermite—Fejérsche Interpolation, *Acta Math. Acad. Sci. Hungar.* (to appear).
[23] P. TURÁN, Some open problems of approximation theory, *Mat. Lapok,* **25** (1974), 21—75 (Hungarian).
[24] J. BALÁZS, On the convergence of Hermite—Fejér interpolation process, *Acta Math. Acad. Sci. Hungar.,* **9** (1958), 259—267.
[25] T. M. MILLS, A. K. VARMA, On a theorem of E. Egerváry and P. Turán on the stability of interpolation (preprint).
[26] J. SZABADOS, *Convergence and saturation problems of approximation processes,* Thesis, 1975 (Hungarian).
[27] V. KUMAR, K. K. MATHUR, On the rapidity of convergence of a quasi-Hermite—Fejér interpolation polynomial, *Studia Sci. Math. Hungar.*, **9** (1974), 313—319.
[28] T. M. MILLS, A convergent quasi-Hermite—Fejér interpolation process, *Bull. Austral. Math. Soc.,* **12** (1975), 267—276.
[29] V. KUMAR, Convergence of Hermite—Fejér interpolation polynomials on the extended nodes, *Publ. Math. Debrecen,* **24** (1977), 31—39.
[30] E. EGERVÁRY, P. TURÁN, Notes on interpolation. V, *Acta Math. Acad. Sci. Hungar.,* **9** (1958), 259—267.

(Received July 12, 1977)

MATHEMATICAL INSTITUTE OF THE
HUNGARIAN ACADEMY OF SCIENCES
1053 BUDAPEST, REÁLTANODA U. 13—15.

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 85—90.

GEWISSE ABSCHÄTZUNGEN ÜBER BESCHRÄNKTE ORTHONORMIERTE SYSTEME

Von

K. TANDORI (Szeged), Mitglied der Akademie

1. Es sei $1 \leqq K \leqq \infty$. Mit $\Omega(K)$ bezeichnen wir die Klasse der im Intervall (0, 1) orthonormierten Funktionensysteme $\varphi = \{\varphi_n(x)\}_1^\infty$ mit

$$|\varphi_n(x)| \leqq K \quad (x \in (0, 1);\ n = 1, 2, \dots).$$

(Im Falle $K=\infty$ ist $\Omega(K)$ die Klasse aller im Intervall (0, 1) orthonormierten Systeme.)

Für eine reelle Zahlenfolge $a = \{a_n\}_1^\infty$ setzen wir

$$\|a;\, K\| = \sup_{\varphi \in \Omega(K)} \sqrt{\int_0^1 \sup_{1 \leqq i \leqq j} |a_i \varphi_i(x) + \dots + a_j \varphi_j(x)|^2 \, dx}.$$

Auf Grund der Definition von $\|\cdot;\, K\|$ ist klar, daß für jede Folge a

$$\|a;\, 1\| \leqq \|a;\, K_1\| \leqq \|a;\, K_2\| \leqq \|a;\, \infty\| \quad (1 < K_1 < K_2 < \infty)$$

besteht.

Für beliebige natürliche Zahlen $N \leqq M$ setzen wir

$$a(N, \infty) = \{0, \dots, 0,\ a_N, a_{N+1}, \dots\},$$

$$a(N, M) = \{0, \dots, 0,\ a_N, \dots, a_M,\ 0, \dots\}.$$

Für jede Folge a und für jede natürliche Zahl $N < M$ gilt

$$\|a(N, M);\, K\| \leqq \|a;\, K\| \quad (1 \leqq K \leqq \infty).$$

Weiterhin, für jede Folge a besteht

$$\|a(1, N);\, K\| \nearrow \|a;\, K\| \quad (N \to \infty;\ 1 \leqq K \leqq \infty);$$

es gilt auch die Dreiecksungleichung

$$\|a+b;\, K\| \leqq \|a;\, K\| + \|b;\, K\| \quad (1 \leqq K \leqq \infty).$$

In [5] haben wir den folgenden Satz gezeigt.

SATZ A. *Es sei $1 < K < \infty$. Für jede Folge a sind die Relationen*

$$\lim_{N \to \infty} \|a(N, \infty);\, K\| = 0, \quad \lim_{N \to \infty} \|a(N, \infty);\, 1\| = 0$$

äquivalent.

2. In dieser Note werden wir zuerst die folgende Behauptung zeigen.

Satz I. *Für jede Zahl K $(1<K<\infty)$ gibt es eine nur von K abhängige positive Zahl $C(K)$ derart, daß*

$$\|a;\ 1\| \geqq C(K)\|a;\ K\|$$

für jede Folge a gilt.

Durch Anwendung der bekannten Abschätzungen in den Arbeiten [2], [3], [4] bekommen wir die folgenden Abschätzungen.

Folgerungen. 1. Ist $|a_1|\geqq \ldots \geqq |a_n| \geqq \ldots$, dann gilt

$$\|a;\ 1\| \geqq C_1\left\{a_1^2+\sum_{n=1}^{\infty} a_n^2 \log^2 n\right\}^{1/2}$$

mit einer positiven Konstante C_1.

2. Für jede Folge a besteht

$$\|a;\ 1\| \geqq C_2\left\{\sum_{n=2}^{\infty} a_n^2 \log_+^2 \frac{a_1^2+a_2^2+\ldots}{a_n^2}\right\}^{1/2}$$

mit einer positiven Konstante C_2, wobei

$$\log_+ x = \begin{cases} \log x, & \text{für } \log x \geqq 1, \\ 1 & \text{sonst} \end{cases}$$

bedeutet.

3. Für eine Folge a sei

$$S = \begin{cases} |a_1^*|+|a_2^*|+\sum\limits_{\nu=0}^{\infty} \sqrt{\sum\limits_{n=2^{2^\nu}+1}^{2^{2^{\nu+1}}} (a_n^*)^2 \log^2 n}, & a \in l_2, \\ \infty & \text{sonst}, \end{cases}$$

wobei $\{a_n^*\}_1^\infty$ eine im absoluten Betrage monoton abnehmende Anordnung der Folge $a=\{a_n\}_1^\infty$ bezeichnet. Für jede Folge a gilt

$$\sup_P \|a;\ 1\| \geqq C_3 S$$

mit einer positiven Konstante C_3, wobei $\sup\limits_P$ das Supremum für jede Anordnung der Folge $a=\{a_n\}_1^\infty$ bedeutet.

Beweis des Satzes I. Den Satz I werden wir „ad Absurdum" beweisen. Es sei K $(1<K<\infty)$ eine Zahl. Gilt der Satz I nicht, dann gibt es für jede natürliche Zahl m eine Folge $a^{(m)}=\{a_n^{(m)}\}_1^\infty$ mit

$$\|a^{(m)};\ 1\| < \|a^{(m)};\ K\|/m^2. \tag{1}$$

Nach obigem folgt aus (1), daß für jeden Index m eine natürliche Zahl N_m mit

$$\|a^{(m)}(1, N_m);\ 1\| \leqq \|a^{(m)}(1, N_m);\ K\|/m^2 \tag{2}$$

existiert. Ohne Beschränkung der Allgemeinheit können wir

$$\|a^{(m)}(1, N_m);\ K\| = 1 \tag{3}$$

annehmen.

Wir setzen

$$b_{n+N_1+\ldots+N_{m-1}} = a_n^{(m)} \quad (n = 1, \ldots, N_m;\ m = 1, 2, \ldots).$$

Für die so definierte Folge $\{b_n\}_1^\infty$ gilt für jede natürliche Zahl N mit $N_1+\ldots+N_{m_0}< <N\leqq N_1+\ldots+N_{m_0+1}$ die Abschätzung

$$\|b(N, \infty;\ 1\| \leqq \sum_{m=m_0}^{\infty} \|b(N_m+1, N_{m+1});\ 1\| \leqq \sum_{m=m_0}^{\infty} \frac{1}{m^2} = O\left(\frac{1}{m_0}\right)$$

auf Grund von (2) und (3). Daraus folgt

$$\lim_{N\to\infty} \|b(N, \infty);\ 1\| = 0. \tag{4}$$

Es sei N eine natürliche Zahl und m ein solcher Index, für welchen $N<N_1+\ldots+N_m$ gilt. Dann ist nach obigem und nach (3)

$$\|b(N, \infty);\ K\| \geqq \|b(N_1+\ldots+N_m+1, N_1+\ldots+N_{m+1});\ K\| =$$
$$= \|a^{(m)}(1, N_{m+1});\ K\| = 1,$$

woraus sich

$$\lim_{N\to\infty} \|b(N, \infty);\ K\| \neq 0 \tag{5}$$

ergibt.

(4) und (5) widersprechen aber dem Satz A.

3. In der Arbeit [1] haben wir bewiesen, daß im Falle

$$|c_n| \leqq |d_n| \quad (n = 1, 2, \ldots) \tag{6}$$

die Ungleichung

$$\|c;\ \infty\| \leqq \|d;\ \infty\|$$

besteht. Weiterhin haben wir in [2] gezeigt, daß im Falle (6)

$$\|c;\ K\| \leqq 2\frac{K}{K-1}\|d;\ K\| \quad (1 < K < \infty) \tag{7}$$

gilt.

In dieser Note werden wir noch folgendes zeigen.

SATZ II. *Es gibt eine positive Konstante C derart, daß im Falle* (6)

$$\|c;\ K\| \leqq C\|d;\ K\| \quad (1 \leqq K \leqq \infty)$$

gilt.

BEWEIS DES SATZES II. Wir zeigen erstens, daß mit einer positiven Konstante C_1

$$\|c;\ 1\| \leqq C_1\|d;\ 1\| \quad (|c_n| \leqq |d_n|;\ n = 1, 2, \ldots) \tag{8}$$

besteht.

Im entgegengesetzten Falle gibt es Folgen $c(m)=\{c_n(m)\}_1^\infty$, $d(m)=\{d_n(m)\}_1^\infty$ mit

$$|c_n(m)| \leqq |d_n(m)| \quad (n = 1, 2, \ldots;\ m = 1, 2, \ldots),$$
$$\|c(m);\ 1\| > m^2\|d(m);\ 1\| \quad (m = 1, 2, \ldots).$$

Dann gibt es eine natürliche Zahl N_m mit

$$\|\{c_n(m)\}_1^{N_m}; 1\| \geqq m^2 \|\{d_n(m)\}_1^{N_m}; 1\| \quad (m = 1, 2, \ldots). \tag{9}$$

Ohne Beschränkung der Allgemeinheit können wir

$$\|d_n(m)\}_1^{N_m}; 1\| = 1/m^2 \quad (m = 1, 2, \ldots) \tag{10}$$

voraussetzen.

Es seien

$$C_{n+N_1+\ldots+N_{m-1}} = c_n(m) \quad (n = 1, \ldots, N_m;\ m = 1, 2, \ldots),$$

$$D_{n+N_1+\ldots+N_{m-1}} = d_n(m) \quad (n = 1, \ldots, N_m;\ m = 1, 2, \ldots).$$

Nach obigem gilt für $N_1+\ldots+N_{m_0} < N \leqq N_1+\ldots+N_{m_0+1}$

$$\|D(N, \infty); 1\| \leqq \sum_{m=m_0}^{n} \|\{D_n\}_{N_1+\ldots+N_m+1}^{N_1+\ldots+N_{m+1}}; 1\| \leqq \sum_{m=m_0}^{\infty} \frac{1}{m^2} = O\left(\frac{1}{m_0}\right),$$

woraus

$$\lim_{N\to\infty} \|D(N, \infty); 1\| = 0 \tag{11}$$

folgt.

Es sei $\varepsilon > 0$ beliebig vorgegeben. Dann gibt es ein orthonormiertes System $\varphi_1(x), \ldots, \varphi_{N_m}(x)$ aus $\Omega(1)$, für welches

$$\int_0^1 \max_{1\leqq i\leqq j\leqq N_m} |c_i(m)\varphi_i(x)+\ldots+c_j(m)\varphi_j(x)|^2\,dx \geqq \|\{c_n(m)\}_1^{N_m}; 1\|^2 - \varepsilon \tag{12}$$

gilt. Wir setzen

$$\psi_n(x) = \begin{cases} \sqrt{2}\,\dfrac{c_n(m)}{d_n(m)}\,\varphi_n(2x), & x \in (0, 1/2), \\ \sqrt{2}\left(1 - \dfrac{c_n^2(m)}{d_n^2(m)}\right)^{1/2} \varphi_n(2x-1), & x \in (1/2, 1) \end{cases} \quad (n = 1, \ldots, N_m).$$

Es ist klar, daß $\{\psi_n(x)\}_1^{N_m} \in \Omega(\sqrt{2})$. Weiterhin gilt

$$\|\{d_n(m)\}_1^{N_m}; \sqrt{2}\|^2 \geqq \int_0^1 \max_{1\leqq i\leqq j\leqq N_m} |d_i(m)\psi_i(x)+\ldots+d_j(m)\psi_j(x)|\,dx^2 \geqq$$

$$\geqq 2\int_0^{1/2} \max_{1\leqq i\leqq j\leqq N_m} |c_i(m)\varphi_i(2x)+\ldots+c_j(m)\varphi_j(2x)|^2\,dx =$$

$$= \int_0^1 \max_{1\leqq i\leqq j\leqq N_m} |c_i(m)\varphi_i(x)+\ldots+c_j(m)\varphi_j(x)|^2\,dx \geqq \|\{c_n(m)\}_1^{N_m}; 1\|^2 - \varepsilon$$

auf Grund von (12), woraus sich

$$\|\{d_n(m)\}_1^{N_m}; \sqrt{2}\| \geqq \|\{c_n(m)\}_1^{N_m}; 1\| \quad (m = 1, 2, \ldots)$$

ergibt. Daraus und aus (9), (10) folgt

$$\|\{d_n(m)\}_1^{N_m};\ \sqrt{2}\| \geqq 1 \quad (m = 1, 2, \ldots).$$

Nach obigem erhalten wir daraus

$$\lim_{N\to\infty} \|D(N, \infty);\ \sqrt{2}\| \neq 0.$$

Auf Grund des Satzes A bekommen wir

$$\lim_{N\to\infty} \|D(N, \infty);\ 1\| \neq 0,$$

was (11) widerspricht. Damit haben wir (8) bewiesen.

Den Satz II werden wir auch „ad Absurdum" zeigen. Ist der Satz II nicht richtig, dann gibt es — auf Grund von (7) und (8) — Zahlen $1 < \ldots < K_m < \ldots \ldots < K_1 < \infty$ und Folgen $\{c_n(m)\}_1^\infty$, $\{d_n(m)\}_1^\infty$ mit

$$|c_n(m)| \leqq |d_n(m)| \quad (n = 1, 2, \ldots;\ m = 1, 2, \ldots),$$

$$\|c(m);\ K_m\| > m^2 \|d(m);\ K_m\| \quad (m = 1, 2, \ldots).$$

Nach obigem gibt es Indizes N_m mit

$$\|\{c_n(m)\}_1^{N_m};\ K_m\| \geqq m^2 \|\{d_n(m)\}_1^{N_m};\ K_m\| \quad (m = 1, 2, \ldots). \tag{13}$$

Ohne Beschränkung der Allgemeinheit können wir

$$\|\{d_n(m)\}_1^{N_m};\ K_m\| = 1/m^2 \quad (m = 1, 2, \ldots) \tag{14}$$

voraussetzen. Es sei

$$C_{n+N_1+\ldots+N_{m+1}} = c_n(m) \quad (n = 1, \ldots, N_m;\ m = 1, 2, \ldots),$$

$$D_{n+N_1+\ldots+N_{m+1}} = d_n(m) \quad (n = 1, \ldots, N_m;\ m = 1, 2, \ldots).$$

Für $N_1 + \ldots + N_{m_0} < N \leqq N_1 + \ldots + N_{m_0+1}$ erhalten wir

$$\|D(N;\ \infty);\ 1\| \leqq$$

$$\leqq \sum_{m=m_0}^{\infty} \|\{D_n\}_{N_1+\ldots+N_m+1}^{N_1+\ldots+N_{m+1}};\ 1\| \leqq \sum_{m=m_0}^{\infty} \|\{D_n\}_{N_1+\ldots+N_m+1}^{N_1+\ldots+N_{m+1}};\ K_m\| = \sum_{m=m_0}^{\infty} \frac{1}{m^2} = O\left(\frac{1}{m_0}\right)$$

auf Grund von (14); woraus sich

$$\lim_{N\to\infty} \|D(N, \infty);\ 1\| = 0 \tag{15}$$

ergibt.

Ferner folgt auf Grund von (13) und (14) nach obigem

$$\varliminf_{N\to\infty} \|C(N, \infty);\ K_1\| \geqq \varliminf_{m\to\infty} \|\{c_n(m)\}_1^{N_m};\ K_m\| \geqq 1.$$

Nach dem Satz I erhalten wir daraus:

$$\lim_{N\to\infty} \|C(N, \infty);\ 1\| \neq 0, \tag{16}$$

und nach (8), (15), wegen $|C_n| \leqq |D_n|$ $(n=1, 2, \infty)$ bekommen wir

$$\lim_{N\to\infty} \|C(N, \infty); 1\| = 0,$$

was (16) widerspricht.

Damit haben wir auch den Satz II bewiesen.

Literaturverzeichnis

[1] K. TANDORI, Über die Konvergenz der Orthogonalreihen. II, *Acta Sci. Math. Szeged,* **25** (1964), 219—232.

[2] K. TANDORI, Über die Konvergenz der Orthogonalreihen, III, *Publ. Math. Debrecen,* **12** (1965), 127—157.

[3] K. TANDORI, Über das Maximum der Summen orthogonaler Funktionen, *Acta Sci. Math. Szeged,* **32** (1971), 317—326.

[4] K. TANDORI, Über die unbedingte Konvergenz der Orthogonalreihen, *Acta Sci. Math. Szeged,* **32** (1971), 11—40.

[5] K. TANDORI, Über beschränkte orthonormierte Systeme, *Acta Math. Acad. Sci. Hungar.,* **31** (1978), 279—285.

(Eingegangen am 8. August 1977.)

JÓZSEF ATTILA UNIVERSITÄT
BOLYAI INSTITUT
H—6720 SZEGED, ARADI VÉRTANÚK TERE 1.

Acta Mathematica Academiae Scintiarum Hungaricae
Tomus 34 (1—2), (1979), 91—100.

ON *f*-REPRESENTATION OF REAL NUMBERS

By
M. L. MEHTA (Gif-sur-Yvette)

Introduction

BISSINGER [1], EVERETTE [2], RÉNYI [3] and others have considered representing a real number x by a finite or infinite sequence of digits called the "f-expansion" of x,

$$x = \varepsilon_0+f(\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots) \equiv \varepsilon_0+f\big(\varepsilon_1+f(\varepsilon_2+f(\varepsilon_3+\dots))\big). \tag{0.1}$$

The "digits" $\varepsilon_j=\varepsilon_j(x)$ and the "remainders" $r_j(x)$, $j=0, 1, 2, \dots$ are obtained by iterations of a positive monotone function $f(x)$ taking values from 0 to 1,

$$\begin{cases} \varepsilon_0(x) = [x], & r_0(x) = (x), \\ \varepsilon_{j+1}(x) = \big[\varphi\big(r_j(x)\big)\big], & r_{j+1}(x) = \big(\varphi(r_j(x))\big) \quad (j = 0, 1, 2, \dots), \end{cases} \tag{0.2}$$

where $[z]$ denotes the integral part and (z) the fractional part of a real number z, while $x=\varphi(y)$ is the inverse function of $y=f(x)$. If for some n, $r_n=0$, then by definition $r_{n+k}=0$, all $k>0$, and x is represented by the finite sequence

$$x = \varepsilon_0+f(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) \equiv \varepsilon_0+f\big(\varepsilon_1+f(\varepsilon_2+\dots+f(\varepsilon_n)\dots)\big). \tag{0.3}$$

The representation (0.1) or (0.3) reduces for $f(x)=x/q$, q an integer $\geqq 2$, to the q-adic expansion $x=\sum_j \varepsilon_j/q^j$, and for $f(x)=1/x$ to the continued fraction expansion.

Given the numbers as sequences of integers $\varepsilon_1, \varepsilon_2, \dots$, a criteria is given to recognize their relative magnitudes.

To this mode of representing real numbers is associated a transformation of the interval (0, 1) to itself,

$$f(\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots) \to f(\varepsilon_2, \varepsilon_3, \dots)$$

or

$$Tx = \big(\varphi(x)\big). \tag{0.4}$$

Under certain conditions on $f(x)$, there is a measure invariant under this transformation; we derive an equation convenient for its numerical evaluation. The case $f(x)=(1+x)^{1/m}-1$ is treated as an example.

1. Uniqueness of f-expansions

Let $f(x)$ satisfy the following conditions.

(1.1) $f(0)=0$,

(1.2) $f(x)$ is positive, continuous and strictly increasing for $0\leqq x\leqq X$, $f(X)=1$, where $1<X\leqq\infty$. If X is finite, it is an integer, $X=\infty$ means $\lim_{x\to\infty} f(x)=1$.

(1.3) $$|f(x_1)-f(x_2)|<|x_1-x_2| \quad \text{for} \quad 0\leqq x_1<x_2.$$

Then RÉNYI [3] showed that every real number x, $0\leqq x<1$, has a representation (0.1) with $0\leqq\varepsilon_j<X$, $j>0$.

In case X is a finite integer, a tail of recurring $X-1$ can be eliminated. In other words, if for some fixed j, $\varepsilon_j<X-1$, while $\varepsilon_{j+k}=0$ for every $k>0$, then because

$$f(X-1, X-1, \dots)\equiv f\big(X-1+f(X-1+\dots)\big)=1,$$

one has

(1.4) $$f(\varepsilon_1, \dots, \varepsilon_j, X-1, X-1, \dots)=f(\varepsilon_1, \dots, \varepsilon_j+1, 0, 0, \dots).$$

Except for this ambiguity, the representation of x by the integers $\varepsilon_1, \varepsilon_2, \dots$, is unique. Also any sequence of positive integers $\varepsilon_1, \varepsilon_2, \dots$, $0\leqq\varepsilon_j<X$, represents a unique number x, $0\leqq x\leqq 1$. In particular, for any infinite sequence of integers $\varepsilon_1, \varepsilon_2, \dots$, $0\leqq\varepsilon_j\leqq X$, limit of $f(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$ as $n\to\infty$ exists and lies between 0 and 1.

Next let $f(x)$ satisfy the conditions,

(1.5) $f(1)=1$,

(1.6) $f(x)$ is positive, continuous and strictly decreasing for $1\leqq x\leqq X$, $f(X)=0$, where $2<X\leqq\infty$. If X is finite, it is an integer; $X=\infty$ means $\lim_{x\to\infty} f(x)=0$.

(1.7) $|f(x_1)-f(x_2)|<|x_1-x_2|$ for $1\leqq x_1<x_2$ and $|f(x_1)-f(x_2)|<\lambda|x_1-x_2|$, $0<\lambda<1$, if $1+f(2)<x_1<x_2$.

Then RÉNYI [3] also showed that every real number x has a representation (0.1), with $0\leqq\varepsilon_j<X$ for $j>0$, ε_j integer, and if some $\varepsilon_j=0$, $j>0$, then $\varepsilon_{j+k}=0$ for every $k>0$.

This representation is unique provided that if $\varepsilon_j=0$, $j>1$, then $\varepsilon_{j-1}\neq 1$. This is necessary, because

$$f(1, 0, 0, \dots)=f(1)=1,$$

and therefore

(1.8) $$f(\varepsilon_1, \dots, \varepsilon_j, 1, 0, 0, \dots)=f(\varepsilon_1, \dots, \varepsilon_j+1, 0, 0, 0, \dots).$$

Also any sequence of integers $\varepsilon_1, \varepsilon_2, \dots$, represents a unique number x, $0\leqq x<1$, provided that $1\leqq\varepsilon_j<X$ for non-zero ε_j; if $\varepsilon_j=0$, then $\varepsilon_{j+k}=0$ for every $k>0$; and if $\varepsilon_j=0$, $j>1$, then $\varepsilon_{j-1}\neq 1$. In particular, for any infinite sequence of integers $\varepsilon_1, \varepsilon_2, \dots$, $1\leqq\varepsilon_j<X$, the limit of $f(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$ as $n\to\infty$ exists and lies between 0 and 1.

If X appearing in conditions (1.2) or (1.6) is finite but not an integer, then a sequence of integers $\varepsilon_1, \varepsilon_2, \dots$ (with $0\leqq\varepsilon_j<X$ for increasing $f(x)$, or $1\leqq\varepsilon_j<X$,

$\varepsilon_j \neq 0$; $\varepsilon_j = 0 \rightarrow \varepsilon_{j+k} = 0$, $k > 0$, for decreasing $f(x)$) does not necessarily correspond to a real number x. An interesting example is $f(x) = x/\beta$, $\beta > 1$, β not an integer, studied in detail by RÉNYI [3] and PARRY [4].

If the conditions (1.3) or (1.7) on the slope of $f(x)$ is not satisfied, then the correspondance between x and its "representation" $\varepsilon_0 + f(\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots)$ determined by (0.1), (0.2) may not be one to one. An interesting example due to EVERETT [2] is

$$f(x) = \begin{cases} x/4, & 0 \leqq x \leqq 4/3 \\ x-1, & 4/3 \leqq x \leqq 5/3 \\ (x+1)/4, & 5/3 \leqq x \leqq 3, \end{cases} \tag{1.9}$$

$$\varphi(y) = \begin{cases} 4y, & 0 \leqq y \leqq 1/3 \\ y+1, & 1/3 \leqq y \leqq 2/3 \\ 4y-1, & 2/3 \leqq y \leqq 1. \end{cases} \tag{1.10}$$

Any number x with $4/3 \leqq x \leqq 5/3$ has the f-expansion $x = f(1, 1, 1, \ldots)$.

2. Ordering of numbers and ordering of sequences

Let $(a_1, a_2, \ldots)$ and $(b_1, b_2, \ldots)$ be finite or infinite sequences of non-negative integers. If they are of unequal lengths, we complete the smaller one by a number of zeros.

Let $f(x)$ satisfy conditions (1.1), (1.2) and (1.3). Let $(a_1, a_2, \ldots)$ and $(b_1, b_2, \ldots)$ be two sequences of integers with $0 \leqq a_j < X$, $0 \leqq b_j < X$, and neither of them has a tail of recurring $X-1$. Let $a = f(a_1, a_2, \ldots)$ and $b = f(b_1, b_2, \ldots)$. Then a is greater (less) than b if the first non-zero integer in the sequence $a_1 - b_1$, $a_2 - b_2$, $a_3 - b_3$, ... is positive (negative).

PROOF. (i) If a sequence of integers $(\varepsilon_1, \varepsilon_2, \ldots)$, $0 \leqq \varepsilon_j < X$ does not have a tail of recurring $X-1$, then from the uniqueness of the f-expansion one has $0 \leqq f(\varepsilon_1, \varepsilon_2, \ldots) = (x) < 1$.

(ii) If for some $n \geqq 1$, $a_n > b_n$, then $X > a_n + f(a_{n+1}, a_{n+2}, \ldots) \geqq a_n \geqq b_n + 1 > b_n + f(b_{n+1}, b_{n+2}, \ldots) \geqq 0$. If $n \geqq 1$ and $a_j = b_j$ for $1 \leqq j < n$, then as $f(x)$ is strictly increasing between 0 and X, we get step by step for $j = n, n-1, \ldots, 1$, $f(a_j, a_{j+1}, \ldots) > f(b_j, b_{j+1}, \ldots)$. Q. e. d.

Next let $f(x)$ satisfy conditions (1.5), (1.6) and (1.7). Let $(a_1, a_2, \ldots, a_n)$ be a finite or infinite sequence of integers with $1 \leqq a_j < X$ $j = 1, 2, \ldots, n$, $a_n \neq 1$. Similarly for the sequence of integers $(b_1, b_2, \ldots, b_m)$. Let $a = f(a_1, a_2, \ldots, a_n)$ and $b = f(b_1, b_2, \ldots, b_m)$. Without loss of generality one may assume $n \geqq m$. Then a is greater (less) than b if the first nonzero integer in the sequence $b_1 - a_1$, $a_2 - b_2$, $b_3 - a_3$, $a_4 - b_4$, ..., $(-)^m(a_m - b_m)$, $(-)^m a_{m+1}$ is positive (negative).

The proof depends on the uniqueness of the f-expansion, so that $0 < f(a_j, a_{j+1}, \ldots) < 1$, if $1 \leqq a_j, a_{j+1}, \ldots < X$. As $f(x)$ is strictly decreasing in $(1, X)$, one has $f(a_j, a_{j+1}, \ldots) > f(b_j, b_{j+1}, \ldots)$, if $a_j < b_j$. A recurrence on j then completes the proof.

Whether $f(x)$ is increasing or decreasing and satisfies the corresponding conditions in section 1 above for the uniqueness of f-expansions, then the set of numbers with finite sequences, i.e. $x=f(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$, are everywhere dense on (0, 1).

Actually any number x, $0\leqq x<1$ has a unique representation $x=f(\varepsilon_1, \varepsilon_2, \dots)$. If for some j and every $k>0$, $\varepsilon_{j+k}=0$, then x has a finite sequence. If not, choose $x_n=f(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$, so that x_n has a finite sequence and $\lim_{n\to\infty} x_n=x$.

3. Ergodicity of f-expansions and the invariant measure

Let $f(x)$ satisfy either the conditions (1.1), (1.2) and (1.3) or the conditions (1.5), (1.6) and (1.7). Then $f(x)$ is absolutely continuous and almost everywhere differentiable. To an f-expansion one may associate a mapping of the interval (0, 1) on itself as follows:

$$Tx = (\varphi(x)), \quad 0 \leqq x < 1 \tag{3.1}$$

where $y=\varphi(x)$ is the function inverse to $x=f(y)$ and the extra bracket means the fractional part of $\varphi(x)$. To make sure that this mapping T is ergodic and has an invariant measure, one assumes some extra conditions on $f(x)$.

Write

$$D_n(x, t) = \frac{d}{dt} f(\varepsilon_1(x), \varepsilon_2(x), \dots, \varepsilon_n(x)+t), \tag{3.2}$$

whenever $\varepsilon_n(x)$ is defined. Let in addition $f(x)$ satisfy the condition

$$\operatorname*{Sup}_{0\leqq t\leqq 1} |D_n(x, t)| \leqq C \operatorname*{Inf}_{0\leqq t\leqq 1} |D_n(x, t)|, \tag{3.3}$$

where the constant $C\geqq 1$ does not depend neither on x nor on n. Then RÉNYI [3] showed that there exists an integrable density $h(x)$ invariant under the mapping Tx of (0, 1) on itself, satisfying the inequalities

$$\frac{1}{C} \leqq h(x) \leqq C, \quad 0 \leqq x \leqq 1. \tag{3.4}$$

Also for any function $\alpha(x)$ integrable on (0, 1) and for almost all z, one has

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \alpha(T^k z) = M(\alpha) = \int_0^1 \alpha(x) h(x)\, dx, \tag{3.5}$$

where $M(\alpha)$ is a finite constant, and $T^k z$ is defined by the recursion $T^k z=T(T^{k-1}z)$ along with equation (3.1).

This theorem of Rényi assures us of the existence of $h(x)$, but unfortunately does not give any hint of how to find it. One may write an equation for $h(x)$, or still better for the integral of $h(x)$, as follows.

For a given x, $0\leqq x\leqq 1$, there are several solutions of $x=(\varphi(y))$, $0\leqq y\leqq 1$. Let them be y_j, $j=0, 1, \dots, X-1$. In other words $x+j=\varphi(y_j)$, or $y_j=f(x+j)$,

$j=0, 1, \ldots, X-1$. As $h(x)$ is invariant, we have

$$|h(x)\,dx| = \sum_{j=0}^{X-1} |h(y_j)\,dy_j|, \tag{3.6}$$

or

$$h(x) = \sum_{j=0}^{X-1} |f'(x+j)|\,h\big(f(x+j)\big), \tag{3.7}$$

where $f'(x)$ is the derivative of $f(x)$.

This equation is in general non-linear and its solution may not be trivial. However, it can be simplified and put in a form suitable at least for a numerical solution. In the rest of this section we take $f(x)$ to be an increasing function. For decreasing $f(x)$ the reasoning is almost the same with a few minor changes.

Put

$$h(x) = \varphi'(x)g\big(\varphi(x)\big) \quad \text{or} \quad f'(x)h\big(f(x)\big) = g(x), \tag{3.8}$$

in equation (3.7) to get

$$\varphi'(x)g\big(\varphi(x)\big) = \sum_{j=1}^{X-1} g(x+j). \tag{3.9}$$

As $g(x)$ is integrable on $(0, X)$, we set

$$G(x) = \int_0^x g(y)\,dy, \tag{3.10}$$

so that $G(x)$ is continuous, almost everywhere differentiable and satisfies

$$G\big(\varphi(x)\big) = \sum_{j=0}^{X-1} \{G(x+j)-G(j)\}, \quad 0 \leqq x < 1. \tag{3.11}$$

The boundary conditions are

$$G(0) = 0, \quad G(X) = 1. \tag{3.12}$$

As $g(x) \geqq 0$, $G(x)$ is monotonous non-decreasing.

The right hand side of equation (3.11) refers to values of $G(x)$ in disjoint intervals $(j, j+1)$ for $j=0, 1, 2, \ldots, X-1$. A continuous solution may therefore be obtained by choosing $G(x)$ arbitrarily on the range $(0, X)$ except for one interval, say $(j, j+1)$, on which it is then determined by equation (3.11). This situation does not mean that the invariant measure is arbitrary to a large extent, since the $G(x)$ so obtained is not necessarily non-decreasing. The monotonic non-decreasing character of $G(x)$ surprisingly forces it to be almost unique as required by Rényi's theorem.

To evaluate $G(x)$ numerically, we start with a non-decreasing function, say $G_0(x)=x/X$, then calculate successively $G_n(x)$ for $n=1, 2, \ldots$ by the equation

$$G_n(x) = \sum_{j=0}^{X-1} \big\{G_{n-1}\big(f(x)+j\big)-G_{n-1}(j)\big\}. \tag{3.13}$$

It is easy to verify that $G_n(x)$ satisfies (3.12) and is non-decreasing for every n. Moreover, the limit of $G_n(x)$ as $n\to\infty$ exists and is independent of $G_0(x)$,

$$\lim_{n\to\infty} G_n(x) = G(x). \tag{3.14}$$

Actually,

$$G_0(x) = \int_0^X G_0'(y)\theta(x-y)\,dy, \tag{3.15}$$

where a prime denotes the derivative and θ is the step function, $\theta(x)=1$ if $x>0$ while $\theta(x)=0$ if $x<0$. As $f(x)$ is strictly increasing in $(0, X)$,

$$\theta(x-y) = \theta\big(f(x)-f(y)\big) \tag{3.16}$$

and the above equation may be written as

$$G_0(x) = \int_0^X dy\, G_0'(y)\theta\big(f(x)-z\big) \tag{3.17}$$

with $z=f(y)$ or $y=\varphi(z)$. Substituting (3.15) in (3.13) we get

$$G_1(x) = \int_0^X dy\, G_0'(y) \sum_{j=0}^{X-1} \big\{\theta\big(f(x)+j-y\big)-\theta(j-y)\big\} = \tag{3.18}$$

$$= \int_0^X dy\, G_0'(y)\theta\big(f(x)-y+[y]\big) = \int_0^X dy\, G_0'(y)\theta\big(f(x)-Tz\big),$$

where $[y]$ is the integral part of y. In the last line we have used the fact, equation (3.1), that $Tz=\varphi(z)-[\varphi(z)]=y-[y]$. Iterating several times we get therefore,

$$G_n(x) = \int_0^X dy\, G_0'(y)\theta\big(f(x)-T^n z\big), \tag{3.19}$$

with $T^n z$ defined recursively by $T^n z = T(T^{n-1}z)$.

We can now take over Rényi's arguments and show that for almost every z the limit of $\theta\big(f(x)-T^n z\big)$ exists,

$$\lim_{n\to\infty} \theta\big(f(x)-T^n z\big) = G(x), \tag{3.20}$$

and it is independent of z. From equations (3.19) and (3.20) we get (3.14).

We can assert a little more. If $f(x)$ is convex from below then so is $G(x)$. In symbols, let $0<\lambda<1$, $0<x<X$, $0<y<X$ and $z=\lambda x+(1-\lambda)y$. If $f(z)\geqq\lambda f(x)+$ $+(1-\lambda)f(y)$ then $G(z)\geqq\lambda G(x)+(1-\lambda)G(y)$.

To see this it will suffice to convince oneself that the convexity of $G(x)$ is preserved through the iterations (3.13). In fact, as $G_{n-1}(x)$ is non-decreasing and $f(x)$ is convex from below, one has

$$G_{n-1}\big(f(z)+j\big) \geqq G_{n-1}\big(\lambda f(x)+(1-\lambda)f(y)+j\big).$$

Now from the convexity of G_{n-1} this last quantity is greater than or equal to $\lambda G_{n-1}(f(x)+j)+(1-\lambda)G_{n-1}(f(y)+j)$. A summation over j gives then

$$G_n(z) \geqq \lambda G_n(x)+(1-\lambda)\,G_n(x).$$

A numerical answer whose precision can be increased at will is next best to an analytical answer. The later in some cases may even not be expressible in terms of elementary functions.

4. Examples

A few examples follow.

EXAMPLE 1. $f(x)=x/q$ for $0\leqq x\leqq q$, where q is an integer greater than one.

$$D_n(x,t) = q^{-n}, \quad h(x) = 1, \quad G(x) = x.$$

EXAMPLE 2. $f(x)=(q-x)/(q-1)$, for $1\leqq x\leqq q$, where q is an integer greater than two.

$$D_n(x,t) = (-)^n(q-1)^{-n}, \quad h(x) = 1, \quad G(x) = x.$$

EXAMPLE 3. $f(x)=1/x$, for $1\leqq x<\infty$.

From the theory of continued fractions $D_n(x, t)$ can be calculated (cf. RÉNYI [3], pp. 487—488) and the constant C in equation (3.3) can be seen to be at most 4. Therefore from equation (3.4), $\frac{1}{4}\leqq h(x)<4$. Equation (3.10) reads

$$\frac{1}{x^2}\,g\left(\frac{1}{x}\right) = \sum_{j=1}^{\infty} g(x+j). \tag{4.1}$$

For this case $h(x)$ is known from other sources (cf. LEVY [5] and DENJOY [6]):

$$h(x) = (\ln 2)^{-1}(x+1)^{-1}, \quad g(x) = (\ln 2)^{-1}\{x(x+1)\}^{-1}. \tag{4.2}$$

EXAMPLE 4. (Bolyai's algorithm). $f(x)=(1+x)^{1/m}-1$, $0\leqq x\leqq 2^m-1$, where m is an integer greater than one. (BOLYAI [7] considered the particular case of $m=2$.) Here again the constant C in equation (3.3) can be seen (see RÉNYI [3], pp. 487—488) to be at most 2, and $\frac{1}{2}\leqq h(x)\leqq 2$. Equation (3.11) reads with a little simplification,

$$\mathcal{G}(x^m) = \sum_{j=0}^{2^m-2} \{\mathcal{G}(x+j)-\mathcal{G}(j+1)\}, \quad 1 \leqq x \leqq 2 \tag{4.3}$$

with

$$\mathcal{G}(1) = 0, \quad \mathcal{G}(2^m) = 1. \tag{4.4}$$

If we require $\mathcal{G}(x)$ satisfying (4.3) and (4.4) to be only continuous, then as noted in section 3 one may choose $\mathcal{G}(x)$ arbitrarily for example on $1\leqq x\leqq 2^m-1$

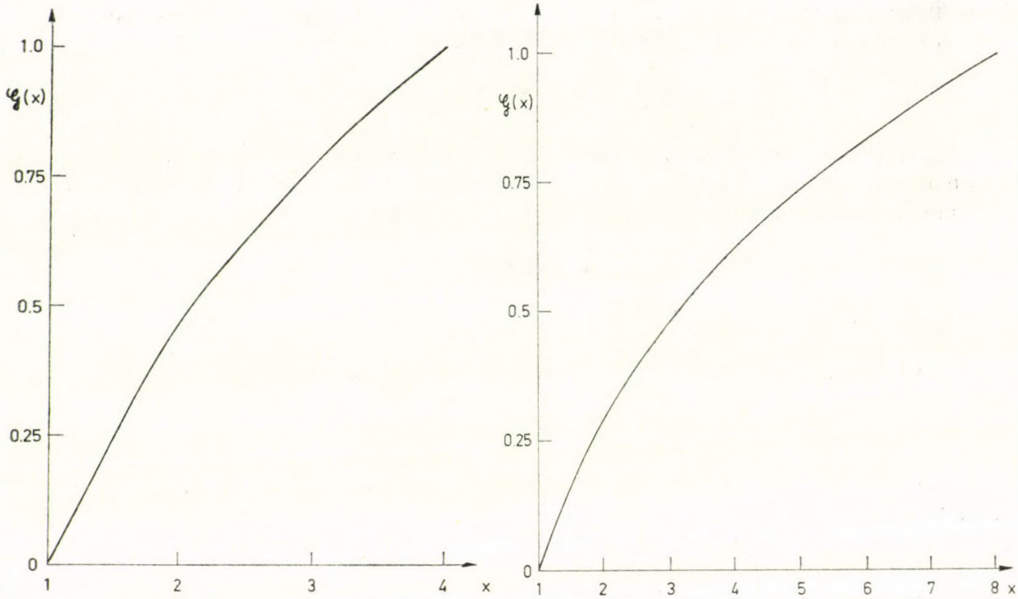


Fig. 1. The function $\mathcal{G}(x)$ for $f(x)=(1+x)^{1/m}-1$. The invariant density for the transformation $Tx=(1+x)^m-[(1+x)^m]$ is $\mathcal{G}'(x^{1/m}-1)$ where prime denotes differentiation. (a) the case $m=2$, (b) the case $m=3$.

and determine it on $2^m-1\leqq x\leqq 2^m$ by (4.3). However this will not be a monotonous non-decreasing function in $1\leqq x\leqq 2^m$ even if we choose it so in $1\leqq x\leqq 2^m-1$.

Iterations of equation (3.13) converge very rapidly. Numerical solutions for the cases $m=2$ and $m=3$ are shown on Figures 1a and 1b. They are convex from below as they should.

5. β-expansions

In the discussion above we restricted ourselves to the case that the X appearing in conditions (1.2) or (1.6), if finite, is an integer. One particular case not satisfying this condition, namely $f(x)=x/3$, $\beta>1$, β not an integer, has been extensively studied (see [1]–[4]). With this $f(x)$ every real number x, $0\leqq x<1$, is represented as the "β-expansion"

$$x=\frac{\varepsilon_1(x)}{\beta}+\frac{\varepsilon_2(x)}{\beta^2}+\dots \tag{5.1}$$

where $\varepsilon_1(x), \varepsilon_2(x), \dots$ are integers and $0\leqq\varepsilon_j(x)<\beta$. The effect of non-integral β is that not all possible sequences of integers less than β correspond to a number.

For the related transformation of the interval (0, 1) to itself, namely

$$Tx=(\beta x)\equiv \text{fractional part of } \beta x,$$

RÉNYI [3] showed that an invariant measure $h(x)$ exists, and if normalized

$$\int_0^1 h(x)\,dx = 1, \tag{5.2}$$

it satisfies the inequalities

$$1-\frac{1}{\beta} \leqq h(x) \leqq \left(1-\frac{1}{\beta}\right)^{-1}. \tag{5.3}$$

PARRY [4] derived for $h(x)$ the equation,

$$\beta h(x) = \sum_{j=0}^{[\beta+x]} h\left(\frac{x+j}{\beta}\right), \tag{5.4}$$

equivalent to our (3.7) and used it to study some of its properties. However, the overall constant $F(\beta)$ naturally arising in his investigation has peculiar properties. For example, $F(\beta)$ is discontinuous on points belonging to an everywhere dense set and it tends to ∞ as $\beta\to 1$.

Our discussion in section 3 above does not require that every possible sequence of integers should correspond to a number. So we can apply the same reasoning to this particular case and determine the invariant measure $h(x)$ numerically for any given β. Equation (3.11) is now replaced by

$$G(\beta x) = \begin{cases} \sum_{j=0}^{[\beta]} \{G(x+j)-G(j)\}, & \text{if} \quad 0 \leqq x \leqq \beta-[\beta], \\ \sum_{j=0}^{[\beta]-1} \{G(x+j)-G(j)\}+G(\beta)-G([\beta]), & \text{if} \quad \beta-[\beta] \leqq x \leqq 1 \end{cases} \tag{5.5}$$

while (3.12) becomes

$$G(0) = 0, \quad G(\beta) = 1. \tag{5.6}$$

To get a numerical solution for $G(x)$ we may again take a non-decreasing $G_0(x)$ satisfying (5.6), but otherwise arbitrary, and determine $G_n(x)$ successively for $n=1, 2, \dots$ by the equation

$$G_n(\beta x) = \begin{cases} \sum_{j=0}^{[\beta]} \{G_{n-1}(x+j)-G_{n-1}(j)\}, & 0 \leqq x \leqq \beta-[\beta] \\ \sum_{j=0}^{[\beta]-1} \{G_{n-1}(x+j)-G_{n-1}(j)\}+1-G_{n-1}([\beta]), & \beta-[\beta] \leqq x \leqq 1 \end{cases}$$

Since the normalization (5.6) is maintained through each iteration, it does not cause any trouble as it did to PARRY [4]. The convergence of $G_n(x)$ as $n\to\infty$ is assured by the arguments of RÉNYI [3], $\lim_{n\to\infty} G_n(x) = G(x)$. Again $G(x)$ is convex from below.

As an example we give the graph of $G(x)$ for $\beta=\pi$ on Figure 2. In the terminology of PARRY [4] π is not "β-simple", nor does it contain a recurring tail in its "β-expansion" with $\beta=\pi$. Therefore (cf. [4], p. 304, corollary to Theorem 2). $G(x)$ does not consist of a finite number of straight line segments. However, it may

Fig. 2. The function $G(x)$ for $f(x)=\pi/x$, $\pi\cong 3.1416$. The invariant density for the transformation $Tx=\pi x-[\pi x]$ is the derivative of $G(x)$.

possibly have an infinite number of straight line segments of varying lengths and this will be hard to recognize from the numerical values of $G(x)$.

I am thankful to my colleagues, G. Baker, J. des Cloizeaux and Y. Pomeau in particular, for many helpful discussions.

References

[1] B. H. Bissinger, A generalization of continued fractions, *Bull. Amer. Math. Soc.*, **50** (1944), 868—876.
[2] C. J. Everett, Representations for real numbers, *ibid.*, **52** (1946), 861—869.
[3] A. Rényi, Representations for real numbers and their ergodic properties, *Acta Math. Acad. Sci. Hungar.*, **8** (1957), 477—493.
[4] W. Parry, On the β-expansion of real numbers, *ibid.*, **11** (1960), 401—406.
[5] P. Levy, *Bull. Soc. Math. France,* **57** (1929), 178—194.
[6] A. Denjoy, *C. R. Acad. Sc. Paris,* **202** (1936), 371—374, 537—540.
[7] W. Bolyai, *Tentamen iuventutem studiosam in elementa matheseos purae elementaris ac sublimioris methodo intuitiva evidentiaque huic propria introducendi,* ed. sed. (Budapest, 1897), vol. 1 (as referred to in Rényi's paper, ref. [3] above).

(Received August 29, 1977)

S.P.T., C.E.N. DE SACLAY
91190 GIF-SUR-YVETTE
FRANCE

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 101—103.

ELATIONS OF GRAPHS

By
R. WEISS (Berlin)

Let Γ be an undirected graph with vertex set V and edge set E and let G be a subgroup of aut (Γ). For each $x \in V$ we denote by $\Gamma(x)$ the set of vertices adjacent to x and by $G(x)^{\Gamma(x)}$ the permutation group induced by the stabilizer $G(x)$ of x on $\Gamma(x)$. For each $x \in V$ and each natural number i let $G_i(x) = \{a \in G \mid a \in G(u)$ for each $u \in V$ with $\partial(x, u) \leqq i\}$, where $\partial(x, u)$ denotes the distance between x and u, and $G_i(x, y) = G_i(x) \cap G_i(y)$ for each $y \in \Gamma(x)$.

In [2, (2.3)] GARDINER proved the following theorem, earlier versions of which may be attributed to THOMPSON [5] and WIELANDT [7, (6.6)]; see also [4, (2.1)].

THEOREM 1. *Let Γ be a connected, finite, undirected graph with vertex set V, $\{x, y\}$ an edge of Γ and G a subgroup of* aut (Γ) *acting transitively on V such that $G(x)^{\Gamma(x)}$ is primitive. Then $|G_1(x, y)|$ is divisible by at most one prime.*

In this paper we prove the following generalization of Theorem 1:

THEOREM 2. *Let Γ be a connected, finite, undirected graph, $\{x, y\}$ an edge of Γ and G a subgroup of* aut (Γ) *such that $G(u)^{\Gamma(u)}$ is primitive for $u = x$ and y. Let π be the set of primes dividing $|G_1(x, y)|$. Then either $|\pi| \leqq 1$ or there exists a $p \in \pi$ such that for either $u = y$ and $v = y$ or $u = y$ and $v = x$, $O^p(G_1(x, y)) \leqq G_2(u)$, $G_2(v)$ is a p-group and $G_2(v) = G_3(v)$.*

We recall that $O^p(G_1(x, y))$ denotes the intersection of all the normal subgroups of $G_1(x, y)$ of index a power of p.

Our proof of Theorem 2 is completely elementary and self-contained. Note that Theorem 2 asserts that if $|\pi| > 1$ then $G_2(v)$ but not $G_2(u)$ is a p-group and so G acts intransitively on V; thus Theorem 2 is in fact a generalization of Theorem 1. To show that $|\pi| > 1$ can actually occur, we take any graph Δ with a vertex-transitive automorphism group G such that for each vertex x, $G(x)^{\Delta(x)}$ is primitive and $|G_1(x)|$ divisible by more than one prime (for instance, the incidence graph of the projective plane $PG(2, q)$ with $q \neq 2$) and let Γ be the graph obtained from Δ by inserting a new vertex at the midpoint of each edge of Δ. Then the pair (Γ, G) fulfils the hypotheses of Theorem 2 with $|\pi| > 1$. Unfortunately, I do not know of any examples without vertices of valency two.

If we let Γ be the incidence graph of a projective plane $\mathcal{P}$ and $G = \text{aut}(\Gamma)$, then the nontrivial elements of the subgroups $G_1(x, y)$, $\{x, y\} \in E$, are precisely the elations of $\mathcal{P}$. Thus Theorems 1 and 2 are related to the result [3, (1.3)]; in particular, when Γ is an arbitrary undirected graph, it seems natural to refer to the nontrivial elements of the subgroups $G_1(x, y)$ where $G = \text{aut}(\Gamma)$ and $\{x, y\} \in E$ as elations of Γ.

Theorem 2 should play a role in the study of locally transitive graphs (see [1], [6]) where it is not assumed that G acts transitively on V.

We begin the proof. Let (Γ, G) be a pair fulfilling the hypotheses of Theorem 2 and let $K=\langle G(x), G(y)\rangle$. For each $w\in V$ let $V(w)=\{z\in V|\partial(w, z)$ is even$\}$. We claim the K acts transitively on E, $V(x)$ and $V(y)$; in particular, K acts transitively on V if Γ is not bipartite. To see this, let $(x_0, x_1, x_2, \dots, x_s)$ be an arbitrary path in Γ with $x_0=x$, $x_1=y$ and s even. There exists an element $a_1\in G(x_1)$ mapping x_0 to x_2. Thus $G(x_2)=a_1^{-1}G(x_0)a_1\leqq K$ and so $K(x_2)$ contains an element a_2 mapping x_1 to x_3. Continuing, we obtain elements $a_i\in K(x_i)$ for $i=3, \dots, s-1$ mapping x_{i-1} to x_{i+1}. Letting $a=a_1\dots a_{s-1}$, we have $(x)a=x_s$ and $(y)a=x_{s-1}$. Since Γ is connected, the claim follows.

Let $H=G(x, y)$ where $G(x, y)=G(x)\cap G(y)$. We may assume $H\neq 1$. Let A be a minimal normal subgroup of H. Suppose $A\not\leqq G_1(x)$. Choose a vertex $z\in\Gamma(x)$ and an element $a\in A$ not in $G(z)$. Since $G_1(x)\trianglelefteq H$, we have $A\cap G_1(x)=1$ and thus $[A, G_1(x)]=1$. In particular, $[a, G_1(x)]=1$ so that $G_2(z)=a^{-1}G_2(z)a=G_2((z)a)$ and hence $G_2(z)\trianglelefteq\langle G(x, z), G(x, (z)a)\rangle$. Since $A\not\leqq G_1(x)$, $G(x)^{\Gamma(x)}$ is not regular. Since $G(x)^{\Gamma(x)}$ is primitive, it follows that $\langle G(x, z), G(x, (z)a)\rangle=G(x)$. Thus $G_2(z)\trianglelefteq\langle G(x), G(z)\rangle$. Since $\langle G(x), G(z)\rangle$ acts transitively on E, $G_2(z)=1$ and thus $G_2(y)=1$. Since $[a, G_1(x, z)]\leqq[A, G_1(x)]=1$, we also have $G_1(x, z)=G_1(x, (z)a)$ and thus $G_1(x, z)\trianglelefteq\langle G(x, z), G(x, (z)a)\rangle=G(x)$; it follows that $G_1(x, y)\leqq G_2(x)$ (and in fact $G_1(x, y)=G_2(x)$). Thus $A\not\leqq G_1(x)$ implies that G fulfils the conclusions of Theorem 2 with $u=x$ and $v=y$. We may therefore assume that $A\leqq G_1(x)$ and similarly $A\leqq G_1(y)$ for every minimal normal subgroup A of H.

Let A be a minimal normal subgroup of H and B a minimal normal subgroup of $G_1(x)$ contained in A. Let $B^*=\langle h^{-1}Bh\,|\,h\in H\rangle$. Then $B^*\leqq A$ and $B^*\trianglelefteq H$; thus $B^*=A$. Since each $h^{-1}Bh$ is a minimal normal subgroup of $G_1(x)$, it follows that $A=B^*\leqq\operatorname{soc}(G_1(x))$ where $\operatorname{soc}(G_1(x))$ denotes the subgroup of $G_1(x)$ generated by all the minimal normal subgroups of $G(x)$. Thus $\operatorname{soc}(H)\leqq\operatorname{soc}(G_1(x))$. Similarly, $\operatorname{soc}(H)\leqq\operatorname{soc}(G_1(y))$.

If $\operatorname{soc}(G_1(x))=\operatorname{soc}(H)=\operatorname{soc}(G_1(y))$, then $\operatorname{soc}(H)\trianglelefteq\langle G(x), G(y)\rangle$ and thus $\operatorname{soc}(H)=1$. Hence we conclude that $\operatorname{soc}(H)\not\leqq\operatorname{soc}(G_1(w))$ for $w=x$ or y. Let B be a minimal normal subgroup of $G_1(w)$ not contained in $\operatorname{soc}(H)$. Let A be a minimal normal subgroup of H contained in $B^*=\langle h^{-1}Bh|h\in H\rangle$. Since B is minimal, either $A\cap B=1$ or $B\leqq A$. Since, however, $A\leqq\operatorname{soc}(H)$ and $B\not\leqq\operatorname{soc}(H)$, we must have $A\cap B=1$ and thus $[A, B]=1$. Hence $1=h^{-1}[A, B]h=[A, h^{-1}Bh]$ for each $h\in H$ and so $[A, B^*]=1$. Since $A\leqq B^*$, A is abelian. Since A is a minimal normal subgroup, $|A|$ must be a power of some $p\in\pi$. In particular, $O_p(H)\neq 1$ where $O_p(H)$ denotes the subgroup of H generated by all the normal subgroups of H of order a power of p.

If $O_p(H)\leqq G_1(x, y)$, then $O_p(G_1(x))=O_p(H)=O_p(G_1(y))$ and thus $O_p(H)\trianglelefteq\langle G(x), G(y)\rangle$. Since $O_p(H)\neq 1$, we conclude that $O_p(H)\not\leqq G_1(u)$ with $u=x$ or y. Choose a vertex $z\in\Gamma(u)$ and an element $a\in O_p(H)$ not in $G(z)$. Let $M=O^p(G_2(z))$ or $O^p(G_1(u, z))$ and set $N=a^{-1}Ma$. Then $N\trianglelefteq G_1(u)$ and MN/N is a subgroup of $(MO_p(H)\cap G_1(u))/N$. Since $O^p(M)=M$ and $|MO_p(H)|/|N|$ is a power of p, it follows that $M=N$. Thus $M\trianglelefteq\langle G(u, z), G(u, (z)a)\rangle$. Since $O_p(H)\not\leqq G_1(u)$, $G(u)^{\Gamma(u)}$ is not regular and hence $\langle G(u, z), G(u, (z)a)\rangle=G(u)$. Therefore $O^p(G_2(z))\trianglelefteq G(u)$ and $O^p(G_1(u, z))\trianglelefteq G(u)$. In particular, if $O_p(H)\not\leqq G_1(w)$

for $w=x$ *and* y then $O^p(G_1(x, y)) \trianglelefteq \langle G(x), G(y) \rangle$ and thus $O^p(G_1(x, y))=1$. Hence we may assume that $O_p(H) \leqq G_1(v)$ for $v \in \{x, y\}$, $v \neq u$. Since $O^p(G_1(u, z)) \trianglelefteq \trianglelefteq G(u)$, $O^p(G_1(u, z)) \leqq G_2(u)$. Since $O^p(G_2(z)) \trianglelefteq G(u)$, we have $O^p(G_2(z)) \trianglelefteq \trianglelefteq \langle G(u), G(z) \rangle$ and hence $O^p(G_2(z))=1$. Let $z_1 \in \Gamma(z)$ and $z_2 \in \Gamma(z_1)-\{z\}$ be arbitrary. Since $G_2(z) \trianglelefteq G_1(z_1) \trianglelefteq G(z_1, z_2)$ and $G_2(z)$ is a p-group, we have $G_2(z) \leqq \leqq O_p(G(z_1, z_2))$. Since $O_p(G(z_1, z_2)) \leqq G_1(z_2)$, $G_2(z) \leqq G_1(z_2)$ and thus $G_2(z)=G_3(z)$.

References

[1] I. Z. Bouwer and D. Ž. Djoković, On regular graphs III, *J. Comb. Th.*, **B14** (1973), 268—277.
[2] A. Gardiner, Arc transitivity in graphs, *Quart. J. Math. Oxford,* **(2)24** (1973), 399—407.
[3] A. M. Gleason, Finite Fano planes, *Amer. J. Math.*, **78** (1956), 797—707.
[4] W. Knapp, On the point stabilizer in a primitive permutation group, *Math. Z.*, **133** (1973), 137—168.
[5] J. G. Thompson, Bounds for orders of maximal subgroups, *J. Algebra,* **14** (1970), 135—138.
[6] R. Weiss, Groups with a (B, N)-pair and locally transitive graphs, *Nagoya Math. J.*, **74** (1979), 1—21.
[7] H. Wielandt, *Subnormal subgroups and permutation groups,* Lecture Notes, The Ohio State University (1971).

(Received August 29, 1977)

II. MATH. INSTITUT DER FU BERLIN
KÖNIGIN-LUISE-STR. 24/26
1000 BERLIN 33

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 105—116.

ON THE MEASURABLE SOLUTION OF A FUNCTIONAL EQUATION ARISING IN INFORMATION THEORY

By

Z. DARÓCZY and A. JÁRAI (Debrecen)

§ 1. Introduction

In the characterization problems of the measure of information the following concepts are well-known and play an important role (see J. ACZÉL and Z. DARÓCZY [2]). Let

$$\Gamma_n := \left\{(p_1, p_2, \dots, p_n):\ p_i \geqq 0,\ \sum_{i=1}^{n} p_i = 1\right\}$$

denote the set of discrete probability distributions consisting of n numbers and let

$$I_n\colon \Gamma_n \to \mathbf{R} \quad (n = 2, 3, \dots) \tag{1.1}$$

be a sequence of unknown functions, where $\mathbf{R}$ denotes the set of real numbers. We say that the sequence (1.1) has the *sum property* if there exists a function $f\colon [0, 1]\to\mathbf{R}$ such that

$$I_n(p_1, p_2, \dots, p_n) = \sum_{i=1}^{n} f(p_i) \tag{1.2}$$

for every $(p_1, p_2, \dots, p_n)\in\Gamma_n$ and $n=2, 3, \dots$. In this case the function $f\colon [0, 1]\to\mathbf{R}$ is said to be the *generating function* of the sequence $\{I_n\}$ having the sum property. Let $k\geqq 2$ and $l\geqq 2$ be fixed natural numbers. The sequence (1.1) is said to be (k, l)-additive if for every $(p_1, p_2, \dots, p_k)\in\Gamma_k$ and $(q_1, q_2, \dots, q_l)\in\Gamma_l$ we have

$$\begin{aligned} I_{kl}(p_1q_1, \dots, p_1q_l;\ p_2q_1\ \dots, p_2q_l; \dots; p_kq_1, \dots, p_kq_l) &= \\ = I_k(p_1, \dots, p_k)+I_l(q_1, \dots, q_l). \end{aligned} \tag{1.3}$$

T. W. CHAUNDY and J. B. MCLEOD [3] have proved the following result: If $I_n\colon \Gamma_n\to\mathbf{R}$ $(n=2, 3, \dots)$ has the sum property with the continuous generating function $f\colon [0, 1]\to\mathbf{R}$ and it is (k, l)-additive for every $k\geqq 2$ and $l\geqq 2$, then there exists $c\in\mathbf{R}$ such that

$$I_n(p_1, p_2, \dots, p_n) = cH_n(p_1, p_2, \dots, p_n) \tag{1.4}$$

holds for every $(p_1, p_2, \dots, p_n)\in\Gamma_n$ $(n=2, 3, \dots)$. Here

$$H_n(p_1, p_2, \dots, p_n) := -\sum_{i=1}^{n} p_i \log_2 p_i \tag{1.5}$$

is the Shannon entropy of the distribution $(p_1, p_2, \dots, p_n)\in\Gamma_n$ $(0 \log_2 0:=0)$.

J. ACZÉL and Z. DARÓCZY [1] have shown that the preceding theorem remains valid supposing only (k, k)-additivity for each $k\geqq 2$. A further reduction of the conditions can be found in [4]. In this respect the strongest result obtained up to the present is the following one (see Z. DARÓCZY [6] and J. ACZÉL—Z. DARÓCZY [2]): If $I_n\colon \Gamma_n \to \mathbf{R}$ $(n=2, 3, \dots)$ has the sum property and it is (2, 3)-additive with the measurable generating function $f\colon [0, 1] \to \mathbf{R}$ satisfying $f(1)=0$, then (1.4) holds for every $(p_1, p_2, \dots, p_n) \in \Gamma_n$ $(n=2, 3, \dots)$.

The following natural question seems not to have been investigated: What can we say about the (2, 2)-additive sequences $I_n\colon \Gamma_n \to \mathbf{R}$ having the sum property if the generating function has some additional property? In this paper this problem is completely solved in case of measurable generating function. In our investigations we rely on the method of A. JÁRAI worked out in [9] and [10] which is very useful in solving problems connected with other functional equations.

§ 2. A general functional equation

Let X be a nonvoid set and $(Y, \mathcal{B}, m)$ a finite measure space. Let Δ denote the class of all sets $D \subset X \times Y$ for which $\inf_{x \in X} m(D_x) > 0$ where

$$D_x := \{y\colon\ y \in Y,\ (x, y) \in D\}$$

is measurable for every $x \in X$.

Let $D \in \Delta$ be fixed and let $\Gamma(D)$ denote the set of functions $g\colon D \to Y$ with the following property: for every $\varepsilon > 0$ there exists $\delta > 0$ such that $E \in \mathcal{B}$ and $m(E) < \delta$ imply

$$E_x^{(g)} := \{y\colon\ y \in D_x,\ g(x, y) \in E\} \in \mathcal{B}$$

and $m(E_x^{(g)}) < \varepsilon$ for every $x \in X$.

The next result plays a fundamental role in our investigations.

LEMMA 1. *Let $D \in \Delta$ and $g_i \in \Gamma(D)$ $(i=1, 2, \dots, n)$. Then there exists a $\delta > 0$ such that $E \in \mathcal{B}$ and $m(E) < \delta$ imply that*

$$\bigcap_{i=1}^{n} C_{D_x} E_x^{(g_i)} \neq \emptyset$$

for every $x \in X$, where $C_A B$ denotes the complement of the set $B \subseteq A$ with respect to the set A.

PROOF. Let $\varepsilon := \inf_{x \in X} m(D_x) > 0$. Then $g_i \in \Gamma(D)$ implies the existence of a $\delta > 0$ for which $E_x^{(g_i)} \in \mathcal{B}$ and $m(E_x^{(g_i)}) < \varepsilon/n$ whenever $E \in \mathcal{B}$, $m(E) < \delta$, $x \in X$ $(i=1, 2, \dots, n)$. Suppose, that there exists $x_0 \in X$ such that

$$\bigcap_{i=1}^{n} C_{D_{x_0}} E_{x_0}^{(g_i)} = \emptyset.$$

Then

$$D_{x_0} = C_{D_{x_0}} \bigcap_{i=1}^{n} C_{D_{x_0}} E_{x_0}^{(g_i)} = \bigcup_{i=1}^{n} E_{x_0}^{(g_i)}$$

implies that

$$\varepsilon \leqq m(D_{x_0}) = m\left(\bigcup_{i=1}^{m} E_{x_0}^{(g_i)}\right) \leqq \sum_{i=1}^{n} m(E_{x_0}^{(g_i)}) < n\frac{\varepsilon}{n} = \varepsilon$$

which is a contradiction.

THEOREM 1. *Let* $D\in\Delta$, $g_i\in\Gamma(D)$ *and* $a_i\in\mathbf{R}$ $(i=1, 2, \dots, n)$. *Suppose that the functions* $F\colon X\to\mathbf{R}$ *and* $f\colon Y\to\mathbf{R}$ *satisfy the functional equation*

$$F(x) = \sum_{i=1}^{n} a_i f[g_i(x, y)] \tag{2.1}$$

for every $(x, y)\in D$. *If* f *is measurable on the set* Y, *then* F *is bounded on the set* X.

PROOF. By the measurability of f

$$\lim_{n\to\infty} m\{y\colon y\in Y,\ |f(y)|>k\}=0.$$

Let $\delta>0$ chosen as in Lemma 1. Then there exists $K>0$ for which the set

$$E:=\{y\colon y\in Y,\ |f(y)|>K\}$$

is measurable and $m(E)<\delta$. By Lemma 1 for every $x\in X$ there exists $y_x\in\bigcap_{i=1}^{n} C_{D_x}E_x^{(g_i)}$, that is $y_x\in C_{D_x}E_x^{(g_i)}$ $(i=1, 2, \dots, n)$. As

$$C_{D_x}E_x^{(g_i)} = C_{D_x}\{y\colon y\in D_x,\ g_i(x,y)\in E\} = \{y\colon y\in D_x,\ g_i(x,y)\notin E\} =$$
$$= \{y\colon y\in D_x,\ |f[g_i(x, y)]|\leqq K\} \quad (i=1, 2, \dots, n),$$

for every $x\in X$ there exists $y_x\in D_x$ (that is $(x, y_x)\in D$), for which by (2.1)

$$|F(x)| = \left|\sum_{i=1}^{n} a_i f[g_i(x, y_x)]\right| \leqq \sum_{i=1}^{n} |a_i|\,|f[g_i(x, y_x)]| \leqq \sum_{i=1}^{n} |a_i| K =: K^*,$$

i.e. F is bounded on X.

§ 3. Classes of functions in $\Gamma(D)$

The next result is well-known (see e.g. E. HEWITT—K. STROMBERG [8]).

LEMMA 2. *Let* $\varphi\colon\ ]a, b[\to\mathbf{R}$ *be a one-to-one and differentiable function on* $]a, b[$, *for which there exists* $\varrho>0$ *with the property* $|\varphi'(t)|\leqq\varrho$ *for every* $t\in]a, b[$. *Then for every Lebesgue-measurable set* $E\subseteqq]a, b[$ *the set* $\varphi(E):=\{\varphi(t)\colon t\in E\}$ *is Lebesgue-measurable and* $m[\varphi(E)]\leqq\varrho m(E)$ *where* m *denotes the Lebesgue measure.*

In the next part of this paper let $X:=[\alpha, \beta]$, where $0<\alpha<\beta<1$ and $Y:=\ ]0, 1[$. Let $\mathscr{B}$ denote the class of all Lebesgue-measurable sets of Y and m the Lebesgue-measure on $\mathscr{B}$.

Let

$$D:=\{(x, y)\colon x\in X,\ x<y<1\}.$$

Clearly, the set D belongs to the class Δ (see Fig. 1).

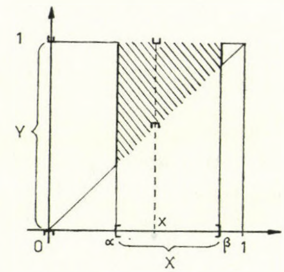


Fig. 1

LEMMA 3. *Let* $g\colon D\to Y$ *be a function with the property that for every* $x\in X$ *the function* $g(x,\cdot)\colon D_x\to Y$ $(D_x=]x,1[)$ *is one-to-one and differentiable on* D_x *and* $|D_2g(x,y)|\geqq\varrho$ *for every* $(x,y)\in D$, *where* $\varrho>0$. *Then* $g\in\Gamma(D)$.

REMARK. Here $D_2g(x,y)$ denotes the partial derivative of g with respect to the second variable (see the notations of J. DIEUDONNÉ [7]).

PROOF. For every fixed $x\in X$ let $\gamma(y):=g(x,y)$ $(y\in D_x)$. By the condition on g, γ is one-to-one and differentiable on D_x, and so γ has an inverse function $\varphi:=\gamma^{-1}\colon \gamma(D_x)\to Y$ which is also one-to-one and differentiable on the open interval $\gamma(D_x)$ and $|\varphi'(t)|\leqq 1/\varrho$ for every $t\in\gamma(D_x)$. If $E\in\mathscr{B}$, then by Lemma 2

$$m(E_x^{(g)})=m\{y\colon y\in D_x,\ g(x,y)\in E\}=m\{y\colon y\in D_x,\ \gamma(y)\in E\}=$$

$$=m\{\gamma^{-1}(t)\colon t\in E\cap\gamma(D_x)\}=m\{\gamma^{-1}[E\cap\gamma(D_x)]\}\leqq\frac{1}{\varrho}m\{E\cap\gamma(D_x)\}\leqq\frac{1}{\varrho}m(E),$$

which implies $g\in\Gamma(D)$.

§ 4. Examples

Let the functions $g_i\colon D\to Y$ $(i=1,2,\dots,7)$ be defined as follows:

$$g_1(x,y):=\frac{x}{y}-x;\quad g_2(x,y):=y-x;\quad g_3(x,y):=1-\frac{x}{y}-y+x;$$

$$g_4(x,y):=\frac{x}{y};\quad g_5(x,y):=1-\frac{x}{y};\quad g_6(x,y):=y;\quad g_7(x,y):=1-y.$$

It is easy to see that the functions g_1, g_2, g_4, g_5, g_6 and g_7 satisfy the conditions of Lemma 3, hence they belong to $\Gamma(D)$. Now let $g(x,y):=g_3(x,y)$.

LEMMA 4. $g\in\Gamma(D)$.

PROOF. It is obvious that for every $x\in X$ and $E\in\mathscr{B}$ the set $E_x^{(g)}$ belongs to $\mathscr{B}$ (see Fig. 2).

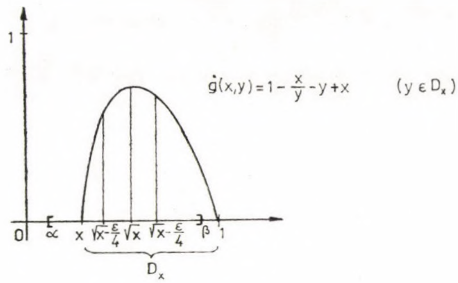


Fig. 2

Let $\varepsilon > 0$ be arbitrary. Without loss of generality suppose that $\varepsilon < 4\sqrt{\alpha}\,(1-\sqrt{\beta})$. For fixed $x \in X$ let

$$I_x^1 := \left]x, \sqrt{x}-\frac{\varepsilon}{4}\right[, \quad I_x^2 := \left]\sqrt{x}+\frac{\varepsilon}{4}, 1\right[, \quad I_x^3 := \left[\sqrt{x}-\frac{\varepsilon}{4}, \sqrt{x}+\frac{\varepsilon}{4}\right].$$

Then for every $E \in \mathscr{B}$

$$E_x^{(g)} = \{y\colon\, y \in D_x,\ g(x, y) \in E\} = A_1 \cup A_2 \cup A_3,$$

where

$$A_i := \{y\colon\, y \in I_x^i,\ g(x, y) \in E\} \quad (i = 1, 2, 3).$$

The sets A_i are measurable and disjoint and $m(A_3) \leqq \varepsilon/2$. If $y \in I_x^1 \cup I_x^2$ then

$$|D_2\, g(x, y)| = \left|\frac{x}{y^2} - 1\right| = \frac{|\sqrt{x}-y|\,|\sqrt{x}+y|}{y^2} \geqq \frac{\varepsilon}{4}\alpha > 0.$$

Hence with the notation $\gamma(y) := g(x, y)$ we get

$$\left|\frac{d}{dy}\gamma^{-1}(y)\right| \leqq \frac{4}{\varepsilon\alpha}$$

which implies, by Lemma 2,

$$m(A_1 \cup A_2) = m(A_1) + m(A_2) = m[\gamma^{-1}(E \cap I_x^1)] +$$

$$+ m[\gamma^{-1}(E \cap I_x^2)] \leqq \frac{4}{\varepsilon\alpha}[m(E \cap I_x^1) + m(E \cap I_x^2)] \leqq \frac{8}{\varepsilon\alpha} m(E).$$

I.e. for $m(E) < \delta := \varepsilon^2\alpha$ we have

$$m(A_1 \cup A_2) < \frac{8}{\varepsilon\alpha}\frac{\varepsilon^2\alpha}{16} = \frac{\varepsilon}{2}.$$

Summarizing we have that $m(E) < \delta$ implies $m(E_x^{(g)}) < \varepsilon$, that is $g \in \Gamma(D)$.

§ 5. Measurable solutions

Let $f\colon]0,1[\to\mathbf{R}$ be an unknown function for which the functional equation

(5.1)
$$f(pq)+f(p(1-q))+f((1-p)q)+f((1-p)(1-q))=f(p)+f(1-p)+f(q)+f(1-q)$$

holds for every $p, q\in]0,1[$. With the substitutions $x:=pq$ and $y:=q$ we obtain from (5.1) that

$$f(x)=-f\left(\frac{x}{y}-x\right)-f(y-x)-f\left(1-\frac{x}{y}-y+x\right)+f\left(\frac{x}{y}\right)+f\left(1-\frac{x}{y}\right)+f(y)+f(1-y)$$
(5.2)

holds for every $x\in]0,1[$ and $x<y<1$. Now let $X:=[\alpha,\beta]\subset]0,1[$ be an arbitrary fixed closed interval, then from (5.2) with the notations $F(t):=f(t)$ $(t\in X)$ and $a_1=a_2=a_3=-1$, $a_4=a_5=a_6=a_7=1$ we get

$$F(x)=\sum_{i=1}^{7}a_i f[g_i(x,y)] \tag{5.3}$$

for every $(x,y)\in D$ where by our former investigations $g_i\in\Gamma(D)$ $(i=1,2,\dots,7)$. By Theorem 1 we have the following

THEOREM 2. *If $f\colon]0,1[\to\mathbf{R}$ is a measurable solution of the functional equation* (5.1) *then f is bounded on each closed interval* $[\alpha,\beta]\subset]0,1[$.

This implies the next important

COROLLARY. *If $f\colon]0,1[\to\mathbf{R}$ is a measurable solution of the functional equation* (5.1) *then f is Lebesgue integrable on each closed interval* $[\alpha,\beta]\subset]0,1[$.

The following result strongly depends on some basic facts of integration theory.

THEOREM 3. *If $f\colon]0,1[\to\mathbf{R}$ is a measurable solution of the functional equation* (5.1) *then f is infinitely differentiable on* $]0,1[$.

PROOF. (i) In the first step we show that f is continuous on $]0,1[$. Let $x_0\in]0,1[$ be arbitrary fixed and choose the numbers λ,μ,α,β such that $0<\alpha<x_0<\beta<1$ and $0<\sqrt{\beta}<\lambda<\mu<1$ (Fig. 3).

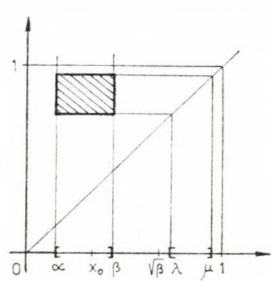


Fig. 3

Then $[\alpha, \beta]\times[\lambda, \mu]\subset D$ and by (5.3)

$$f(x) = \sum_{i=1}^{7} a_i f[g_i(x, y)] \tag{5.4}$$

holds for every $(x, y)\in[\alpha, \beta]\times[\lambda, \mu]$. Let us define the following sets:

$$J_i := \{g_i(x, y)\colon (x, y)\in[\alpha, \beta]\times[\lambda, \mu]\} \quad (i = 1, 2, \dots, 7).$$

Clearly J_i is a closed interval in $]0, 1[$ for every i.

Now let $x\in[\alpha, \beta]$ be fixed and vary y in $[\lambda, \mu]$. Then $g_i(x, \cdot)$ maps the closed interval onto a closed interval, on which f is integrable by the corollary of Theorem 2. Hence by (5.4)

$$\int_{\lambda}^{\mu} f(x)\,dy = \sum_{i=1}^{7} a_i \int_{\lambda}^{\mu} f[g_i(x, y)]\,dy = \sum_{i=1}^{7} a_i \int_{g_i(x,\lambda)}^{g_i(x,\mu)} f(t) D_2 h_i(x, t)\,dt,$$

where $y=h_i(x, t)$ is the function defined by the condition $g_i(x, y)=t$. This is uniquely determined by the invertibility of $g_i(x, \cdot)$ on $[\lambda, \mu]$. Using this we have

$$f(x) = \sum_{i=1}^{7} A_i \int_{\gamma_i(x)}^{\Gamma_i(x)} f(t) H_i(x, t)\,dt \tag{5.5}$$

for each $x\in[\alpha, \beta]$ where

$$A_i := \frac{a_i}{\mu-\lambda}, \quad \gamma_i(x) := g_i(x, \lambda), \quad \Gamma_i(x) := g_i(x, \mu), \quad H_i(x, t) := D_2 h_i(x, t)$$
$$(i = 1, 2, \dots, 7).$$

Now we show that the functions

$$Q_i(x) := \int_{\gamma_i(x)}^{\Gamma_i(x)} f(t) H_i(x, t)\,dt \quad (x\in[\alpha, \beta],\ i = 1, 2, \dots, 7)$$

are continuous at x_0. By the boundedness of f on J_i and by the continuity of H_i there exist positive constants K_i, L_i $(i=1, 2, \dots, 7)$ such that $|f(t)|\leqq K_i$, $|H_i(x, t)|\leqq L_i$ whenever $x\in[\alpha, \beta]$ and $t\in J_i$. Thus for every $x\in[\alpha, \beta]$

$$|Q_i(x)-Q_i(x_0)| = \Big| \int_{\gamma_i(x)}^{\Gamma_i(x)} f(t) H_i(x, t)\,dt - \int_{\gamma_i(x_0)}^{\Gamma_i(x_0)} f(t) H_i(x_0, t)\,dt\Big| \leqq$$

$$\leqq \Big| \int_{\gamma_i(x)}^{\alpha_i(x_0)} f(t) H_i(x, t)\,dt\Big| + \Big| \int_{\Gamma_i(x_0)}^{\Gamma_i(x)} f(t) H_i(x, t)\,dt\Big| + \Big| \int_{\gamma_i(x_0)}^{\Gamma_i(x_0)} f(t)[H_i(x, t)-H_i(x_0, t)]\,dt\Big| \leqq$$

$$\leqq |\gamma_i(x)-\gamma_i(x_0)|\,K_i L_i + |\Gamma_i(x)-\Gamma_i(x_0)|\,K_i L_i + |\Gamma_i(x_0)-\gamma_i(x_0)|\,K_i \sup_{t\in J_i} H_i(x,t)-H_i(x_0,t)|$$

which implies

$$\lim_{x\to x_0} |Q_i(x)-Q_i(x_0)| = 0;$$

i.e. Q_i is continuous at x_0. Hence by (5.5) f is continuous at x_0 and so f continuous in $]0, 1[$.

(ii) Next we show that f is differentiable in $]0, 1[$. By (i), f is continuous in $]0, 1[$, therefore Q_i is differentiable on the closed interval $[\alpha, \beta]$ and by the theorem concerning the differentiation of parametric integrals (see J. DIEUDONNÉ [7]) (5.5) implies

$$f'(x) = \sum_{i=1}^{7} A_i \Big[\int_{\gamma_i(x)}^{\Gamma_i(x)} f(t) D_1 H_i(x, t)\, dt +$$

$$+ f(\Gamma_i(x)) H_i(x, \Gamma_i(x)) \Gamma_i'(x) - f(\gamma_i(x)) H_i(x, \gamma_i(x)) \gamma_i'(x)\Big]$$

for every $x \in [\alpha, \beta]$. The right side of the former equation is differentiable for the same reason, that is f is twice differentiable. Continuing this process we have that f is infinitely differentiable.

§ 6. Differentiable solutions

Consider the functional equation

$$(6.1) \qquad f(pq) + f((1-p)(1-q)) + f(p(1-q)) + f((1-p)q) =$$
$$= f(p) + f(1-p) + f(q) + f(1-q)$$

valid for every $p, q \in]0, 1[$, where $f\colon]0, 1[\to \mathbf{R}$ is an unknown function.

LEMMA 5. *If* $f\colon]0, 1[\to \mathbf{R}$ *is a twice differentiable solution of* (6.1) *in* $]0, 1[$ *then the function*

$$(6.2) \qquad f_1(t) := f'(t) + t f''(t) \quad (t \in]0, 1[)$$

satisfies the functional equation

$$(6.3) \qquad f_1(pq) + f_1((1-p)(1-q)) - f_1(p(1-q)) - f_1((1-p)q) = 0$$

for every $p, q \in]0, 1[$.

PROOF. Differentiating (6.1) with respect to p, then the resulting equation with respect to q, we have (6.3) with the notation (6.2).

LEMMA 6. *If* $f\colon]0, 1[\to \mathbf{R}$ *is a solution of the functional equation* (6.3) *and it is twice differentiable in* $]0, 1[$, *then the function*

$$(6.4) \qquad f_2(t) := f_1'(t) + t f_1''(t) \quad (t \in]0, 1[)$$

satisfies the equation

$$(6.5) \qquad f_2(pq) + f_2((1-p)(1-q)) + f_2(p(1-q)) + f_2(q(1-p)) = 0$$

for every $p, q \in]0, 1[$.

PROOF. Differentiating (6.3) with respect to p, then the resulting equation with respect to q, we have (6.5) with the notation (6.4).

LEMMA 7. *If* f_2:]0, 1[$\to\mathbf{R}$ *is a solution of* (6.5) *and it is twice differentiable in*]0, 1[*then*

$$f_2''(t) = 0 \quad (t\in]0, 1[). \tag{6.6}$$

PROOF. Let $q=\frac{1}{2}$ in (6.5) then

$$f_2\left(\frac{p}{2}\right)+f_2\left(\frac{1-p}{2}\right) = 0 \quad (p\in]0, 1[)$$

therefore

$$f_2''\left(\frac{p}{2}\right)+f_2''\left(\frac{1-p}{2}\right) = 0 \quad (p\in]0, 1[) \tag{6.7}$$

is valid. Let us differentiate the equation (6.5) twice with respect to q then we have

$$p^2f_2''(pq)+(1-p)^2f_2''((1-p)(1-q))+p^2f_2''(p(1-q))+(1-p)^2f_2''(q(1-p)) = 0 \tag{6.8}$$

for every $p, q\in]0, 1[$. Let $q=\frac{1}{2}$ in (6.8) then we have

$$p^2f_2''\left(\frac{p}{2}\right)+(1-p)^2f_2''\left(\frac{1-p}{2}\right) = 0 \quad (p\in]0, 1[)$$

which by (6.7) implies

$$f_2''\left(\frac{p}{2}\right)(2p-1) = 0$$

for every $p\in]0, 1[$. Hence for $p\neq\frac{1}{2}$, $p\in]0, 1[$

$$f_2''\left(\frac{p}{2}\right) = 0. \tag{6.9}$$

On the other hand, from (6.7) we see that (6.9) holds for $p=\frac{1}{2}$ as well. Hence

$$f_2''(t) = 0 \quad \text{for} \quad t\in\left]0, \frac{1}{2}\right[. \tag{6.10}$$

Now let $t\in\left[\frac{1}{2}, 1\right[$ arbitrary. Then there exist $p, q\in]0, 1[$ such that $t=pq$, and by $p>\frac{1}{2}$ and $q>\frac{1}{2}$ we have that $(1-p)(1-q)$, $p(1-q)$, $q(1-p)$ belong to $\left]0, \frac{1}{2}\right[$, i.e. from (6.8) by (6.10), $f_2''(t)=f_2''(pq)=0$. Thus we have proved (6.6) for every $t\in]0, 1[$.

THEOREM 4. *Let* f:]0, 1[$\to\mathbf{R}$ *be a solution of the functional equation* (6.1) *and suppose that* f *is six times differentiable in*]0, 1[. *Then there exist constants* $a_i\in\mathbf{R}$ $(i=0, 1, 2, 3)$ *and* $b_k\in\mathbf{R}$ $(k=1, 2)$ *such that*

$$f(t) = a_3t^3+a_2t^2+a_1t+a_0+b_1t\ln t+b_2\ln t \tag{6.11}$$

for every $t\in]0, 1[$.

PROOF. Using the former notations, we have by Lemma 7 $f_2''(t)=0$ ($t\in]0,1[$), hence

$$f_2(t) = At+B \quad (t\in]0,1[;\ A,B\in\mathbf{R}).$$

By Lemma 6,

$$f_2(t) = [tf_1'(t)]' = At+B,$$

hence

$$f_1(t) = \frac{A}{4}t^2+Bt+C\ln t+D \quad (C,D\in\mathbf{R}).$$

Finally, by Lemma 5

$$f_1(t) = [tf'(t)]' = \frac{A}{4}t^2+Bt+C\ln t+D,$$

hence

$$f(t) = \frac{A}{36}t^3+\frac{B}{4}t^2+Ct\ln t+(E-C)\ln t+Dt+F-C$$

for every $t\in]0,1[$, where $E,F\in\mathbf{R}$. This implies (6.11).

THEOREM 5. *Let* $f\colon]0,1[\to\mathbf{R}$ *be a measurable solution of the functional equation* (6.1). *Then*

$$f(t) = A(4t^3-9t^2+5t)+Bt\ln t+C \tag{6.12}$$

holds for every $t\in]0,1[$ *where* A, B, C *are real constants. All functions of the form* (6.12) *satisfy* (6.1).

PROOF. By Theorem 3, f is infinitely differentiable in $]0,1[$; hence by Theorem 4 f has the form (6.11). Substituting (6.11) in (6.1) it is easy to see that the function $t\to a_0+b_1t\ln t$ satisfies (6.1) and therefore $b_2=0$. Thus it is sufficient to know in which case does the cubic polynomial

$$\varphi(t) := a_3t^3+a_2t^2+a_1t$$

satisfy equation (6.1). By simple calculation we obtain

$$\varphi(pt)+\varphi\big((1-p)t\big) = (3a_3t^3+2a_2t^2)(p^2-p)+\varphi(t),$$

from which

$$\varphi(pq)+\varphi\big((1-p)q\big)+\varphi\big(p(1-q)\big)+\varphi\big((1-p)(1-q)\big) =$$
$$= (p^2-p)[3a_3q^3+2a_2q^2+3a_3(1-q)^3+2a_2(1-q)^2]+\varphi(q)+\varphi(1-q).$$

Hence φ satisfies (6.1) if and only if

$$(p^2-p)[3a_3q^3+2a_2q^2+3a_3(1-q)^3+2a_2(1-q)^2] = (p^2-p)(3a_3+2a_2)+a_3+a_2+a_1$$

holds for every $p, q\in]0,1[$. This implies

$$a_3+a_2+a_1 = 0. \tag{6.13}$$

On the other hand, dividing by (p^2-p) shows that

$$(q^2-q)(9a_3+4a_2) = 0,$$

i.e.

$$9a_3+4a_2 = 0. \tag{6.14}$$

This and (6.13) imply that φ satisfies (6.1) if and only if

$$\varphi(t) = 4At^3-9At^2+5At$$

where $A:=\frac{a_3}{4}$. This means that f is of the form (6.12).

§ 7. Entropies

Let $\alpha>0$ and $\alpha\neq 1$. It is well-known that the quantities

$$H_n^\alpha(p_1, p_2, \dots, p_n) := (2^{1-\alpha}-1)^{-1}\left(\sum_{i=1}^{n} p_i^\alpha-1\right) \tag{7.1}$$

defined for all $(p_1, p_2, \dots, p_n)\in\Gamma_n$ $(n=2, 3, \dots)$ are called entropies of type α (Z. DARÓCZY [4], J. ACZÉL and Z. DARÓCZY [2]). With the definition

$$H_n^1(p_1, p_2, \dots, p_n) := H_n(p_1, p_2, \dots, p_n) \tag{7.2}$$

we have

$$\lim_{\alpha\to 1} H_n^\alpha(p_1, p_2, \dots, p_n) = H_n^1(p_1, p_2, \dots, p_n)$$

i.e. the Shannon entropy can be considered to be of type 1.

Our preceding results can be summarized in the next interesting theorem.

THEOREM 6. *Let* $I_n\colon \Gamma_n\to\mathbf{R}$ $(n=2, 3, \dots)$ *have them sum-property, be* (2, 2)-*additive, and have a measurable generating function. Then there are real constants* A, B, C, D *such that for every*

$$(p_1, p_2, \dots, p_k, 0, \dots, 0)\in\Gamma_n \quad (p_i>0,\ i=1, 2, \dots, k\colon\ k\geqq 2,\ n\geqq 2,\ k\leqq n)$$

the equation

$$I_n(p_1, p_2, \dots, p_k, 0, \dots, 0) = -3AH_k^3(p_1, p_2, \dots, p_k)+ \tag{7.3}$$

$$+\frac{9}{2}AH_k^2(p_1, p_2, \dots, p_k)-BH_k^1(p_1, p_2, \dots, p_k)+kC+(n-k)D$$

holds. For $(1, 0, \dots, 0)\in\Gamma_n$ $(n\geqq 2)$ *we have*

$$I_n(1, 0, \dots, 0) = nD. \tag{7.4}$$

Conversely, for all real constants A, B, C, D *the sequence* $\{I_n\}$ *defined by equations* (7.3) *and* (7.4) *has the sum property and is* (2, 2)-*additive.*

PROOF. Let $f\colon [0, 1]\to\mathbf{R}$ be the generating function of the sequence $\{I_n\}$. By the (2, 2)-additivity of $\{I_n\}$, f satisfies (5.1) (or (6.1)) on the open interval $]0, 1[$.

As f is measurable in $]0, 1[$, we obtain from Theorem 5 that

$$f(t) = \begin{cases} 4At^3 - 9At^2 + 5At + Bt\log_2 t + C, & \text{if} \quad t \in]0, 1[\\ D, \quad \text{if} \quad t = 0 \quad \text{or} \quad 1 \end{cases} \tag{7.5}$$

since (2, 2)-additivity implies $f(1) = f(0)$. As $I_n(p_1, p_2, \dots, p_n)$ is a symmetric function in p_i, an arbitrary element of Γ_n can be written in the form

$$(p_1, p_2, \dots, p_k, 0, \dots, 0) \in \Gamma_n \quad (k \geqq 2;\ p_i > 0,\ i = 1, \dots, k)$$

or $(1, 0, \dots, 0) \in \Gamma_n$. Hence in the first case by (7.5)

$$I_n(p_1, p_2, \dots, p_k, 0, \dots, 0) = \sum_{i=1}^{k} f(p_i) + (n-k) f(0) =$$

$$= 4A \sum_{i=1}^{k} p_i^3 - 9A \sum_{i=1}^{n} p_i^2 + 5A \sum_{i=1}^{k} p_i + B \sum_{i=1}^{k} p_i \log_2 p_i + kC + (n-k) D$$

which implies (7.3) according to the notations (7.1) and (7.2). In the second case we clearly get (7.4).

By Theorem 6 we can say that (2, 2)-additive sequences with the sum-property in the presence of a very weak condition (e.g. measurability) can be obtained essentially by the linear combination of entropies of types 1, 2 and 3.

References

[1] J. Aczél und Z. Daróczy, Charakterisierung der Entropien positiver Ordnung und der Shannonschen Entropie, *Acta Math. Acad. Sci. Hungar.*, **14** (1963), 95—121.

[2] J. Aczél and Z. Daróczy, *On measures of information and their characterizations,* Academic Press (New York—San Francisco—London, 1975).

[3] T. W. Chaundy and J. B. McLeod, On a functional equation, *Proc. Edinburgh Math. Soc.*, **(2)12** (*Edinburgh Math. Notes,* **43**, 7—8) (1960).

[4] Z. Daróczy, Über die Charakterisierung der Shannonschen Entropie. *Proc. Colloq. Information Theory* (Debrecen, 1967) **1**, 135—139. J. Bolyai Math. Soc. (Budapest, 1968).

[5] Z. Daróczy, Generalized information functions. *Information and Control,* **16** (1970), 36—51.

[6] Z. Daróczy, On the measurable solutions of a functional equation, *Acta Math. Acad. Sci. Hungar.*, **22** (1971), 11—14.

[7] J. Dieudonné, *Foundations of modern analysis,* Academic Press (New York and London, 1960).

[8] E. Hewitt—K. Stromberg, *Real and abstract analysis,* Springer-Verlag (Berlin—Heidelberg—New York, 1965).

[9] A. Járai, *Ph. D. thesis.* University Debrecen, 1976.

[10] A. Járai, On measurable solutions of functional equations, *Publ. Math. Debrecen,* **26** (1979), 17—35.

(Received August 29, 1977)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF L. KOSSUTH,
4010 DEBRECEN, P. F. 12.
HUNGARY

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 117—119.

A NOTE ON MAPPING POLYDISCS INTO BALLS AND VICE VERSA

By

L. LEMPERT (Budapest)

1. Let B_n and D^n denote the unit ball and unit polydisc in $\mathbf{C}^n$, respectively;

$$B_n = \left\{z = (z_1, \dots, z_n) \in \mathbf{C}^n \colon |z|^2 = \sum_1^n |z_j|^2 < 1\right\},$$

$$D^n = \{z = (z_1, \dots, z_n) \in \mathbf{C}^n \colon |z_j| < 1 \text{ for } j = 1, 2, \dots, n\}.$$

More generally, let

$$rB_n = \{rz\colon z \in B_n\}, \quad rD^n = \{rz\colon z \in D^n\}.$$

For a domain $U \subset \mathbf{C}^n$ a mapping $\Phi\colon U \to \mathbf{C}^n$ is called biholomorphic if it is holomorphic and injective. In this case the image $\Phi(U)$ is a domain, too, and the inverse mapping Φ^{-1} is holomorphic on $\Phi(U)$.

In [1], J. E. FORNAESS and E. L. STOUT considered biholomorphic mappings Φ of D^n *into* B_n (though in a much more general setting). Poincaré's theorem tells us that the image cannot be the whole B_n (at least if $n > 1$, which will be assumed throughout this paper). The above authors proved, however, that the image can be very large in one sense. (The measure of $B_n \setminus \Phi(D^n)$ can be zero.) At the same time, it cannot be too large in some other sense; indeed, they proved the existence of an $R = R(n) < 1$, such that $\Phi(D^n)$ cannot contain any ϱB_n with $\varrho > R$. Their proof was indirect and did not give any effective value for R. Of course, $R \geqq 1/\sqrt{n}$, since the image of the mapping

$$D^n \ni z \mapsto z/\sqrt{n} \in B_n$$

actually covers $B_n/\sqrt{n}$.

The authors of [1] have raised the very natural problem of finding the minimal $R(n)$. In this note we shall prove that $R(n) = 1/\sqrt{n}$ is the minimal value; in other words, the following holds:

THEOREM 1. *If Φ is a biholomorphic mapping of D^n into B_n and $\Phi(D^n)$ contains ϱB_n with some $\varrho > 0$, then $\varrho \leqq 1/\sqrt{n}$.*

The dual of this will be proved, too:

THEOREM 2. *If Φ maps B_n biholomorphically into D^n and $\Phi(B_n)$ covers ϱD^n with some $\varrho > 0$, then $\varrho \leqq 1/\sqrt{n}$.*

2. Proof of Theorem 1. To avoid multiindices, we are going to treat the case $n=2$ only; z_1, z_2 will be substituted by x and y, respectively, and Φ will have the components f and g.

We can assume that $\Phi(0)=0$. Indeed, some $(\xi, \eta)\in D^2$ must be mapped on $(0, 0)$, and then

$$\Psi(x, y) = \Phi\left(\frac{\xi-x}{1-x\bar{\xi}}, \frac{\eta-y}{1-y\bar{\eta}}\right),$$

while having the same range as Φ, would keep the origin fixed.

Now, if $\Phi(0)=0$, then

$$f(x, y) = \sum{}' a_{jk}x^j y^k, \quad g(x, y) = \sum{}' b_{jk}x^j y^k,$$

where $\sum{}'$ indicates that summation is taken over nonnegative j-s and k-s, $j=k=0$ excluded.

Since $|f|^2+|g|^2<1$ throughout D^2, we have for any $r<1$

$$4\pi^2 > \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\left(|f(re^{i\varphi}, re^{i\psi})|^2+|g(re^{i\varphi}, re^{i\psi})|^2\right) d\varphi\, d\psi =$$

$$= 4\pi^2 \sum{}' (|a_{jk}|^2+|b_{jk}|^2)\, r^{2(j+k)},$$

and, therefore

$$\sum{}' (|a_{jk}|^2+|b_{jk}|^2) \leqq 1. \tag{1}$$

On the other hand, Φ, being biholomorphic, is proper, too; that is, if $z^m\in D^2$ $(m=1, 2\ldots)$ and $z^m\to\partial D^2$, then $\Phi(z^m)\to\partial\Phi(D^2)$. It is evident that $\partial\Phi(D^2)$ lies outside ϱB_2: hence for $\varphi\in[-\pi, \pi]$

$$\liminf_{r\to 1} |\Phi(re^{i\varphi}, 0)| \geqq \varrho,$$

i.e.:

$$\liminf_{r\to 1}\left(|f(re^{i\varphi}, 0)|^2+|g(re^{i\varphi}, 0)|^2\right) \geqq \varrho^2.$$

f and g being bounded,

$$\liminf_{r\to 1} \int_{-\pi}^{\pi} \left(|f(re^{i\varphi}, 0)|^2+g(re^{i\varphi}, 0)|^2\right) d\varphi \geqq 2\pi\varrho^2$$

follows. Therefore

$$\sum_{j=1}^{\infty} (|a_{j0}|^2+|b_{j0}|^2) \geqq \varrho^2.$$

Similarly,

$$\sum_{k=1}^{\infty} (|a_{0k}|^2+|b_{0k}|^2) \geqq \varrho^2.$$

These two, added up and compared with (1), yield the conclusion of our theorem.

3. Proof of Theorem 2. The two theorems are equivalent via the Schwarz Lemma. One can reason like that:

Suppose $\Phi(B_n)\subset D^n$ covers ϱD^n. This means that Φ^{-1} is defined and biholomorphic on ϱD^n and $\Phi^{-1}(\varrho D^n)\subset B_n$. On the other hand, by virtue of the Schwarz Lemma, $\Phi(\varrho B_n)\subset \varrho D^n$, that is, $\Phi^{-1}(\varrho D^n)\supset \varrho B_n$. Thus, Φ^{-1} maps a polydisc into B_n in such a way that the image of the polydisc covers ϱB_n. By Theorem 1 ϱ is at most $1/\sqrt{n}$, then.

Theorem 1 could be applied since the size of our polydisc is of no importance; at least as far as biholomorphic mappings are concerned, all polydiscs are equal.

References

[1] J. E. Fornaess, E. L. Stout, Polydiscs in complex manifolds, *Math. Ann.*, **227** (1977), 145—154.

(Received September 2, 1977)

DEPARTMENT I OF ANALYSIS
EÖTVÖS LORÁND UNIVERSITY
1088 BUDAPEST, MÚZEUM KRT. 6—8.

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 121—127.

CONTINUOUS PATHS AND HOMOLOGY

By
M. BOGNÁR (Budapest)

This paper deals with a connection between the *i*-category of the continuous paths of a Hausdorff space (see [2], 2.8) and the homology groups of the bodies of these paths. It is closely related to the category-homomorphisms (see [1], 6.1).

1. Let **R** be the real line space. Denote by T_0 or by I the unit interval subspace of **R** and by $\dot{T}_0$ the discrete two-point space $\{0, 1\}$. For any closed interval $[b, c]$ of **R** where $b<c$ let $[b, c]^{\#}=(T, \dot{T})$ where T is the subspace of **R** with the underlying set $[b, c]$ and $\dot{T}$ is the discrete two-point space $\{b, c\}\cdot[b, c]^{\#}=(T, \dot{T})$ is a compact pair.

We can consider $[b, c]$ as a closed 1-simplex in the Euclidean 1-space **R** and $\{b, c\}$ is the frontier of $[b, c]$. $[b, c]$ is at the same time a 1-cell and its frontier is a 0-sphere.

Remark that for any two closed intervals $[b, c]$ and $[b', c']$ of **R** and for any two strictly monotone increasing surjective functions $s\colon [b, c]\to[b', c']$ and $s'\colon[b, c]\to[b', c']$ we have obviously $s(b)=s'(b)=b'$, $s(c)=s'(c)=c'$ and s and s' are continuous. Moreover, s and s' are homotopic maps rel $\{b, c\}$.

In fact let $h\colon [b, c]^{\#}\times I\to[b', c']^{\#}$ be defined by

$$h(x, t) = (1-t).s(x)+t.s'(x).$$

h establishes clearly a homotopy of s to s' rel $\{b, c\}$.

The strictly monotone increasing surjective functions of the type $s\colon [b, c]\to[b', c']$ appear later several times. They will be called *simple functions*.

2. Let p be a prime and G an elementary cyclic p-group, i.e. G is isomorphic to the group J_p of integers mod p. Let H be a continuous homology theory defined on the category $\mathcal{A}_C$ of all compact pairs and all continuous maps of such pairs and based on the coefficient group G. Thus H is isomorphic on $\mathcal{A}_C$ to the Čech homology theory over G.

The group G can be considered as a compact commutative topological group. Hence the homology theory H is exact (see [3], p. 248).

3. The group $H_1(T_0, \dot{T}_0)$ is isomorphic to G (see [3], p. 45). Let u be a fixed nonzero element of $H_1(T_0, \dot{T}_0)$. Let $[b, c]$ be a closed interval in **R**, choose a simple function $s\colon [0, 1]\to[b, c]$ and put $[b, c]_*=s_*(u)$ where s_* is the homomorphism $s_*\colon H_1([0, 1]^{\#})\to H_1([b, c]^{\#})$ induced by s. As we have seen in Section 1, all the maps s of this kind are homotopic. Hence the homomorphism s_* and the member $[b, c]_*$ of $H_1([b, c]^{\#})$ do not depend on the special choice of s.

s is a topological map, consequently s_* is an isomorphism and this yields $[b, c]_* \neq 0$.

Observe that if $b=0$ and $c=1$ then s is homotopic to the identical map of $[0, 1]^\#$. Thus $[0, 1]_* = s_*(u) = u$.

It is to be noted that $[b, c]_*$ depends obviously on the special choice of $u=[0, 1]_*$. However this dependence can be described very easily. Given two non-zero elements $[0, 1]_*$ and $[0, 1]_{*'}$ of $H_1(T_0, \dot{T}_0)$ there exists evidently an integer m relative prime to p such that $[0, 1]_{*'} = m[0, 1]_*$ and then for any closed interval $[b, c]$ of $\mathbf{R}$ $[b, c]_{*'} = m[b, c]_*$ holds.

4. *Let* $[b, c]$ *and* $[b', c']$ *be closed intervals in* $\mathbf{R}$ *and* $s\colon [b, c] \to [b', c']$ *a simple function. Then*

$$s_*([b, c]_*) = [b', c']_*.$$

In fact let $s'\colon [0, 1] \to [b, c]$ be a simple function. Then $ss'\colon [0, 1] \to [b', c']$ is simple as well. Consequently

$$[b', c']_* = (ss')_*([0, 1]_*) = s_* s'_*([0, 1]_*) = s_*([b, c]_*).$$

5. *Let* $[b, c]$ *and* $[b', c']$ *be two closed intervals in* $\mathbf{R}$ *such that* $[b, c] \subset [b', c']$. *Moreover let* $[b, c]^\# = (T, \dot{T})$, $[b', c')^\# = (T', \dot{T}')$ *and* $Q = T' \setminus \operatorname{int} T = T' \setminus (T \setminus \dot{T})$. *Let* $i\colon (T, \dot{T}) \to (T', Q)$ *and* $i'\colon (T', \dot{T}') \to (\dot{T}', Q)$ *be injections. Then*

$$i_*([b, c]_*) = i'_*([b', c']_*).$$

Proof. Let $s\colon [b', c'] \to [b, c]$ be a simple function and define $h\colon [b', c']^\# \times I \to (T', Q)$ by

$$h(x, t) = (1-t).\, s(x) + t.x, \quad (x, t) \in [b', c'] \times I.$$

h establishes a homotopy of $i\,s$ to i', where s can be considered as a map $s\colon [b', c']^\# \to [b, c]^\#$. Hence $i\,s$ and i' are homotopic maps, consequently $(i\,s)_* = i_* s_* = i'_*$, and this yields by Section 4

$$i'_*([b', c']_*) = i_* s_*([b', c']_*) = i_*([b, c]_*)$$

as required.

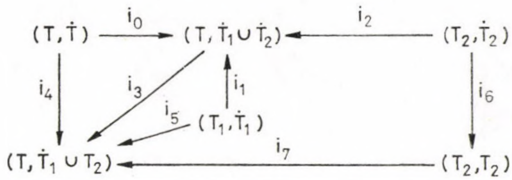
6. *Let* $b<c<d$ $(b, c, d \in \mathbf{R})$ *and* $[b, c]^\# = (T_1, \dot{T}_1)$, $[c, d]^\# = (T_2, \dot{T}_2)$, $[b, d]^\# = (T, \dot{T})$. *Let* $i_1\colon (T_1, \dot{T}_1) \to (T, \dot{T}_1 \cup \dot{T}_2)$, $i_2\colon (T_2, \dot{T}_2) \to (T, \dot{T}_1 \cup \dot{T}_2)$ *and* $i_0\colon (T, \dot{T}) \to (T, \dot{T}_1 \cup \dot{T}_2)$ *be injections. Then*

$$i_{0*}([b, d]_*) = i_{1*}([b, c]_*) + i_{2*}([c, d]_*).$$

Proof. $K = \{[b, c], [c, d], \{b\}, \{c\}, \{d\}\}$ is a 1-complex, where $|K| = |K^1| = T$ and $|K^0| = \dot{T}_1 \cup \dot{T}_2$. ($K^i$ is the i-skeleton of K $(i=0, 1)$). Thus $i_{0*}([b, d]_*) \in H_1(|K^1|, |K^0|) = H_1(T, \dot{T}_1 \cup \dot{T}_2)$ has a unique decomposition in the form

$$i_{0*}([b, d]_*) = i_{1*}(u_1) + i_{2*}(u_2) \tag{1}$$

where $u_i \in H_1(T_i, \dot{T}_i)$ $(i=1, 2)$ (see [3], p. 85.). Taking the commutative diagram



where each i_j is an injection and applying Section 5 and $H_1(T_2, T_2)=0$ we obtain

$$i_{5*}([b, c]_*) = i_{4*}([b, d]_*) = i_{3*}i_{0*}([b, d]_*) =$$

$$= i_{3*}i_{1*}(u_1)+i_{3*}i_{2*}(u_2) = i_{5*}(u_1)+i_{7*}i_{6*}(u_2) = i_{5*}(u_1).$$

However the injection $i_5\colon (T_1, \dot{T}_1)\to(T, \dot{T}_1\cup T_2)$ is a relative homeomorphism and so i_{5*} is an isomorphism (see [3], p. 266), consequently $u_1=[b, c]_*$. Similarly $u_2=[c, d]_*$. Hence we have by (1) the statement.

7. In the remainder we shall use the notions and notations of [2].

In particular let Y be a fixed Hausdorff space. Let $k\colon [b, c]\to Y$ be a nondegenerated continuous line in Y, i.e. a continuous map of the interval $[b, c]$ into Y. The body of k — it is indicated by $\tilde{k}$ — is the set $k[b, c]$ endowed with the subspace topology of Y. Let $\dot{\tilde{k}}$ be the discrete space with the underlying set $\{k(b), k(c)\}$. It is important that k can be considered as a map $k\colon [b, c]^\# \to(\tilde{k}, \dot{\tilde{k}})$ and thus it induces a homomorphism $k_*\colon H_1([b, c]^\#)\to H_1(\tilde{k}, \dot{\tilde{k}})$. Let $k_{**}=k_*([b, c]_*)$.

8. *Let* $k\colon [b, c]\to Y$ *be a nondegenerated continuous line. Then* $k^\bullet_{**}=-k_{**}$ ($k^\bullet$ is defined by $k^\bullet(x)=k(b+c-x)$).

In fact, define $g\colon [b, c]\to[b, c]$ by $g(x)=b+c-x$. Then $g(b)=c$, $g(c)=b$, consequently the automorphism g of the 1-simplex $[b, c]$ is odd. Hence for any $v\in H_1([b, c^\#])$ $g_*(v)=-v$ (see [3], p. 82) and this yields

$$k^\bullet_{**} = k^\bullet_*([b, c]_*) = (kg)_*([b, c]_*) = k_*g_*([b, c]_*) = k_*(-[b, c]_*) = -k_{**}.$$

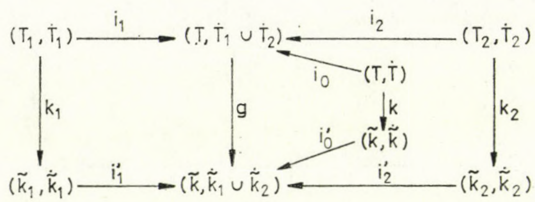
9. *Let* $k_1\colon [b, c]\to Y$ *and* $k_2\colon [c, d]\to Y$ *be nondegenerated continuous lines such that* $k_1(c)=k_2(c)$. *Let* $k=k_1\circ k_2$

$$\left(k_1\circ k_2\colon [b, d]\to Y \quad \text{is defined by} \quad k_1\circ k_2(x)=\begin{cases} k_1(x) & \text{if} \quad b\leqq x\leqq c \\ k_2(x) & \text{if} \quad c\leqq x\leqq d\end{cases}\right).$$

Let $i_1'\colon (\tilde{k}_1, \dot{\tilde{k}}_1)\to(\tilde{k}, \dot{\tilde{k}}_1\cup\dot{\tilde{k}}_2)$, $i_2'\colon (\tilde{k}_2, \dot{\tilde{k}}_2)\to(\tilde{k}, \dot{\tilde{k}}_1\cup\dot{\tilde{k}}_2)$ *and* $i_0'\colon (\tilde{k}, \dot{\tilde{k}})\to(\tilde{k}, \dot{\tilde{k}}_1\cup\dot{\tilde{k}}_2)$ *be injections. Then*

$$i'_{0*}(k_{**}) = i'_{1*}(k_{1**})+i'_{2*}(k_{2**}).$$

PROOF. Let $[b, c]^\# =(T_1, \dot{T}_1)$, $[c, d]^\# =(T_2, \dot{T}_2)$ and $[b, d]^\# =(T, \dot{T})$. Define the map $g\colon (T, \dot{T}_1\cup\dot{T}_2)\to(\tilde{k}, \dot{\tilde{k}}_1\cup\dot{\tilde{k}}_2)$ by $g(x)=k(x)$ $(x\in[b, d])$. Taking the commutative diagram



where i_0, i_1, i_2 are injections too and applying Section 6 we obtain

$$i'_{0*}(k_{**}) = i'_{0*}k_*([b, d]_*) = g_* i_{0*}([b, d]_*) =$$

$$= g_*\big(i_{1*}([b, c]_*) + i_{2*}([c, d]_*)\big) = i'_{1*}k_{1*}([b, c]_*) + i'_{2*}k_{2*}([c, d]_*) = i'_{1*}(k_{1**}) + i'_{2*}(k_{2**})$$

as required.

10. *Let* $k\colon [b, c] \to Y$ *and* $k'\colon [b', c'] \to Y$ *be two nondegenerated equivalent continuous lines, i.e. there exists a simple function* $s\colon [b, c] \to [b', c']$ *such that* $k = k's$. *Then* $k_{**} = k'_{**}$.

In fact, Section 4 shows that

$$k_{**} = k_*([b, c]_*) = k'_* s_*([b, c]_*) = k'_*([b', c']_*) = k'_{**}.$$

11. We now consider continuous paths in Y and assign to each closed one a well defined member of the 1-dimensional homology group of its body.

In particular let K be a nondegenerated continuous path in Y. That means K is an equivalence class (in the sense of Section 10) of nondegenerated continuous lines in Y. The body of $k\colon [b, c] \to Y$, the initial point $k(b)$ and the closing point $k(c)$ of k do not depend on the special choice of $k \in K$. These notions will be indicated by $\tilde{K}$, $\triangle(K)$, $\triangledown(K)$ respectively (see [2]). Let $\dot{\tilde{K}}$ be the discrete two-point or one-point space $\{\triangle(K), \triangledown(K)\}$. If $\triangle(K) = \triangledown(K)$ then K is a *closed path* and $\triangle(K)$ is said to be the *base point* of K.

Take now any $k \in K$, and put $K_{**} = k_{**}$. According to Section 10 K_{**} is well defined. Moreover $(\tilde{k}, \dot{\tilde{k}}) = (\tilde{K}, \dot{\tilde{K}})$.

Suppose now that K is a closed path and let $j_K\colon (\tilde{K}, \emptyset) \to (\tilde{K}, \dot{\tilde{K}})$ be an injection. Since $\dot{\tilde{K}}$ is a singleton it is homologically trivial, consequently $j_{K*}\colon H_1(\tilde{K}) \to H_1(\tilde{K}, \dot{\tilde{K}})$ is an isomorphism ([3], p. 23). Let now $K_* = (j_{K*})^{-1}(K_{**})$. We define K_* also in the degenerate case, i.e. when k is a degenerated line for any $k \in K$ (that is a map of any singleton of $\mathbf{R}$ into Y). In this case $\tilde{K}$ is a one-point space. Whence $H_1(\tilde{K}) = 0$, and we choose then for K_* the 0-element of this $H_1(\tilde{K})$.

12. K_* depends obviously on the special choice of $[0, 1]_*$. However if we take two nonzero elements $[0, 1]_*$ and $[0, 1]_{*'}$ of $H_1([0, 1]^{\#})$ then according to Section 3 there exists an integer m such that for any closed interval $[b, c]$ in $\mathbf{R}$ $[b, c]_{*'} = m[b, c]_*$. Hence we have for every nondegenerated continuous line k in Y $k_{**'} = mk_{**}$, This implies $K_{**'} = mK_{**}$ for every nondegenerated continuous path K in Y, and thus for each closed K in Y $K_{*'} = mK_*$. This last equality is evidently true also for degenerated closed paths.

13. It is to be noted that for any closed Jordan path K of Y (see [2], 5.9) $K_* \neq 0$.

In fact let k: $[b, c] \to Y \in K$. Then k: $[b, c]^\# \to (\tilde{K}, \dot{\tilde{K}})$ is a relative homeomorphism, thus k_* is an isomorphism and this yields, by $[b, c]_* \neq 0$, $K_{**} = k_{**} = k_*([b, c]_*) \neq 0$. Whence

$$K_* = (j_{K*})^{-1}(K_{**}) \neq 0.$$

We are formulating now three important properties of the operation $K \mapsto K_*$.

14. *Let K be a continuous path in Y* (K need not be closed). *Then* $(KK^\bullet)_* = 0$. (Recall [2], 2.6 and 2.5.)

Proof. First of all we formulate and prove the following

Lemma. *Let (X, A) be a compact pair such that A is a one-point or a two-point space. Let j: $(X, \emptyset) \to (X, A)$ be an injection. Then j_*: $H_1(X) \to H_1(X, A)$ is a monomorphism.*

In fact $H_1(A) = 0$ (see [3], pp. 45 and 46.) Hence j_* is monomorphic by the exactness of the sequence

$$H_1(X, A) \xleftarrow[j_*]{} H_1(X) \leftarrow H_1(A).$$

We are going now to prove $(KK^\bullet)_* = 0$.

If K is a degenerated path, then the assertion is obviously true. Suppose that K is nondegenerated. Let k_1: $[b, c] \to Y \in K$ and k_2: $[c, d] \to Y \in K^\bullet$ (see also [2], 2.3). Then $k = k_1 \circ k_2 \in KK^\bullet$.

We shall use the notations of Section 9. Since $\tilde{k}$ coincides here with $\tilde{k}_1$ and $\tilde{k}_2$ moreover $\dot{\tilde{k}}_1$ coincides with $\dot{\tilde{k}}_2$ it follows that i_1' and i_2' are identities, consequently i_{1*}' and i_{2*}' are also identical maps. However k_2 is equivalent to $k_1^\bullet$. Hence we have according to Sections 8 and 10 $k_{2**} = k^\bullet_{1**} = -k_{1**}$ and thus by Section 9

$$(i_0' j_{KK^\bullet})_*((KK^\bullet)_*) = i_{0*}'(j_{KK^\bullet *}((KK^\bullet)_*) = i_{0*}'((KK^\bullet)_{**}) = i_{0*}'(k_{**}) = k_{1**} - k_{1**} = 0.$$

By the lemma $(i_0' j_{KK^\bullet})_*$ is a monomorphism, consequently $(KK^\bullet)_* = 0$ holds as required.

15. *Let K_1 and K_2 be continuous paths in Y such that both of the products $K_1 K_2$ and $K_2 K_1$ exist.* (Thus $K_1 K_2$ and $K_2 K_1$ are closed paths.) *Then* $(K_1 K_2)_* = (K_2 K_1)_*$.

Proof. If some of the K_i-s $(i = 1, 2)$ are degenerated then $K_1 K_2 = K_2 K_1$ and the assertion is obviously true.

Suppose now that neither K_1 nor K_2 is degenerated. Let $b < c < d < e$ ($b, c, d, e \in \mathbf{R}$). Let $k_1 \colon [b, c] \to Y$ and $k_3 \colon [d, e] \to Y$ be representatives of K_1 and let $k_2 \colon [c, d] \to Y \in K_2$. Since k_1 and k_3 are equivalent we have $\tilde{k}_1 = \tilde{k}_3$, $\dot{\tilde{k}}_1 = \dot{\tilde{k}}_3$ and by Section 10 $k_{1**} = k_{3**}$. Let $i_1' = i_3'$: $(\tilde{k}_1, \dot{\tilde{k}}_1) \to (\widetilde{k_1 \circ k_2}, \dot{\tilde{k}}_1 \cup \dot{\tilde{k}}_2)$, i_2': $(\tilde{k}_2, \dot{\tilde{k}}_2) \to (\widetilde{k_1 \circ k_2}, \dot{\tilde{k}}_1 \cup \dot{\tilde{k}}_2)$, i_0' $(\widetilde{k_1 \circ k_2}, \dot{\widetilde{k_1 \circ k_2}}) \to (\widetilde{k_1 \circ k_2}, \dot{\tilde{k}}_1 \cup \dot{\tilde{k}}_2)$, i_4' $(\widetilde{k_2 \circ k_3}, \dot{\widetilde{k_2 \circ k_3}}) \to (\widetilde{k_1 \circ k_2}, \dot{\tilde{k}}_1 \cup \dot{\tilde{k}}_2)$, j: $(\widetilde{k_1 \circ k_2}, \emptyset) \to (\widetilde{k_1 \circ k_2}, \dot{\tilde{k}}_1 \cup \dot{\tilde{k}}_2)$ be injections. Then evidently

$j = i_0' j_{K_1 K_2} = i_4' j_{K_2 K_1}$ and so according to Section 9

$$j_*((K_1 K_2)_*) = i'_{0*} j_{K_1 K_2 *}((K_1 K_2)_*) = i'_{0*}((K_1 K_2)_{**}) =$$
$$= i'_{0*}((k_1 \circ k_2)_{**}) = i'_{1*}(k_{1**}) + i'_{2*}(k_{2**}) = i'_{2*}(k_{2**}) + i'_{3*}(k_{3**}) = i'_{4*}((K_2 K_1)_{**}) =$$
$$= i'_{4*} j_{K_2 K_1 *}((K_2 K_1)_*) = j_*((K_2 K_1)_*).$$

Hence we have by the lemma in Section 14 $(K_1 K_2)_* = (K_2 K_1)_*$ as required.

16. *Let K_1 and K_2 be continuous closed paths in Y with the same base point. Let $K = K_1 K_2$ and let $i_1'' \colon \tilde{K}_1 \to \tilde{K}$ and $i_2'' \colon \tilde{K}_2 \to \tilde{K}$ be injections. Then*

$$i''_{1*}(K_{1*}) + i''_{2*}(K_{2*}) = K_*.$$

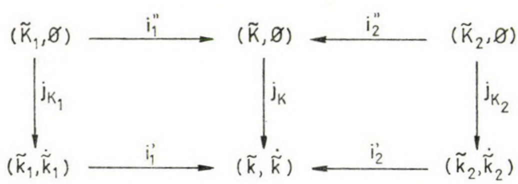
Proof. Suppose that some of the K_i-s $(i=1, 2)$ are degenerated. According to Section 15 there is no loss of generality by assuming that K_1 is a degenerated path. Then $K_2 = K$ and i_2'' is an identical map, consequently i''_{2*} is identical as well. Thus

$$i''_{1*}(K_{1*}) + i''_{2*}(K_{2*}) = i''_{1*}(0) + K_* = K_*$$

is obviously true in this case.

Suppose now that neither K_1 nor K_2 is degenerated. Let $b < c < d$ $(b, c, d \in \mathbf{R})$ and $k_1 \colon [b, c] \to Y \in K_1$, $k_2 \colon [c, d] \to Y \in K_2$. Then $k = k_1 \circ k_2 \in K$. $\dot{\tilde{K}}$ coincides here with $\dot{\tilde{k}}_1$ and $\dot{\tilde{k}}_2$ and thus $\dot{\tilde{k}}_1 \cup \dot{\tilde{k}}_2 = \dot{\tilde{k}}$. Consequently — using the notations of Section 9 — i_0' is an identical map and thus i'_{0*} is an identity as well.

Hence we have by the commutativity of the diagram



$$j_{K*}(i''_{1*}(K_{1*}) + i^*_{2*}(K_{2*})) = i'_{1*} j_{K_1}(K_{1*}) + i'_{2*} j_{K_2 *}(K_{2*}) =$$
$$= i'_{2*}(K_{1**}) + i'_{2*}(K_{2**}) = i'_{1*}(k_{1**}) + i'_{2*}(k_{2**}) = k_{**} = K_{**} = j_{K*}(K_*).$$

However j_{K*} is an isomorphism and so we obtain the desired formula

$$i''_{1*}(K_{1*}) + i''_{2*}(K_{2*}) = K_*.$$

17. Now we can finish our program.

Suppose that Y is a compact Hausdorff space and assign to each closed continuous path K in Y the injection $i_K \colon \tilde{K} \to Y$ and the member $\Phi(K) = i_{K*}(K_*)$ of $H_1(Y)$. Then the map $\Phi \colon \mathcal{K}_Y^{\boldsymbol{\cdot}} \to H_1(Y)$ (see also [2], 2.8) *is a category homomorphism.* (Recall [1], 6.1.)

Proof. We must show that conditions (1), (2), (3) of [1], 6.1 are satisfied. In fact (1) is true by Section 14 and (3) by Section 15.

Let K_1 and K_2 be continuous closed paths in Y with the same base point and let $K=K_1K_2$. Moreover let $i_1''\colon \tilde{K}_1\to\tilde{K}$ and $i_2''\colon \tilde{K}_2\to\tilde{K}$ be injections. Then we have obviously $i_{K_1}=i_K i_1''$ and $i_{K_2}=i_K i_2''$. Applying Section 16 we obtain the required relation

$$\Phi(K_1K_2)=\Phi(K)=i_{K*}(K_*)=i_{K*}\big(i_{1*}''(K_1)+i_{2*}''(K_{2*})\big)=$$
$$=i_{K_1*}(K_{1*})+i_{K_2*}(K_{2*})=\Phi(K_1)+\Phi(K_2).$$

Hence $\Phi\colon \mathscr{K}_Y^{\bullet}\to H_1(Y)$ is a category homomorphism of the i-category of the continuous paths of Y into the 1-dimensional homology group (mod p) of Y.

References

[1] M. Bognár, Über Kategorien mit Involution, *Acta Math. Acad. Sci. Hungar.*, **23** (1972), 147—174.
[2] M. Bognár, Die *i*-Kategorie der stetigen Wege, *Acta Math. Acad. Sci. Hungar.*, **24** (1973), 155—178.
[3] S. Eilenberg and S. Steenrod, *Foundations of algebraic topology* (Princeton, 1952).

(Received September 9, 1977)

DEPARTMENT OF GEOMETRY
EÖTVÖS LORÁND UNIVERSITY
1088 BUDAPEST, MÚZEUM KRT. 6—8.

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 129—139.

MARTINGALE RANDOM CENTRAL LIMIT THEOREMS

By
Z. RYCHLIK (Lublin)

1. Introduction. Let $\{S_n, \mathcal{F}, n \geqq 1\}$ be a martingale on the probability space $(\Omega, \mathcal{A}, P)$, with $S_0=0$, and $X_n=S_n-S_{n-1}$, $n \geqq 1$. $\mathcal{F}_0$ need not be the trivial σ-field $\{\emptyset, \Omega\}$. Let $\varphi_j(t)=E(\exp\{itX_j\}|\mathcal{F}_{j-1})$, and let $\sigma_j^2=E(X_j^2|\mathcal{F}_{j-1})$, $s_n^2=\sum_{j=1}^{n}\sigma_j^2$ for $n=1, 2, \ldots$.

Now let $\{N_n, n \geqq 1\}$ be a sequence of positive integer-valued random variables defined on the same probability space $(\Omega, \mathcal{A}, P)$. Let us denote

$$S_{N_n}=X_1+X_2+\ldots+X_{N_n}, \quad V_n^2=\sum_{j=1}^{N_n}\sigma_j^2, \quad f_n(t)=\prod_{j=1}^{N_n}\varphi_j(t/B_n), \quad b_n=\left(\max_{k \leqq N_n}\sigma_k^2\right)/B_n^2,$$

where $B_n^2=E(X_1^2+X_2^2+\ldots+X_{N_n}^2)$. Throughout the paper B_n is assumed to be finite for all $n \geqq 1$.

In what follows $C=C[0, 1]$ denotes the space of real-valued, continuous functions on [0, 1] and $\mathcal{G}$ denotes a σ-field of Borel sets generated by the open sets of uniform topology. By W we will denote the Wiener measure on $(C, \mathcal{C})$ with the corresponding Wiener process $\{W(t): 0 \leqq t \leqq 1\}$, (cf. [1], Sec. 9).

Let $Y_n(t)$, $0 \leqq t \leqq 1$, be the random function defined as follows:

$$Y_n(t)=S_k/V_n+X_{k+1}(tV_n^2-s_k^2)/V_n\sigma_{k+1}^2 \tag{1}$$

for $0 \leqq t \leqq 1$ and $s_k^2 \leqq tV_n^2 \leqq s_{k+1}^2$, $k=0, 1, 2, \ldots, N_n-1$, where $s_0^2=0$. It is obvious that $Y_n(t)$ is continuous with probability one, being composed of straight line segments joining the points $(s_k^2/V_n^2, S_k/V_n)$, $k=0, 1, 2, \ldots, N_n$. Thus there is a measure P_n in the space $(C, \mathcal{C})$, according to which the stochastic process $\{Y_n(t), 0 \leqq t \leqq 1\}$ is distributed.

In this paper we use an approach developed by BROWN [2], to generate random central limit theorems for martingales. Section 2 defines the random Lindeberg condition for martingales and gives its several equivalent forms. Theorems 1 and 2 generalize Lindeberg—Feller's central limit theorem to random sums. Section 3 contains an invariance principle for a certain class of martingales. From this result we obtain some new limit theorems concerning of sums with random indices. The results obtained are generalizations of that of given in [5—7].

Throughout, we use the notations $X_+=\max(0, X)$ and $X_-=\max(0, -X)$, while Re z is used to denote the real part of z. $\Phi(x)$ denotes the standard normal distribution function, and the various kinds of convergence, in L_p norm, in probability, and weak (in distribution) are denoted by $\xrightarrow{L_p}$, $\xrightarrow{P}$ and $\xrightarrow{\mathcal{D}}$, respectively.

2. Random central limit theorems. Throughout the paper we say that the random Lindeberg condition is satisfied if, for all $\varepsilon>0$,

$$L(n,\varepsilon)=B_n^{-2}E\left\{\sum_{i=1}^{N_n}X_i^2I(|X_i|\geqq\varepsilon B_n)\right\}\to 0 \quad \text{as} \quad n\to\infty, \tag{2}$$

where $I(A)$ denotes the indicator function of the set A. Furthermore, we consider martingales for which

$$V_n^2/B_n^2\xrightarrow{P}1 \quad \text{as} \quad n\to\infty. \tag{3}$$

First we shall prove the following

LEMMA 1. *If N_n is stopping time with respect to $\{\mathscr{F}_i,\, i\geqq 0\}$, then* (3) *is equivalent to*

$$V_n^2/B_n^2\xrightarrow{L_1}1 \quad as \quad n\to\infty. \tag{4}$$

PROOF. It is obvious that (4) implies (3). On the other hand, for every $\varepsilon>0$,

$$E(V_n^2/B_n^2-1)_-\leqq\varepsilon+P(|V_n^2B_n^{-2}-1|\geqq\varepsilon).$$

Thus, by (3), $\lim_{n\to\infty}E(V_n^2/B_n^2-1)_-=0$. Furthermore,

$$EV_n^2=\sum_{k=1}^{\infty}\int_{[N_n=k]}\sum_{j=1}^{k}\sigma_j^2\,dP=\sum_{k=1}^{\infty}\int_{[N_n\geqq k]}E(X_k^2|\mathscr{F}_{k-1})\,dP=$$

$$=\sum_{k=1}^{\infty}\int_{[N_n\geqq k]}X_k^2\,dP=\sum_{k=1}^{\infty}\int_{[N_n=k]}\sum_{i=1}^{k}X_i^2\,dP=E(X_1^2+\ldots+X_{N_n}^2)=B_n^2.$$

Thus $E(V_n^2/B_n^2-1)=0$, which implies (4) since $E(V_n^2/B_n^2-1)_+=E(V_n^2/B_n^2-1)\to 0$ as $n\to\infty$.

Let us put

$$g(n,\varepsilon)=V_n^{-2}\sum_{j=1}^{N_n}E\big(X_j^2I(|X_j|\geqq\varepsilon B_n)|\mathscr{F}_{j-1}\big),\quad G(n,\varepsilon)=V_n^2B_n^{-2}g(n,\varepsilon),$$

$$h(n,\varepsilon)=V_n^{-2}\sum_{j=1}^{N_n}E\{X_j^2U(|X_j|B_n^{-1}\varepsilon^{-1})|\mathscr{F}_{j-1}\},\quad H(n,\varepsilon)=V_n^2B_n^{-2}h(n,\varepsilon),$$

where $U(x)$ is any continuous nonnegative function of bounded variation on $[0,\infty)$ for which $U(0)=0$ and $U(x)\to\text{const.}\,(>0)$ as $x\to\infty$.

LEMMA 2. *Assume that for every n, N_n is a stopping time with respect to $\{\mathscr{F}_i,\, i\geqq 0\}$. Then, under the condition* (3), *or alternatively* (4), *the random Lindeberg condition is equivalent to the convergence to zero as $n\to\infty$ of $g(n,\varepsilon)$, $G(n,\varepsilon)$, $h(n,\varepsilon)$ or $H(n,\varepsilon)$, for all $\varepsilon>0$, either in probability or in L_1.*

PROOF. It suffices to show the mutual equivalence of convergences in probability of g, G, h and H, then to show that each such convergence in probability implies a corresponding convergence in L_1 norm since the convergence in L_1 is stronger than convergence in probability.

Let us observe that

$$EG(n,\varepsilon) = B_n^{-2}\sum_{k=1}^{\infty}\int_{[N_n=k]}\sum_{i=1}^{k} E\{X_i^2 I(|X_i|\geqq \varepsilon B_n)|\mathscr{F}_{i-1}\} =$$

$$= B_n^{-2}\sum_{k=1}^{\infty}\int_{[N_n\geqq k]} E\{X_k^2 I(|X_k|\geqq \varepsilon B_n)|\mathscr{F}_{k-1}\}\, dP=$$

$$= B_n^{-2}\sum_{k=1}^{\infty}\int_{[N_n\geqq k]} X_k^2 I(|X_k|\geqq \varepsilon B_n)\, dP = L(n,\varepsilon).$$

Thus $L(n,\varepsilon)\to 0$ as $n\to\infty$ if, and only if, $EG(n,\varepsilon)\to 0$ as $n\to\infty$. Moreover, taking into account (3), we see that $g(n,\varepsilon)$ converges in probability if and only if $G(n,\varepsilon)$ does. Also, by (3), $h(n,\varepsilon)\xrightarrow{P}0$ if and only if $H(n,\varepsilon)\xrightarrow{P}0$. Furthermore, noting that there exist constants $a, b>0$ for which

$$bI(x\geqq a)\leqq U(x) \quad \text{for all} \quad x\geqq 0 \tag{5}$$

we see that $h(n,\varepsilon)\xrightarrow{P}0$ implies $g(n,\varepsilon)\xrightarrow{P}0$. On the other hand, if K is the total variation of U, then

$$h(n,\varepsilon) = \int_0^{\infty} g(n,\varepsilon y)\, dU(y) \quad \text{pointwise a.e.,}$$

$$\leqq \int_0^{\delta} dU(y) + Kg(n,\varepsilon\delta) \quad \text{a.e.}$$

since $g(n,\varepsilon)\leqq 1$ a.e. and $g(n,y)$ is a nonincreasing function with respect to y, for each n. Thus, by choosing δ small, it is obvious that $h(n,\varepsilon)\xrightarrow{P}0$ as $n\to\infty$ for all $\varepsilon>0$, if $g(n,\varepsilon)\xrightarrow{P}0$ as $n\to\infty$ for all $\varepsilon>0$.

Now we are going to show that convergences in probability of g, G, h and H imply corresponding convergence in mean. In the case of g and h this is trivial since the random variables g and h are uniformly bounded. Thus we can assume that

$$\lim_{n\to\infty} Eh(n,\varepsilon) = 0 \quad \text{for all} \quad \varepsilon>0. \tag{6}$$

It will suffice now to show that (6) implies

$$\lim_{n\to\infty} EH(n,\varepsilon) = 0 \quad \text{for all} \quad \varepsilon>0, \tag{7}$$

since $H(n,\varepsilon)\xrightarrow{P}0$ implies (6). Moreover, by (5), $\lim_{n\to\infty} EG(n,\varepsilon)=0$ provided that (7) holds. Thus the proof will be completed if we show that (6) implies (7).

By (4) we have $V_n^2/B_n^2=1+d_n$, where $E|d_n|\to 0$ as $n\to\infty$. Therefore

$$EH(n,\varepsilon) = Eh(n,\varepsilon)+Ed_n h(n,\varepsilon)\to 0 \quad \text{as} \quad n\to\infty$$

since $E|d_n h(n,\varepsilon)|\leqq KE|d_n|$ and (6) holds. K stands, as above, for the total variation of U. The proof is complete.

THEOREM 1. *Let N_n, for all n, be a stopping time with respect to* $\{\mathscr{F}_i,\ i\geqq 0\}$ *such that* (3) *holds. Then*

$$f_n(t)\xrightarrow{P}\exp(-t^2/2)\quad\text{as}\quad n\to\infty,\quad\text{for all } t, \tag{8}$$

and

$$b_n\xrightarrow{P}0\quad\text{as}\quad n\to\infty \tag{9}$$

if and only if the random Lindeberg condition (2) *holds.*

PROOF. Let $\varepsilon>0$ be given. Then

$$B_n^{-2}E\{\max_{k\leqq N_n}E(X_k^2|\mathscr{F}_{k-1})\}\leqq\varepsilon^2+B_n^{-2}V_n^2g(n,\varepsilon)=\varepsilon^2+L(n,\varepsilon),$$

and as $\varepsilon>0$ can be given arbitrarily small, the random Lindeberg condition implies that $b_n\xrightarrow{L_1}0$ and hence

$$b_n\xrightarrow{P}0\quad\text{as}\quad n\to\infty. \tag{10}$$

Now the basic inequality (4.1) in [3], page 485, shows, for every $k\leqq N_n$,

$$|\varphi_k(tB_n^{-1})-1|\leqq t^2\sigma_k^2/2B_n^2\leqq t^2b_n/2\xrightarrow{L_1}0\quad\text{as}\quad n\to\infty. \tag{11}$$

From the Taylor series it is seen that

$$\sum_{k=1}^{N_n}|\log(1+z_k)-z_k|\leqq\sum_{k=1}^{N_n}|z_k|\leqq\max_{k\leqq N_n}|z_k|\sum_{k=1}^{N_n}|z_k|.$$

With $z_k=\varphi_k(t/B_n)-1$, by (11) this yields for fixed t

$$\sum_{k=1}^{N_n}|\log\varphi_k(t/B_n)+1-\varphi_k(t/B_n)|\leqq t^4b_n(V_n^2/B_n^2)/4.$$

Thus, by (3) and (10),

$$\log f_n(t)=\sum_{k=1}^{N_n}\log\varphi_k(t/B_n)=-\sum_{k=1}^{N_n}\big(1-\varphi_k(t/B_n)\big)+A_n(t),$$

where for each fixed t, $A_n(t)\xrightarrow{P}0$ if $b_n\xrightarrow{P}0$ as $n\to\infty$.

Define, as in [2], functions Q and M by writing

$$\exp(ix)=1+ix+x^2Q(x)/2-x^2/2,$$

and $M(x)=\min(|x|/3,\,2)$. Then

$$|1-Q(x)|\leqq 1 \tag{12}$$

and

$$|Q(x)|\leqq M(x)\quad\text{for all } x. \tag{13}$$

Thus, by the relations given above, we obtain

$$\text{(14)} \qquad \log f_n(t) = -t^2 V_n^2/2B_n^2 + A_n(t) + t^2 B_n^{-2} \sum_{j=1}^{N_n} E\{X_j^2 Q(tX_j B_n^{-1})|\mathscr{F}_{j-1}\}/2.$$

Theorem 1 will then follow if the convergence in probability to zero as $n\to\infty$ of the third term of (14) is equivalent to the random Lindeberg condition. But the function M has the properties required for the function U in Lemma 2. So, by (13) and Lemma 2, the third term of (14) converges in probability to zero if the random Lindeberg condition will be satisfied. On the other hand, the convergence in probability to zero of the third term of (14) implies the corresponding convergence of its real part. But the function Re Q has the properties required for U in Lemma 2. Thus, by Lemma 2, (2) holds. Q. e. d.

The results of this section remain valid also when N_n is, for each n, independent of $X_1, X_2, \ldots$. From Theorem 1 we obtain the following generalization of Lindeberg—Feller's central limit theorem [4], page 280.

THEOREM 2. *Suppose* $\{X_n,\ n\geqq 1\}$ *is a sequence of independent random variables such that* $EX_n=0$, $EX_n^2=\sigma_n^2<\infty$. *Let* $\{N_n,\ n\geqq 1\}$ *be a sequence of positive integer-valued random variables independent of* $X_1, X_2, \ldots,$ *and let* $M_n^2=\sum_{k=1}^{N_n}\sigma_k^2$. *If* $M_n^2/B_n^2 \xrightarrow{P} 1$, *then*

$$\lim_{n\to\infty} P(S_{N_n} < xB_n) = \Phi(x) \quad \textit{and} \quad \lim_{n\to\infty} E\big(\max_{j\leqq N_n} \sigma_j^2/B_n^2\big) = 0$$

if and only if for every $\varepsilon>0$

$$\lim_{n\to\infty} B_n^{-2} E\left\{\sum_{k=1}^{N_n} EX_k^2 I(|X_k| \geqq \varepsilon B_n)\right\} = 0.$$

PROOF. Let $\mathscr{F}_j=\sigma(X_1, X_2, \ldots, X_j)$. Then

$$\varphi_j(t) = E\{\exp(itX_j)|\mathscr{F}_{j-1}\} = E\exp(itX_j),$$

$$f_n(t) = \prod_{j=1}^{N_n} E\exp(itX_j), \quad E(X_j^2|\mathscr{F}_{j-1}) = EX_j^2$$

and $V_n^2=M_n^2$. Moreover,

$$B_n^2 = E(X_1^2+X_2^2+\ldots+X_{N_n}^2) = E\left(\sum_{k=1}^{N_n}\sigma_k^2\right) = EM_n^2,$$

since N_n is independent of $X_1, X_2, \ldots$.

Now let us observe that $E\exp(itS_{N_n})=Ef_n(t)$, and $|f_n(t)|\leqq 1$. Thus, (8) holds if and only if $E\exp(itS_{N_n})\to\exp(-t^2/2)$. On the other hand, for every $\varepsilon>0$,

$$B_n^{-2}\big(\max_{k\leqq N_n} EX_k^2\big) \leqq \varepsilon + B_n^{-2} \sum_{k=1}^{N_n} EX_k^2 I(|X_k| \geqq \varepsilon B_n) \xrightarrow{L_1} \varepsilon.$$

Thus Theorem 2 follows from Theorem 1 since, by the assumptions,

$$L(n, \varepsilon) = B_n^{-2} E\left\{\sum_{k=1}^{N_n} EX_k^2 I(|X_k| \geqq \varepsilon B_n)\right\}.$$

Theorem 2 also generalizes the main result of H. Robbins [5].

Remark. Recently several papers have appeared which were devoted to the study of the limit distribution of S_{N_n} as $n \to \infty$. The results obtained have the following form. Suppose the partial sums S_n obey the central limit theorem. If $N_n/k_n \xrightarrow{P} \lambda$, where λ is a positive random variable and the k_n are constants tending to infinity then, under some additional assumptions, the sequence $\{S_{N_n}, n \geqq 1\}$ is asymptotically normal (cf. Billingsley [1]). Thus in the proof it is an important fact that the partial sums S_n for a non-random index n obey the central limit theorem. Our approach is different. We simply ask when the sequence $\{S_{N_n}, n \geqq 1\}$ is asymptotically normal. From the assumptions of Theorem 2 it does not follow that $\{S_n, n \geqq 1\}$ is asymptotically normal. To see this let us consider the following

Example. Let $\{X_k, k \geqq 1\}$ be a sequence of independent random variables defined as follows:

$$P[S_{2^{2^n}} = 2^{2^{n-1}}] = P[X_{2^{2^n}} = -2^{2^{n-1}}] = 1/2$$

for $n = 1, 2, \ldots$, and assume that X_k, for $k \neq 2^{2^n}$ $(n = 1, 2, \ldots)$ $k \geqq 1$, has the normal distribution function with mean zero and variance one. Let us put $S_n = X_1 + X_2 + \ldots + X_n$. Then

$$\sigma^2 S_{2^{2^n}} \approx 2 \cdot 2^{2^n} \quad \text{and} \quad \sigma^2 S_{2^{2^n}-1} \approx 2^{2^n} \quad \text{as} \quad n \to \infty.$$

Now let us put, for example,

$$P[N_n = 2^{2^n} - 1] = 1 - 1/n, \quad P[N_n = 2^{2^n}] = 1/n.$$

Then

$$B_n^2 = EM_n^2 \approx 2^{2^n} \quad \text{as} \quad n \to \infty.$$

Moreover, it is easy to see that $M_n^2/B_n^2 \xrightarrow{P} 1$ and, for every $\varepsilon > 0$,

$$B_n^{-2} E\left\{\sum_{k=1}^{N_n} EX_k^2 I(|X_k| \geqq \varepsilon B_n)\right\} \to 0 \quad \text{as} \quad n \to \infty.$$

Thus, by Theorem 2,

$$\lim_{n \to \infty} P[S_{N_n} < xB_n] = \Phi(x)$$

and

$$E\left(\max_{1 \leqq k \leqq N_n} \sigma^2 X_k / B_n^2\right) \to 0 \quad \text{as} \quad n \to \infty.$$

On the other hand, taking into account that

$$\prod_{k=1}^{n-1} \cos\left(t 2^{2^{k-1}} / \sigma(S_{2^{2^n}})\right) \to 1 \quad \text{as} \quad n \to \infty,$$

we get

$$E\exp\left(itS_{2^{2^n}}/\sigma(S_{2^{2^n}})\right) = \cos\left(t2^{2^{n-1}}/\sigma(S_{2^{2^n}})\right)\prod_{k=1}^{n-1}\cos\left(t2^{2^{k-1}}/\sigma(S_{2^{2^n}})\right)\cdot$$

$$\cdot\exp\{-t^2(2^{2^n}-n)/2\sigma(S_{2^{2^n}})\}\to\cos(t/\sqrt{2})\exp(-t^2/4)\neq\exp(-t^2/2).$$

Thus $P[S_n<x\sigma S_n]$ does not converge to the standard normal distribution function.

3. Random functional central limit theorems for martingales. Throughout this section we assume that N_n is, for each n, independent of $X_1, X_2, \ldots$.

THEOREM 3. *Let $\{S_n, \mathscr{F}_n, n\geqq 1\}$ be a martingale, and let $\{N_n, n\geqq 1\}$ be a sequence of positive integer-valued random variables such that N_n is, for each n, independent of $X_n=S_n-S_{n-1}$, $n=1,2,\ldots$. Let $\{P_n, n\geqq 1\}$ be the sequence of probability measures on $(C,\mathscr{C})$ determined by the distribution $\{Y_n(t), 0\leqq t\leqq 1\}$. If* (3) *and the random Lindeberg condition hold, then $P_n\xrightarrow{\mathscr{D}}W$.*

PROOF. We may assume that the sequence $\{X_n, n\geqq 1\}$ satisfies the following relation:

$$P(V_n^2/B_n^2\leqq C)=1\quad\text{for all } n, \tag{15}$$

where C is a constant greater than 1. Of course, there is no loss of generality, because if $\{X_n, n\geqq 1\}$ does not satisfy (15), then we can choose any constant $C>1$ and for each n set

$$X_{nj}=X_jI(s_j^2\leqq CB_n^2),\quad j=1,2,\ldots.$$

Then $\{X_{ni}, \mathscr{F}_i, i\geqq 1\}$ will form a martingale difference sequence and

$$\lim_{n\to\infty}P\left(\bigcup_{j=1}^{N_n}[X_j\neq X_{nj}]\right)\leqq\lim_{n\to\infty}P(V_n^2/B_n^2\leqq C)=0$$

because (3) holds. Furthermore, the random Lindeberg condition holds with X_j replaced by X_{nj}. Thus, in oder to prove convergence in distribution properties for $\sum_j X_j/B_n$ (where the summation can be taken over a subset of the integers from 1 to N_n), it suffices to prove them for $\sum_j X_{nj}/B_n$ as $n\to\infty$. So, taking into account the considerations given above, we assume throughout that the random Lindeberg condition, (3) and (15) hold.

We shall first establish that the finite dimensional distributions of P_n converge weakly to the corresponding finite dimensional distributions of W. Let k be a fixed positive integer, and let $0=t_0<t_1<\ldots<t_{k-1}<t_k=1$. We wish to show that

$$\left(Y_n(t_1), Y_n(t_2), \ldots, Y_n(t_k)\right)\xrightarrow{\mathscr{D}}\left(W(t_1), W(t_2), \ldots, W(t_k)\right).$$

But, by Corollary 1 to Theorem 5.1 in [1] it is enough to prove that

$$\left(Y_n(t_1), Y_n(t_2)-Y_n(t_1), \ldots, Y_n(t_k)-Y_n(t_{k-1})\right)\xrightarrow{\mathscr{D}} \tag{16}$$

$$\xrightarrow{\mathscr{D}}\left(W(t_1), W(t_2)-W(t_1), \ldots, W(t_k)-W(t_{k-1})\right).$$

On the other hand, by Theorem 7.7 in [1], (16) holds if and only if

$$(17)\qquad c_1 Y_n(t_1)+c_2\big(Y_n(t_2)-Y_n(t_1)\big)+\ldots+c_k\big(Y_n(t_k)-Y_n(t_{k-1})\big)\xrightarrow{\mathcal{D}}$$
$$\xrightarrow{\mathcal{D}} c_1 W(t_1)+c_2\big(W_n(t_2)-W_n(t_1)\big)+\ldots+c_k\big(W(t_k)-W(t_{k-1})\big),$$

for real numbers $c_1, c_2, \ldots, c_k$ not all zero.
Let us put

$$m_j=\max\{m\geqq 0:\ s_m^2\leqq t_jV_n^2\},\quad s_r=c_j\quad \text{for}\quad m_{j-1}<r\leqq m_j,$$
$$s=\max_{r\leqq N_n}|s_r|=\max_j|c_j|,$$

and observe that

$$|Y_n(t_j)-Y_n(s_{m_j}^2/V_n^2)|\leqq|Y_n(s_{m_j+1}^2/V_n^2)-Y_n(s_{m_j}^2/V_n^2)|=|X_{m_j+1}|/V_n\leqq\big(\max_{k\leqq N_n}|X_k|\big)/V_n.$$

On the other hand, for any given $\varepsilon>0$,

$$\big(\max_{k\leqq N_n}|X_k|\big)/B_n\leqq\varepsilon+\varepsilon^{-1}B_n^{-2}\sum_{k=1}^{N_n}X_k^2 I(|X_k|\geqq\varepsilon B_n)\xrightarrow{L_1}\varepsilon,$$

since (2) holds. Therefore, by (3), $\max_{k\leqq N_n}|X_k|)/V_n\xrightarrow{P}0$. Thus (17) holds if and only if

$$(18)\qquad \lim_{n\to\infty}E\exp\left(i\sum_{k=1}^{N_n}s_kX_kV_n^{-1}\right)=\exp(-\sigma^2/2),$$

where $\sigma^2=\sum_{j=1}^{k}c_j^2(t_j-t_{j-1})$. Again by (3), (18) is equivalent to

$$(19)\qquad \lim_{n\to\infty}E\exp\left(i\sum_{k=1}^{N_n}s_kX_kB_n^{-1}\right)=\exp(-\sigma^2/2).$$

Put

$$T_r=B_n^{-1}\sum_{j=1}^{r}s_jX_j,\quad U_r^2=B_n^{-2}\sum_{j=1}^{r}s_j^2\sigma_j^2.$$

We wish to show that

$$(20)\qquad \lim_{n\to\infty}E\big(\exp(iT_{N_n}+U_{N_n}^2/2)-1\big)=0$$

and

$$(21)\qquad \lim_{n\to\infty}E\,|\exp(U_{N_n}^2/2)-\exp(\sigma^2/2)|=0,$$

since then by simple computations from (20) and (21) (19) easily established.
Define

$$(22)\qquad Z_j=\exp(iT_{j-1}+U_j^2/2)\{\exp(is_jX_j/B_n)-\exp(-s_j^2\sigma_j^2/2B_n^2)\}=$$
$$=\exp(iT_{j-1}+U_{N_n}^2/2)\{is_jX_jB_n^{-1}-s_j^2X_j^2(1-Q(s_jX_jB_n^{-1}))/2B_n^2+$$
$$+s_j^2\sigma_j^2/2B_n^2-Z(s_j^2\sigma_j^2/2B_n^2)\},$$

where $Z(x)=\exp(-x)-1+x$ for $x\geqq 0$. Then

$$|E\exp(iT_{N_n}+U_{N_n}^2/2)-1|=\left|E\sum_{j=1}^{N_n}Z_j\right|.$$

Hence, using (13), (15), $\max|s_j|\leqq s$, the inequality $Z(x)\leqq x^2/2$ and the independence N_n of $X_1, X_2, \ldots$, we get

$$\left|E\sum_{j=1}^{N_n} Z_j\right| \leqq E\sum_{j=1}^{N_n}\left|E(Z_j|\mathscr{F}_{j-1})\right| \leqq$$

$$\leqq \exp(s^2C/2)E\left\{\sum_{j=1}^{N_n}\left|E(s_j^2X_j^2Q(s_jX_jB_n^{-1})|\mathscr{F}_{j-1})/2B_n^2-Z(s_j^2\sigma_j^2/2B_n^2)\right|\right\}\leqq$$

$$\leqq s^2\exp(s^2C/2)E\left\{\sum_{j=1}^{N_n}E\big(X_j^2M(|sX_j|B_n^{-1})|\mathscr{F}_{j-1}\big)\right\}\Big/2B_n^2+$$

$$+s^4\exp(s^2C/2)\,E\left(\sum_{j=1}^{N_n}\sigma_j^4\right)\Big/8B_n^4\leqq s^2\exp(s^2C/2)EH(n, s^{-1})/2+$$

$$+s^4\exp(s^2C/2)E(b_nV_n^2/B_n^2)/8,$$

and, by Lemma 2, $EH(n, s^{-1})\to 0$ as $n\to\infty$. Furthermore, for every $\varepsilon>0$, $Eb_n\leqq \leqq\varepsilon^2+EG(n, \varepsilon)$. Thus, by Lemma 2 and (15), $E(b_nV_n^2B_n^{-2})\to 0$ as $n\to\infty$, and this proves (20).

For any $1\leqq j\leqq k$,

$$|(s_{m_j}^2-s_{m_{j-1}}^2)/B_n^2-(t_j-t_{j-1})|\leqq b_n.$$

Thus $(s_{m_j}^2-s_{m_{j-1}}^2)/B_n^2\xrightarrow{P}(t_j-t_{j-1})$. It follows that $U_{N_n}^2\xrightarrow{P}\sigma^2$ as $n\to\infty$, and also, since $U_{N_n}^2\leqq s^2C$, that (21) holds.

We now show that the sequence of measures $\{P_n,\ n\geqq 1\}$, in $(C, \mathscr{C})$, is tight. As $P\big(Y_n(0)=0\big)=1$, for the tightness of this sequence it is enough to prove (cf. Theorem 8.2, [1]) that, for every $\varepsilon>0$,

(23) $$\lim_{h\to 0}\limsup_{n\to 0} P\big(\sup_{|t-s|\leqq h}|Y_n(s)-Y_n(t)|\geqq\varepsilon\big)=0.$$

But, by (3), this is equivalent to

(24) $$\lim_{h\to 0}\limsup_{n\to\infty} P\big(\sup_{|t-s|\leqq h}V_n|Y_n(s)-Y_n(t)|\geqq\varepsilon B_n\big)=0.$$

To verify (24) the proof proceeds as in [6] or [2]. So, for any given $\varepsilon>0$, we have

(25)

$$P\big(\sup_{|t-s|\leqq h}V_n|Y_n(s)-Y_n(t)|\geqq\varepsilon B_n\big)\leqq\sum_{k:\,kh<1}P\big(\sup_{kh<t\leqq(k+1)h}V_n|Y_n(t)-Y_n(kh)|\geqq\varepsilon B_n/4\big).$$

On the other hand, by (1),

$$\sup_{kh<t\leqq(k+1)h}V_n|Y_n(t)-Y_n(kh)|\leqq 2\max_{q_k<r\leqq q_{k+1}}\left|\sum_{j=q_k}^{r}X_j\right|,$$

where $q_k=\max\{i:\ s_i^2\leqq khV_n^2\}$. Hence

$$P\big(\sup_{kh<t\leqq(k+1)h}V_n|Y_n(t)-Y_n(kh)|\geqq\varepsilon B_n/4\big)\leqq P\left(\max_{q_k<r\leqq q_{k+1}}\left|\sum_{j=q_k}^{r}X_j\right|\geqq\varepsilon B_n/4\right)=P_k.$$

Let $q_{1k}=\max\{i\colon s_i^2\leqq khs_l^2\}$. Then, taking into account that N_n is independent of X_k, $k\geqq 1$, and then applying Lemma 4 [2], we obtain

$$P_k \leqq 8\sum_{i=1}^{\infty} P(N_n=i) \int\limits_{\left[\left|\sum_{j=q_{ik}}^{q_{ik+1}} X_j\right| \geqq \varepsilon B_n/8\right]} \left|\sum_{j=q_{ik}}^{q_{ik+1}} X_j\right| dP/\varepsilon B_n \leqq$$

$$\leqq (8/\varepsilon)\left\{\sum_{i=1}^{\infty} P(N_n=i) P\left(\left|\sum_{j=q_{ik}}^{q_{ik+1}} X_j\right| \geqq \varepsilon B_n/8\right\}^{1/2} = (8/\varepsilon)\left\{P\left(\left|\sum_{j=q_k}^{q_{k+1}} X_j\right| \geqq \varepsilon B_n/8\right)\right\}^{1/2}.$$

But

$$P\left(\left|\sum_{j=q_k}^{q_{k+1}} X_j\right| \geqq \varepsilon B_n/8\right) = P\left(V_n|Y_n(s^2_{q_{k+1}}/V_n^2)-Y_n(s^2_{q_k}/V_n^2)| \geqq \varepsilon B_n/8\right).$$

Thus from the convergence of the finite dimensional distributions and by (3) we have

$$\limsup_{n\to\infty} P\left(\left|\sum_{j=q_k}^{q_{k+1}} X_j\right| \geqq \varepsilon B_n/8\right) = \frac{1}{\sqrt{2\pi h}} \int\limits_{|u|\geqq \varepsilon/8} \exp(-u^2/2h)\,du.$$

In view of the relations given above we now obtain

$$\limsup_{n\to\infty} P\big(\sup_{|t-s|\leqq h} V_n|Y_n(s)-Y_n(t) \geqq \varepsilon B_n\big) \leqq (8/\varepsilon)\frac{1}{h}\left(\frac{1}{\sqrt{2\pi}} \int\limits_{|u|>\varepsilon/8\sqrt{h}} \exp(-u^2/2)\,du\right)^{1/2},$$

and here the right-hand side tends to zero as $h\to 0$. Thus we have proved the tightness of the sequence $\{P_n,\ n\geqq 1\}$. Hence, by Theorem 8.1 in [1], $P_n \xrightarrow{\mathscr{D}} W$.

Theorem 3 enables us to give the following generalizations of the results given in [5], [6] and [7].

THEOREM 4. *Let $\{S_n, \mathscr{F}_n,\ n\geqq 1\}$ be a martingale, and let $\{N_n,\ n\geqq 1\}$ be a sequence of positive integer-valued random variables such that N_n is, for each n, independent of $X_n=S_n-S_{n-1}$, $n=1, 2, \ldots$. If* (2) *and* (3) *hold, then*

$$\lim_{n\to\infty} P(S_{N_n}<xV_n)=\Phi(x), \quad \textit{and} \quad \lim_{n\to\infty} P(S_{N_n}<xB_n)=\Phi(x)$$

for all x.

THEOREM 5. *Under the assumptions of Theorem 4, for each $x>0$,*

$$\lim_{n\to\infty} P\big(\max_{0\leqq k\leqq N_n} S_k < xB_n\big) = \frac{2}{\sqrt{2\pi}} \int_0^x \exp(-u^2/2)\,du,$$

and

$$\lim_{n\to\infty} P\big(\max_{k\leqq N_n} |S_k| < xB_n\big) = \frac{1}{\sqrt{2\pi}} \int_{-x}^{x} \sum_{k=-\infty}^{\infty} (-1)^k \exp\left(-\frac{(u-2kx)^2}{2}\right) du.$$

The proof of this theorem is a consequence of Theorem 3 and results given in [1].

THEOREM 6. *Let $\{X_n,\ n\geqq 1\}$ be a sequence of independent random variables with $EX_k=0$ and $EX_k^2=\sigma_k^2<\infty$, $k=1, 2, \ldots$, and let $\{N_n,\ n\geqq 1\}$ be a sequence* *of positive integer-valued random variables such that* N_n *is, for each* $n \geqq 1$, *independent of* $X_1, X_2, \ldots$. *Further on, let* Y_n' *be a random function defined as*

$$Y_n'(t) = S_k/M_n + (M_n^2 t - s_k^2) X_{k+1}/M_n \sigma_{k+1}^2,$$

for $0 \leqq t \leqq 1$, *and* $s_k^2 \leqq M_n^2 t \leqq s_{k+1}^2$, $k = 0, 1, \ldots N_n - 1$, *where* $S_0 = 0$, $s_0^2 = 0$, $s_k^2 = \sum_{i=1}^{k} \sigma_i^2$ *and* $M_n^2 = \sum_{k=1}^{N_n} \sigma_k^2$. *Let* P_n' *denote the distribution of* Y_n' *in* $(C, \mathscr{C})$. *If* $M_n^2/B_n^2 \xrightarrow{P} 1$ *and for every* $\varepsilon > 0$

$$\lim_{n \to \infty} B_n^{-2} E\left\{\sum_{k=1}^{N_n} EX_k^2 I(|X_k| \geqq \varepsilon B_n)\right\} = 0, \quad \textit{then} \quad P_n' \xrightarrow{\mathscr{D}} W.$$

ACKNOWLEDGEMENT. I am grateful to the referee for useful comments and to August Lapola for helpful discussion.

References

[1] P. BILLINGSLEY, *Convergence of Probability Measures*. Wiley (New York, 1968).
[2] B. M. BROWN, Martingale central limit theorems, *Ann. Math. Statist.*, **42** (1971), 59—66.
[3] W. FELLER, *An Introduction to Probability Theory and Its Applications* 2. Wiley (New York, 1966).
[4] M. LOÉVE, *Probability Theory,* 3rd ed., Van Nostrand (New York, 1963).
[5] H. ROBBINS, The asymptotic distribution of the sums of a random number of random variables, *Bull. Amer. Math. Soc.*, **54** (1948), 1151—1161.
[6] Z. RYCHLIK, D. SZYNAL, A functional random-sum central limit theorem, *Bull. Acad. Polon. Sci., Sér. Sci. Math. Astronom. Phys.*, **23** (1975), 1013—1018.
[7] Z. RYCHLIK, A central limit theorem for the sums of a random number of independent random variables, *Coll. Math.*, **35** (1976), 147—158.

(Received September 16, 1977; revised March 20, 1978)

INSTITUTE OF MATHEMATICS
MARIA SKŁODOWSKA UNIVERSITY
20—031 LUBLIN, UL. NOWOTKI 10
POLAND

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 141—155.

AN A.S. INVARIANCE PRINCIPLE FOR LACUNARY SERIES $f(n_k x)$

By

I. BERKES (Budapest) and W. PHILIPP (Urbana)

1. Introduction

Let $f(x)$ $(-\infty < x < +\infty)$ be a measurable function such that

$$f(x+1) = f(x), \quad \int_0^1 f(x)\, dx = 0. \tag{1.1}$$

It is well known that if $f(x)$ is smooth enough and the sequence $\{n_k\}$ of integers grows rapidly then the sequence $f(n_k x)$ of functions $(0 \leqq x \leqq 1)$ behaves like a sequence of independent random variables. A typical result in this direction (see [5], [8]) is that if f satisfies (1.1) and the Lipschitz condition and

$$n_{k+1}/n_k \to \infty \tag{1.2}$$

then $f(n_k x)$ obeys the central limit theorem and the law of the iterated logarithm (CLT and LIL in the sequel). Here (1.2) is best possible in the sense that it cannot be weakened to

$$n_{k+1}/n_k \geqq q > 1 \tag{1.3}$$

even with a large q. This is shown by the example of ERDŐS and FORTET (see [5])

$$f(x) = \cos 2\pi x + \cos 2\pi m x, \quad n_k = m^k - 1$$

for which both the central limit theorem and the law of iterated logarithm fail to hold. On the other hand, there exist many sequences $\{n_k\}$ satisfying (1.3) but not (1.2) such that $f(n_k x)$ satisfies the CLT and LIL. E.g. $n_k = 2^k$ is such a sequence (see [4]). It was Gaposhkin who characterized all the sequences $\{n_k\}$ (among the sequences obeying (1.3)) such that $f(n_k x)$ obeys the CLT. Let us say that a sequence $\{n_k\}$ of integers belongs to class

B_2 if there is a constant C such that the number of solutions of the equation $n_k \pm n_l = \nu$ $(k > l)$ is at most C for any integer $\nu > 0$;

D_m if the (set-theoretic) union of the sequences $\{n_k\}$, $\{2n_k\}$, ..., $\{mn_k\}$, considered as a single sequence, belongs to class B_2;

D_∞ if $\{n_k\}$ belongs to class D_m for all integers $m = 1, 2, \ldots$.

GAPOSHKIN showed (see [3]) that if $\{n_k\}$ belongs to D_∞ (and satisfies (1.3)) then $f(n_k x)$ obeys the CLT for all sufficiently smooth f; on the other hand, if $\{n_k\}$ does not belong to D_∞ (but satisfies (1.3)) then there exists a trigonometric polynomial f such that $f(n_k x)$ fails to obey the CLT.

The purpose of the present paper is to extend (the positive half of) Gaposhkin's theorem and to prove an a.s. invariance principle for the sequence $f(n_k x)$ under the assumption that $\{n_k\}$ satisfies (1.3) and belongs to D_∞. Our method (which differs from that of Gaposhkin) makes use of martingale tools; in fact, it is a combination of the methods of [1—2], [6]. In [2], § 3 an a.s. invariance principle was proved for $f(n_k x)$ assuming a condition for $\{n_k\}$ (the so called Λ^* condition) which is slightly more stringent than D_∞. The present improvement (which is now best possible) is obtained by utilizing ideas from [6].

Our main result is the following:

THEOREM 1. *Let $f(x)$ $(-\infty<x<+\infty)$ satisfy* (1.1) *and the Lipschitz condition. Assume that $\{n_k\}$ satisfies* (1.3) *and belongs to class D_∞. Assume finally that there exists a constant $C_1>0$ such that for any $M\geqq 0$, $N\geqq N_0$ we have*

$$\int_0^1 \left(\sum_{j=M+1}^{M+N} f(n_j x)\right)^2 dx \geqq C_1 N. \tag{1.4}$$

Let $S_N=\sum_{k=1}^{N} f(n_k x)$. Then the sequence $\{S_N, N\geqq 1\}$ can be redefined on a new probability space (without changing its distribution) together with a Wiener-process $\zeta(t)$ such that

$$S_N = \zeta(\tau_N)+o(N^{1/2-\lambda}) \quad a.s. \tag{1.5}$$

where $\lambda>0$ is an absolute constant and τ_N is an increasing sequence of random variables such that $\tau_N/b_N\to 1$ a.s. where

$$b_N = \int_0^1 \left(\sum_{k=1}^{N} f(n_k x)\right)^2 dx. \tag{1.6}$$

Condition (1.4) in Theorem 1 cannot be omitted as it is shown by the example $f(x)=\cos 2\pi x-\cos 4\pi x$, $n_k=2^k$ (cf. [5]).

Actually, the proof of Theorem 1 will yield the following result which gives some information about what happens if we replace D_∞ by D_m in Theorem 1.

THEOREM 2. *Let $\varepsilon>0$. Then there exists an $m=m(\varepsilon, f)$ with the property that if we replace the condition $\{n_k\}\in D_\infty$ by $\{n_k\}\in D_m$ in Theorem* 1 *then the statement remains valid with the modification that for the random variable τ_N in* (1.5) *we have*

$$1-\varepsilon \leqq \varliminf_{N\to\infty} \frac{\tau_N}{b_N} \leqq \varlimsup_{N\to\infty} \frac{\tau_N}{b_N} \leqq 1+\varepsilon \quad a.s. \tag{1.7}$$

instead of $\tau_N/b_N\to 1$ a.s.

In other words, if $\{n_k\}$ belongs to D_m with a large m ("large" here depends also on f) then the conclusion of Theorem 1 remains "almost" valid.

It is easy to see (cf. [2], Lemmas (2.1), (2.2) and their proofs) that Theorem 1 implies Donsker's invariance principle (functional CLT) and the functional LIL for $f(n_k x)$. These limit theorems need not be valid under the conclusion of Theorem 2 but even under Theorem 2 we can state at least that $f(n_k x)$ obeys Donsker's invariance principle and the functional LIL "approximately". Roughly speaking, the smaller the ε in (1.7) is, the more precisely $f(n_k x)$ satisfies the above mentioned limit theorems. (For precise details of this statement via "ε-limit theorems" see [2], § 4.)

Let us say that a rational number $r>0$ is of order k if in the reduced form $r=p/q$ the greater of p and q is k. The following lemma is easy to prove (cf. the proof of Lemma (3.1) in [2]):

LEMMA (1.1). *The (set-theoretic) union of the sequences $\{n_k\}, \{2n_k\}, \dots, \{mn_k\}$ satisfies the Hadamard gap condition if and only if for any subsequences n_{k_i}, n_{l_i} and any rational number $r>0$ of order $\leqq m$ the relations*

$$\lim_{i\to\infty} \frac{n_{k_i}}{n_{l_i}} = r, \quad \frac{n_{k_i}}{n_{l_i}} \neq r \quad (i = 1, 2, \dots)$$

are impossible.

The condition of Lemma (1.1) is satisfied e.g. if

a) $n_{k+1}/n_k > m \quad (k = 1, 2, \dots)$,

b) $\lim_{k\to\infty} n_{k+1}/n_k = \alpha$ where α is a rational number of order $> m$.

Since the Hadamard gap condition implies condition B_2, in examples a), b) the sequence $\{n_k\}$ belongs to D_m. For examples for sequences D_∞ see [3] or [2], § 3.

It follows from example a), Theorem 2 and our remarks above that if $\{n_k\}$ satisfies (1.3) with a large q then $f(n_k x)$ almost satisfies the CLT, the LIL and their functional versions (this was also proved in [2], § 4). Example b) shows that the same conclusion holds if n_{k+1}/n_k tends to a rational number of great order (e.g. to a rational number very close to an integer).

2. Two preparatory lemmas

In what follows, $\|f\|$ and $\|f\|_\infty$ will denote the L_2 and L_∞ norm of f, resp. For two numerical sequences a_n, b_n the relation $a_n \sim b_n$ means $\lim_{n\to\infty} a_n/b_n = 1$.

LEMMA (2.1). *Let $g(x)$ $(-\infty<x<+\infty)$ be a measurable function such that*

$$g(x+1) = g(x), \quad \int_0^1 g(x)\,dx = 0.$$

Then for any real $a<b$ and $\lambda>0$ we have

$$\Big|\int_a^b g(\lambda x)\,dx\Big| \leqq \frac{2}{\lambda}\int_0^1 |g(x)|\,dx.$$

This is Lemma (3.2) of [1].

For the formulation of the next lemma we notice that if f satisfies (1.1) and the Lipschitz α condition then

$$\|f - s_n(f)\|_\infty \leqq A n^{-\alpha/2} \tag{2.1}$$

where A is a positive constant and $s_n(f)$ denotes the n-th partial sum of the Fourier-series of f. (See [9], p. 64.)

LEMMA (2.2). *Let $f(x)$ satisfy* (1.1) *and the Lipschitz α condition and let $1 \leqq n_1 < n_2 < \ldots < n_N$ be a sequence of positive numbers satisfying* (1.3). *Then, if $N \geqq N_0$ where N_0 depends on $f(x)$ and q, we have for any real a*

$$\int_a^{a+1} \left(\sum_{k=1}^{N} f(n_k x)\right)^2 dx \leqq C_2(\|f\|^3 + \|f\|^2 + \|f\|)N \tag{2.2}$$

and

$$\int_0^1 \left(\sum_{k=1}^{N} f(n_k x)\right)^4 dx \leqq C_3 N^2 \tag{2.3}$$

where C_2 depends on q and on the numbers A, α in (2.1) *and C_3 depends on $f(x)$ and q.*

In view of the remark preceding Lemma (2.2), relation (2.2) follows from Lemma (3.3) of [1]; on the other hand, by Lemma (3.4) of [1] we have for $N \geqq N_0$

$$\sum_{k=1}^{N} f(n_k x) = \xi_1 + \xi_2 \tag{2.4}$$

where ξ_1 and ξ_2 are random variables (functions) on [0, 1] such that, if P denotes the Lebesgue measure, then we have

$$P(|\xi_1| \geqq y\sqrt{N}) \leqq C_4 e^{-C_5 y} \quad (y \geqq 0) \quad \text{and} \quad \|\xi_2\|_\infty \leqq 1 \tag{2.5}$$

where C_4 and C_5 depend on $f(x)$ and q. (As a matter of fact, Lemma (3.4) of [1] assumes $\|f - s_n(f)\| \leqq An^{-\alpha}$ instead of (2.1) and states correspondingly $\|\xi_2\| \leqq 1$ instead of $\|\xi_2\|_\infty \leqq 1$ but the proof there applies with trivial changes in the present case too.) Evidently (2.4) and (2.5) imply (2.3).

3. Main lemma

We first approximate the functions $f(n_k x)$ by step-functions $\varphi_k(x)$ as follows. By assumption, $f(x)$ satisfies in [0, 1] the Lipschitz α condition for some $0 < \alpha \leqq 1$. Let now $2^l \leqq n_k < 2^{l+1}$, put $p = \left[l + \frac{20}{\alpha} \log k\right]$ and let $\varphi_k(x)$ denote the function in [0, 1] which is constant in the intervals $[i2^{-p}, (i+1)2^{-p})$ $(0 \leqq i \leqq 2^p - 1)$ and these constant values coincide with the respective values of $f(n_k x)$ at the points $i2^{-p}$ $(0 \leqq i \leqq 2^p - 1)$. By the Lipschitz α condition we have

$$|f(n_k x) - \varphi_k(x)| \leqq C\left(\frac{n_k}{2^p}\right)^\alpha \leqq C\left(\frac{2^{l+1}}{2^{l+(20/\alpha)\log k - 1}}\right)^\alpha \leqq C \cdot 2^{-20 \log k} \leqq Ck^{-10}. \tag{3.1}$$

(Here and in the sequel, C will denote positive constants, not always the same, depending (at most) on $f(x)$ and q.) Let us now divide the set of positive integers into disjoint blocks $I_1, J_1, I_2, J_2, \ldots$ in such a way that I_k contains $[k^{1/2}]$ integers, J_k contains $[k^{1/4}]$ integers $(k = 1, 2, \ldots)$. Let

$$T_k = \sum_{\nu \in I_k} f(n_\nu x), \quad D_k = \sum_{\nu \in I_k} \varphi_\nu(x). \tag{3.2}$$

Then by (3.1)

$$|D_k - T_k| \leqq C \sum_{\nu \in I_k} \nu^{-10} \leqq C \sum_{\nu=[(k-1)^{1/2}]}^{\infty} \nu^{-10} \leqq Ck^{-4} \tag{3.3}$$

and thus using $|D_k| \leqq Ck^{1/2}$, $|T_k| \leqq Ck^{1/2}$ and the mean value theorem we get

$$|D_k^2 - T_k^2| \leqq C, \quad |D_k^4 - T_k^4| \leqq C. \tag{3.4}$$

Now we formulate our

MAIN LEMMA. *We have*

$$\left|E(D_k | D_1, \dots, D_{k-1})\right| \leqq Ck^{-2} \quad (k \geqq k_0) \tag{3.5}$$

$$E(D_k^2 | D_1, \dots, D_{k-1}) \leqq Ck^{1/2} \quad (k \geqq k_0) \tag{3.6}$$

$$\sum_{k=1}^{N} E(D_k^2 | D_1, \dots, D_{k-1}) \sim d_N \quad a.s. \quad as \quad N \to \infty \tag{3.7}$$

$$ED_k^4 \leqq Ck \quad (k \geqq k_0) \tag{3.8}$$

where $d_N = \sum_{k=1}^{N} ED_k^2$. *Also,* $CN^{3/2} \leqq d_N \leqq CN^{3/2}$ *for* $N \geqq N_0$.

PROOF. We begin with the proof of (3.5). Let $\mathcal{F}_{k-1}$ denote the σ-field generated by $D_1, \dots, D_{k-1}$. In view of (3.3) it suffices to show that

$$\left|E(T_k | \mathcal{F}_{k-1})\right| \leqq Ck^{-2}. \tag{3.9}$$

Let $b = b(k)$ denote the largest integer of the block I_{k-1}, let l be an integer such that $2^l \leqq n_b < 2^{l+1}$ and put $w = \left[l + \frac{20}{\alpha} \log b\right]$. From the definition of φ_k it follows that every φ_ν, $1 \leqq \nu \leqq b$ takes a constant value on each interval of the form

$$[i2^{-w}, (i+1)2^{-w}) \quad (0 \leqq i \leqq 2^w - 1) \tag{3.10}$$

and thus every set $\{D_1 = a_1, \dots, D_{k-1} = a_{k-1}\}$ where $a_1, \dots, a_{k-1}$ are constants, can be obtained as an union of intervals of the form (3.10). In other terms, $\mathcal{F}_{k-1}$ is purely atomic and each of its atoms is a union of intervals of the form (3.10). Hence to prove (3.9) it suffices to show that

$$\left|2^w \int_{i2^{-w}}^{(i+1)2^{-w}} T_k \, dx\right| \leqq Ck^{-2} \quad (0 \leqq i \leqq 2^w - 1). \tag{3.11}$$

Let $c = c(k)$ denote the smallest integer of the block I_k. By (1.3) we have

$$\sum_{\nu \in I_k} \frac{1}{n_\nu} \leqq \sum_{j=c}^{\infty} \frac{1}{n_\nu} \leqq \frac{1}{n_c}(1 + q^{-1} + q^{-2} + \dots) = \frac{q}{q-1} \frac{1}{n_c}$$

and

$$\frac{n_b}{n_c} \leqq q^{-(c-b)} = q^{-[(k-1)^{1/4}]-1} \leqq q^{-(k-1)^{1/4}}.$$

Hence applying Lemma (2.1) and using the trivial relation $b \leqq 2k^{3/2}$ we get

$$(3.12)\qquad \left|2^w \int_{i2^{-w}}^{(i+1)2^{-w}} T_k\,dx\right| = \left|2^w \int_{i2^{-w}}^{(i+1)2^{-w}} \sum_{\nu\in I_k} f(n_\nu x)\right| dx \leqq$$

$$\leqq 2^w C \sum_{\nu\in I_k} \frac{2}{n_\nu} \leqq C\frac{2^w}{n_c} \leqq C\frac{2^{l+(20/\alpha)\log b}}{n_c} \leqq C\frac{n_b}{n_c} b^{20/\alpha} \leqq Cq^{-(k-1)^{1/4}} k^{30/\alpha} \leqq Ck^{-2}$$

and thus (3.11) is proved.

To prove (3.6) it suffices to show (in view of (3.4)) that $E(T_k^2|\mathcal{F}_{k-1}) \leqq Ck^{1/2}$ and since $\mathcal{F}_{k-1}$ is atomic and each of its atoms is a union of intervals of the form (3.10), the last relation will follow if we show that

$$2^w \int_{i2^{-w}}^{(i+1)2^{-w}} T_k^2\,dx \leqq Ck^{1/2} \quad (0 \leqq i \leqq 2^w-1,\ k \geqq k_0).$$

The integral on the left hand side is equal to

$$(3.13)\qquad 2^w \int_{i2^{-w}}^{(i+1)2^{-w}} \Big(\sum_{\nu\in I_k} f(n_\nu x)\Big)^2 dx = \int_i^{i+1} \Big(\sum_{\nu\in I_k} f(m_\nu t)\Big)^2 dt$$

where $m_\nu = 2^{-w} n_\nu$. Evidently $m_{\nu+1}/m_\nu \geqq q > 1$ for all the ν's appearing here. If $c = c(k)$ denotes the smallest integer of I_k as above, then the smallest of the m_ν's is $m_c = n_c/2^w$ which is at least 1 (in fact it is $\geqq Ck^2$ by a part of the estimate (3.12)). Hence by Lemma (2.2) the integral on the right side of (3.13) is $\leqq Ck^{1/2}$ for $k \geqq k_0$ which was to be proved.

To prove (3.8) it suffices to remark that, by Lemma (2.2), we have $ET_k^4 \leqq Ck$ which, together with (3.4), implies (3.8).

We now turn to the proof of (3.7). We proceed in three steps.

a) Let $\mathcal{L}_1 \subset \mathcal{L}_2 \subset \dots$ be any increasing sequence of σ-fields such that D_k is $\mathcal{L}_k$ measurable. Then the relations

$$(3.14)\qquad \sum_{k=1}^N D_k^2 \sim d_N \quad \text{a.s.}$$

and

$$(3.15)\qquad \sum_{k=1}^N E(D_k^2|\mathcal{L}_{k-1}) \sim d_N \quad \text{a.s.}$$

are equivalent. Indeed, the sequence $H_k = D_k^2 - E(D_k^2|\mathcal{L}_{k-1})$ is a square integrable martingale difference sequence (and consequently orthogonal) with $EH_k^2 \leqq \leqq 4ED_k^4 \leqq Ck$ by Minkowski's inequality and (3.8). Hence the Rademacher—Mensov convergence theorem implies the a.s. convergence of $\sum_{k=1}^\infty k^{-3/2} H_k$ and thus by the Kronecker lemma we have

$$(3.16)\qquad \frac{1}{N^{3/2}} \sum_{k=1}^N H_k \to 0 \quad \text{a.s.}$$

By condition (1.4) of the theorem $ET_k^2 \geqq Ck^{1/2}$ for $k \geqq k_0$ hence by (3.4) $ED_k^2 \geqq$

$\geqq Ck^{1/2}$ and thus $d_N \geqq CN^{3/2}$ for $N \geqq N_0$. Also, by Lemma (2.2), $ET_k^2 \leqq Ck^{1/2}$ for $k \geqq k_0$, hence by (3.4) $ED_k^2 \leqq Ck^{1/2}$ and $d_N \leqq CN^{3/2}$. (We thus proved the last statement of the main lemma.) Therefore (3.16) implies $\sum_{k=1}^{N} H_k = o(d_N)$ which really shows that (3.14) and (3.15) are equivalent.

b) We now prove (3.7) in the special case when f is a trigonometric polynomial:

$$f = \sum_{k=1}^{m} (a_k \cos 2\pi kx + b_k \sin 2\pi kx). \tag{3.17}$$

By a) it suffices to show (3.14) or, what is the same,

$$\sum_{k=1}^{N} (D_k^2 - ED_k^2) = o(d_N) \quad \text{a.s.} \tag{3.18}$$

By (3.4) and $d_N \geqq CN^{3/2}$, (3.18) will follow from

$$\sum_{k=1}^{N} (T_k^2 - ET_k^2) = o(N^{3/2}) \quad \text{a.s.} \tag{3.19}$$

Let us express the left side of (3.19) as a trigonometric polynomial, using (3.17). By (3.17),

$$f(n_\nu x) = \sum_{j=1}^{m} (a_k \cos 2\pi j n_\nu x + b_k \sin 2\pi j n_\nu x)$$

and thus

$$T_k = \sum_{\nu \in I_k} f(n_\nu x) = \sum (c_l \cos 2\pi \lambda_l x + d_l \sin 2\pi \lambda_l x) \tag{3.20}$$

where all the λ_l's are of the form tn_ν, $1 \leqq t \leqq m$, $\nu \in I_k$. Denoting by N_m the (set-theoretic) union of the sequences $\{n_k\}, \{2n_k\}, \ldots, \{mn_k\}$, this means that all the λ_l's belong to N_m. Also, for the coefficients c_l, d_l in (3.20) we have

$$|c_l| \leqq \overline{M}, \quad |d_l| \leqq \overline{M} \tag{3.21}$$

where $\overline{M}$ depends only on $f(x)$ and $\{n_k\}$. Indeed, the trigonometric sums $f(n_\nu x)$ and $f(n_\mu x)$, $\nu < \mu$ can overlap (i.e. contain a term with the same frequency) only if $n_\mu \leqq mn_\nu$ i.e. overlapping is impossible if $\mu - \nu \geqq p$ where p is the smallest integer such that $q^p > m$. This remark shows that $|c_l| \leqq pM_1$, $|d_l| \leqq pM_1$ where $M_1 = \max(|a_1|, |b_1|, \ldots, |a_m|, |b_m|)$ and hence (3.21) is valid.

We notice also that the trigonometric sums in (3.20) are pairwise non-overlapping for $k \geqq k_0$. This follows from the fact that the largest λ_l in T_{k-1} is mn_b and the smallest λ_l in T_k is n_c where b and c are the largest integer of the block I_{k-1} and the smallest integer of the block I_k, resp. By the separation of I_{k-1} and I_k by the block J_{k-1} of length $[(k-1)^{1/4}]$ and because of (1.3) we have $mn_b < n_c$ for $k \geqq k_0$.

Squaring (3.20) and using well known trigonometric identities we get

$$T_k^2 = \frac{1}{2} \sum (c_l^2 + d_l^2) + \sum (e_i \cos 2\pi \varrho_i x + f_i \sin 2\pi \varrho_i x) \tag{3.22}$$

where $|e_i|\leqq\overline{M}^2$, $|f_i|\leqq\overline{M}^2$ and the ϱ_i's are the numbers of the form $\lambda_s\pm\lambda_r$ with the λ's appearing in (3.20). Hence summing (3.22) for $M+1\leqq k\leqq M+N$ (but not collecting the terms with equal frequencies) we get that

$$\sum_{k=M+1}^{M+N} T_k^2 = B+\sum (r_i\cos 2\pi\theta_i x+s_i\sin 2\pi\theta_i x) \tag{3.23}$$

where B is a constant, $|r_i|\leqq\overline{M}^2$, $|s_i|\leqq\overline{M}^2$ and the θ_i's are the numbers of the form $\lambda_s\pm\lambda_r$ where λ_s and λ_r are from the same T_k, $M+1\leqq k\leqq M+N$. Since the T_k's are non-overlapping for $k\geqq k_0$ and N_m satisfies condition B_2, there is a constant C_1 such that at most C_1 of the θ_i's can be equal. Hence collecting the terms with equal frequency on the right hand side of (3.23) we get

$$\sum_{k=M+1}^{M+N} T_k^2 = B+\sum (u_j\cos 2\pi jx+v_j\sin 2\pi jx) \tag{3.24}$$

where the sum on the right hand side is finite and $|u_j|\leqq C_1\overline{M}^2$, $|v_j|\leqq C_1\overline{M}^2$. Also, the number of terms on the right hand side of (3.20) is $\leqq mk^{1/2}$, hence in the second sum of (3.22) is $\leqq m^2k$ and on the right side of (3.23), (3.24) is $\leqq \sum_{k=M+1}^{M+N} m^2k\leqq$ $\leqq m^2N(M+N)\leqq m^2[(M+N)^2-M^2]$. The number B in (3.24) is evidently equal to the expectation of $\sum_{k=M+1}^{M+N} T_k^2$ (since the integral of the trigonometric sum on the right hand side is 0). Hence (3.24) implies

$$E\left(\sum_{k=M+1}^{M+N}(T_k^2-ET_k^2)\right)^2=\frac{1}{2}\sum(u_j^2+v_j^2)\leqq$$

$$\leqq\frac{1}{2}C_1^2\overline{M}^4m^2\big((M+N)^2-M^2\big)=C_2\big((M+N)^2-M^2\big).$$

Applying the Gal—Koksma law of large numbers (see [6], p. 134) for the variables $T_k^2-ET_k^2$, we get

$$\sum_{k=1}^{N}(T_k^2-ET_k^2)=O(N\log^3 N)\quad \text{a.s.}$$

and thus (3.19) is proved.

c) Let now f be any function satisfying (1.1) and the Lipschitz α condition and fix an $\varepsilon>0$. Since the Fourier series of f converges uniformly to f (even (2.1) is valid) we can write $f=f_1+f_2$ where f_1 is a suitable partial sum of the Fourier series of f (hence it is a trigonometric polynomial) and $\|f_2\|_\infty\leqq\varepsilon$. (Evidently f_1 and f_2 also satisfy (1.1) and they are also Lipschitz α functions.) In the same way as we constructed the step-function $\varphi_k(x)$ from $f(n_kx)$, we can construct $\varphi_k^{(1)}(x)$ and $\varphi_k^{(2)}(x)$ from $f_1(n_kx)$ and $f_2(n_kx)$, resp. Then we have

$$T_k=T_k^{(1)}+T_k^{(2)}\quad\text{and}\quad D_k=D_k^{(1)}+D_k^{(2)}$$

where

$$T_k^{(1)} = \sum_{\nu\in I_k} f_1(n_\nu x), \quad T_k^{(2)} = \sum_{\nu\in I_k} f_2(n_\nu x)$$

$$D_k^{(1)} = \sum_{\nu\in I_k} \varphi_\nu^{(1)}(x), \quad D_k^{(2)} = \sum_{\nu\in I_k} \varphi_\nu^{(2)}(x).$$

Evidently (3.3), (3.4) hold for the D_k, T_k's with superscripts, too:

$$|D_k^{(i)}-T_k^{(i)}| \leqq Ck^{-4}, \quad |(D_k^{(i)})^2-(T_k^{(i)})^2| \leqq C, \quad |(D_k^{(i)})^4-(T_k^{(i)})^4| \leqq C \quad i=1,2. \tag{3.25}$$

If $b=b(k)$ is the largest integer of I_{k-1}, $2^l \leqq n_b < 2^{l+1}$, $w=\left[l+\frac{20}{\alpha}\log b\right]$ then, as we showed, every φ_ν, $1\leqq\nu\leqq b$ and therefore also D_{k-1}, takes a constant value on each interval of the form (3.10). In other words, if $\mathcal{G}_{k-1}$ denotes the σ-field generated by the intervals (3.10), then D_{k-1} is $\mathcal{G}_{k-1}$ measurable. Since $\varphi_\nu^{(1)}$ and $\varphi_\nu^{(2)}$ are step-functions with the same intervals of constancy as φ_ν, not only D_{k-1} but also $D_{k-1}^{(1)}$ and $D_{k-1}^{(2)}$ are $\mathcal{G}_{k-1}$ measurable.

Let $d_N^{(1)} = \sum_{k=1}^{N} E(D_k^{(1)})^2$. Since f_1 is a trigonometric polynomial, the relation

$$\sum_{k=1}^{N} (D_k^{(1)})^2 \sim d_N^{(1)} \quad \text{a.s.} \tag{3.26}$$

is exactly what we proved in b). As we remarked above, $D_k^{(1)}$ is $\mathcal{G}_k$ measurable and hence by the equivalence statement of a) (3.26) implies

$$\sum_{k=1}^{N} E\big((D_k^{(1)})^2|\mathcal{G}_{k-1}\big) \sim d_N^{(1)} \quad \text{a.s.} \tag{3.27}$$

We now prove two simple estimates

$$E\big((D_k^{(1)})^2|\mathcal{G}_{k-1}\big) \leqq Ck^{1/2} \quad (k \geqq k_0) \tag{3.28}$$

$$E\big((D_k^{(2)})^2|\mathcal{G}_{k-1}\big) \leqq C\varepsilon k^{1/2} \quad (k \geqq k_0) \tag{3.29}$$

which, together with (3.27), will easily lead to our aim (3.7).

The proofs of (3.28) and (3.29) are the same, we prove e.g. (3.29). In view of (3.25) it suffices to show

$$E\big((T_k^{(2)})^2|\mathcal{G}_{k-1}\big) \leqq C\varepsilon k^{1/2} \quad (k \geqq k_0)$$

and since $\mathcal{G}_{k-1}$ is atomic with atoms of the form (3.10), the last inequality is equivalent to

$$2^w \int_{i2^{-w}}^{(i+1)2^{-w}} (T_k^{(2)})^2\,dx \leqq C\varepsilon k^{1/2} \quad (0 \leqq i \leqq 2^w-1, \ k \geqq k_0). \tag{3.30}$$

Here the left-hand side can be written as

$$2^w \int_{i2^{-w}}^{(i+1)2^{-w}} (T_k^{(2)})^2\,dx = 2^w \int_{i2^{-w}}^{(i+1)2^{-w}} \Big(\sum_{\nu\in I_k} f_2(n_\nu x)\Big)^2 dx = \int_i^{i+1} \Big(\sum_{\nu\in I_k} f_2(m_\nu t)\Big)^2 dt \tag{3.31}$$

where $m_\nu = n_\nu/2^w$. Exactly in the same way as in the case of the second integral in (3.13), the numbers m_ν are all greater than 1 and they satisfy $m_{\nu+1}/m_\nu \geqq q > 1$. Let us also observe that since f_1 is a partial sum of the Fourier series of f i.e. $f_2 = f - s_k(f)$ with a certain k, (2.1) is inherited for f_2 with the same A, α. Since $\|f_2\| \leqq \|f_2\|_\infty \leqq \varepsilon$, an application of Lemma (2.2) gives that the last integral in (3.31) is at most $C\varepsilon k^{1/2}$ for $k \geqq k_0$ and thus (3.30) is valid.

To deduce (3.7) from (3.27), (3.28), (3.29) let us first integrate (3.29) to get $E(D_k^{(2)})^2 \leqq C\varepsilon k^{1/2}$ for $k \geqq k_0$. On the other hand, by assumption (1.4) of the theorem we have $ET_k^2 \geqq Ck^{1/2}$ for $k \geqq k_0$ and thus (3.4) implies $ED_k^2 \geqq Ck^{1/2}$. We thus get $E(D_k^{(2)})^2/ED_k^2 \leqq C\varepsilon$ $(k \geqq k_0)$ whence we obtain, using $D_k = D_k^{(1)} + D_k^{(2)}$ and Minkowski's inequality,

$$1 - C\sqrt{\varepsilon} \leqq E(D_k^{(1)})^2/ED_k^2 \leqq 1 + C\sqrt{\varepsilon} \quad (k \geqq k_0)$$

and consequently

$$(1 - C\sqrt{\varepsilon})d_N < d_N^{(1)} < (1 + C\sqrt{\varepsilon})d_N \quad (N \geqq N_0). \tag{3.32}$$

Summing up (3.29) for $k = 1, 2, \ldots, N$ and using $d_N \geqq CN^{3/2}$ we obtain

$$\sum_{k=1}^{N} E\big((D_k^{(2)})^2 | \mathcal{G}_{k-1}\big) \leqq C\varepsilon d_N \quad (N \geqq N_0). \tag{3.33}$$

Also, (3.28), (3.29) and Schwarz's inequality imply $\left|E(D_k^{(1)} D_k^{(2)} | \mathcal{G}_{k-1})\right| \leqq C\sqrt{\varepsilon}\, k^{1/2}$ whence

$$\left|\sum_{k=1}^{N} E(2D_k^{(1)} D_k^{(2)} | \mathcal{G}_{k-1})\right| \leqq C\sqrt{\varepsilon} d_N \quad (N \geqq N_0). \tag{3.34}$$

Adding (3.27), (3.33), (3.34) and using $D_k = D_k^{(1)} + D_k^{(2)}$ and (3.32) we see that

$$(1 - C\sqrt{\varepsilon})\, d_N \leqq \sum_{k=1}^{N} E(D_k^2 | \mathcal{G}_{k-1}) \leqq (1 + C\sqrt{\varepsilon})\, d_N$$

a.s. for sufficiently large N which implies, since $\varepsilon > 0$ was arbitrary,

$$\sum_{k=1}^{N} E(D_k^2 | \mathcal{G}_{k-1}) \sim d_N \quad \text{a.s.}$$

Since D_k is $\mathcal{G}_k$ measurable, the last relation implies (3.14) and (3.7) by the equivalence statement of a). Hence the proof of the main lemma is complete.

REMARK 1. In the proof of the lemma above, the assumption $\{n_k\} \in D_\infty$ was used only in the proof of (3.7); relations (3.5), (3.6), (3.8) are valid under the mere assumption (1.3). We also see that if f is a trigonometric polynomial of order m:

$$f = \sum_{k=1}^{m} (a_k \cos 2\pi kx + b_k \sin 2\pi kx)$$

then for the validity of (3.7) it suffices to assume $\{n_k\} \in D_m$ instead of $\{n_k\} \in D_\infty$ (see step b) of the proof of (3.7)). Together with step c), this shows that if $\{n_k\} \in D_m$

and $\|f-s_m(f)\|_\infty=\varepsilon_0$ then instead of (3.7) we have

$$(3.35)\qquad (1-C\sqrt{\varepsilon_0})\,d_N \leqq \sum_{k=1}^{N} E(D_k^2|\mathscr{G}_{k-1}) \leqq (1+C\sqrt{\varepsilon_0})\,d_N$$

a.s. for sufficiently large N where $\mathscr{G}_k$ is the σ-field defined above. Since D_k is $\mathscr{G}_k$ measurable and by step a) we have (under (1.3))

$$\sum_{k=1}^{N} D_k^2 - \sum_{k=1}^{N} E(D_k^2|\mathscr{L}_{k-1}) = o(d_N) \quad \text{a.s.}$$

for any increasing sequence $\mathscr{L}_k$ of σ-fields such that D_k is $\mathscr{L}_k$ measurable, (3.35) implies

$$(1-C\sqrt{\varepsilon_0})\,d_N \leqq \sum_{k=1}^{N} E(D_k^2|D_1, \dots, D_{k-1}) \leqq (1+C\sqrt{\varepsilon_0})\,d_N$$

a.s. for sufficiently large N. In other words, if $\{n_k\}\in D_m$ for a large m (here "large" depends also on f) then (3.7) is satisfied "approximately".

Let

$$\bar{D}_k = D_k - E(D_k|D_1, \dots, D_{k-1}).$$

Then for the $\bar{D}_k$'s the main lemma implies the following

LEMMA (3.2). *We have*

$$(3.36)\qquad E(\bar{D}_k|\bar{\mathscr{F}}_{k-1}) = 0$$

$$(3.37)\qquad E(\bar{D}_k^2|\bar{\mathscr{F}}_{k-1}) \leqq Ck^{1/2} \quad (k \geqq k_0)$$

$$(3.38)\qquad \sum_{k=1}^{N} E(\bar{D}_k^2|\bar{\mathscr{F}}_{k-1}) \sim d_N \quad a.s.$$

$$(3.39)\qquad E\bar{D}_k^4 \leqq Ck \quad (k \geqq k_0)$$

where $\bar{\mathscr{F}}_{k-1}$ denotes the σ-field generated by $\bar{D}_1, \dots, \bar{D}_{k-1}$.

PROOF. We have $|\bar{D}_k - D_k| \leqq Ck^{-2}$ (see (3.5)), $|D_k| \leqq Ck^{1/2}$, $|\bar{D}_k| \leqq Ck^{1/2}$ and hence by the mean value theorem

$$(3.40)\qquad |\bar{D}_k^2 - D_k^2| \leqq C, \quad |\bar{D}_k^4 - D_k^4| \leqq C.$$

The second relations of (3.40) and (3.8) evidently imply (3.39), furthermore (3.6) and the first relation of (3.40) imply $E(\bar{D}_k^2|D_1, \dots, D_{k-1}) \leqq Ck^{1/2}$ from which (3.37) follows by taking conditional expectations of both sides with respect to $\bar{\mathscr{F}}_{k-1}$ (which is contained in the σ-field generated by $D_1, \dots, D_{k-1}$). In step a) of the proof of (3.7) we saw that (3.7) is equivalent to $\sum_{k=1}^{N} D_k^2 \sim d_N$ a.s.; now the first relation of (3.40) and $d_N \geqq CN^{3/2}$ show that also $\sum_{k=1}^{N} \bar{D}_k^2 \sim d_N$ a.s. is an equivalent statement. Finally, the martingale argument of step a) and (3.39) show that $\sum_{k=1}^{N} \bar{D}_k^2 \sim d_N$ implies (3.38).

REMARK 2. The main lemma concerns the "long" block sums D_k and T_k. Defining the "short" block sums

$$T_k' = \sum_{\nu \in J_k} f(n_\nu x), \quad D_k' = \sum_{\nu \in J_k} \varphi_\nu(x) \tag{3.41}$$

an analogous statement holds for these sums:

$$\left|E(D_k'|\mathscr{F}_{k-1}')\right| \leqq Ck^{-2}, \quad E\big((D_k')^2|\mathscr{F}_{k-1}'\big) \leqq Ck^{1/4},$$

$$\sum_{k=1}^{N} E\big((D_k')^2|\mathscr{F}_{k-1}'\big) \sim d_N' \quad \text{a.s.}, \quad E(D_k')^4 \leqq Ck^{1/2}$$

where $\mathscr{F}_{k-1}'$ denotes the σ-field generated by $D_1', \dots, D_{k-1}'$ and $d_N' = \sum_{k=1}^{N} E(D_k')^2$. Also, $CN^{5/4} \leqq d_N' \leqq CN^{5/4}$ for $N \geqq N_0$. The analogue of Lemma (3.2) also holds for the centered sums $\bar{D}_k' = D_k' - E(D_k'|\mathscr{F}_{k-1}')$.

4. Conclusion of the proof

Using the main lemma and Lemma (3.2) we can complete the proofs of Theorem 1, 2 in a standard way, following [1] or [6]. We prove here Theorem 1; the proof of Theorem 2 is the same (see Remark 1 after the proof of the main lemma). Let

$$V_N = \sum_{k=1}^{N} E(\bar{D}_k^2|\bar{D}_1, \dots, \bar{D}_{k-1}),$$

then $V_N \sim d_N$ a.s. by (3.38). Also, using (3.39) and $d_N \geqq CN^{3/2}$ (see the main lemma) we see that the sum $\sum_{k=1}^{\infty} d_k^{-3/2} E\bar{D}_k^4$ is convergent. By Beppo Levi's theorem this implies the a.s. convergence of the series

$$\sum_{k=1}^{\infty} d_k^{-3/2} E(\bar{D}_k^4|\bar{D}_1, \dots, \bar{D}_{k-1})$$

and since the general term of the series

$$\sum_{k=1}^{\infty} \frac{1}{V_k^{3/4}} \int_{x^2 > V_k^{3/4}} x^2 \, dP(\bar{D}_k < x|\bar{D}_1, \dots, \bar{D}_{k-1}) \tag{4.1}$$

can be majorized by

$$\frac{1}{V_k^{3/2}} \int_{-\infty}^{+\infty} x^4 \, dP(\bar{D}_k < x|\bar{D}_1, \dots, \bar{D}_{k-1}) = \frac{1}{V_k^{3/2}} E(\bar{D}_k^4|\bar{D}_1, \dots, \bar{D}_{k-1}) \sim$$

$$\sim \frac{1}{d_k^{3/2}} E(\bar{D}_k^4|\bar{D}_1, \dots, \bar{D}_{k-1}),$$

it follows that the series (4.1) is also a.s. convergent. Thus we can apply Theorem (4.4) of [7] to the martingale difference sequence $\bar{D}_k$ with $f(x) = x^{3/4}$ and we get that there exists a Wiener-process $\zeta(t)$ such that

$$\bar{D}_1+\ldots+\bar{D}_k=\zeta(V_k)+o(V_k^{1/2-\eta}) \quad \text{a.s.} \tag{4.2}$$

with an absolute constant $\eta>0$. (Strictly speaking, we first have to redefine the sequence $\bar{D}_k$ on a new, larger probability space and $\zeta(t)$ will be defined over this new space; in the sequel, however, we will speak as if (4.2) were valid for the original sequence. This little inaccuracy essentially simplifies the formulas (we do not have to use stars or superscripts for the "redefined" variables) and does not cause any trouble.) Replacing $\bar{D}_1+\ldots+\bar{D}_k$ with $T_1+\ldots+T_k$ on the left hand side of (4.2), we commit an error $O(1)$ (since $|\bar{D}_k-T_k|\leqq|\bar{D}_k-D_k|+|D_k-T_k|=O(k^{-2})$ by (3.3) and (3.5)); hence (4.2) and $V_k\sim d_k$ a.s. imply

$$T_1+\ldots+T_k=\zeta(V_k)+o(d_k^{1/2-\eta}) \quad \text{a.s.} \tag{4.3}$$

(In what follows η will denote positive absolute constants, possibly different at different places.) We also remark that by replacing the left hand side of (4.3) by $T_1+T_1'+\ldots+T_k+T_k'$ (T_i' are the short block sums defined in (3.41)) we only add a term which is $o(d_k^{1/2-\eta})$, so it does not bother the right side of (4.3). (Indeed, (4.3) has the exact analogue

$$T_1'+\ldots+T_k'=\zeta(V_k')+o(d_k'^{1/2-\eta}) \quad \text{a.s.} \tag{4.4}$$

for the short block sums where $V_k'=\sum_{j=1}^{k}E(\bar{D}_j'^2|\bar{D}_1',\ldots,\bar{D}_{j-1}')$, $d_k'=\sum_{j=1}^{k}E(D_j')^2$, cf. Remark 2 at the end of § 3. Now it is sufficient to observe that the analogue of (3.37) to the 'primed' variables $\bar{D}_k'$ i.e. $E(\bar{D}_k'^2|\mathcal{F}_{k-1}')\leqq Ck^{1/4}$ implies $V_k'=O(k^{5/4})$ and thus $d_k'\leqq Ck^{5/4}$ and the standard estimate $\zeta(t)=o(t^{1/2}\log t)$ show that the right hand side of (4.4) is $o(k^{5/8}\log k)+o(k^{5/8-\eta})$ which is dominated by the remainder term $o(d_k^{1/2-\eta})=o(k^{3/4-3\eta/2})$ in (4.3) if η is small enough.) Hence (4.3) implies

$$T_1+T_1'+\ldots+T_k+T_k'=\zeta(V_k)+o(d_k^{1/2-\eta}) \quad \text{a.s.}$$

which can be rewritten as

$$S_{N_k}=\zeta(V_k)+o(d_k^{1/2-\eta}) \quad \text{a.s.} \tag{4.5}$$

where $S_N=\sum_{\nu=1}^{N}f(n_\nu x)$ and $N_k=\sum_{i=1}^{k}([i^{1/2}]+[i^{1/4}])\sim\frac{2}{3}k^{3/2}$. Since $CN^{3/2}\leqq d_N\leqq CN^{3/2}$ the remainder term in (4.5) can be also written as $o(N_k^{1/2-\eta})$. Hence if we define a sequence τ_n of random variables by

$$\begin{cases}\tau_0=0\\ \tau_{N_k}=V_k \quad \text{for} \quad k=1,2,\ldots\\ \tau_n \quad \text{is linear in the intervals } N_k\leqq n\leqq N_{k+1} \text{ for } k=0,1,\ldots\ (N_0=0)\end{cases} \tag{4.6}$$

then (4.5) simply says that the relation

$$S_N=\zeta(\tau_N)+o(N^{1/2-\eta}) \quad \text{a.s.} \tag{4.7}$$

is valid for the indices $N=N_k$. To get (4.7) for general N it suffices to show

(4.8) $$\max_{N_k \leqq N \leqq N_{k+1}} |S_N - S_{N_k}| = o(N_k^{1/2-\eta}) \quad \text{a.s.}$$

and

(4.9) $$\max_{N_k \leqq N \leqq N_{k+1}} |\zeta(\tau_N) - \zeta(\tau_{N_k})| = o(N_k^{1/2-\eta}) \quad \text{a.s.}$$

The first relation is trivial since

$$|S_N - S_{N_k}| = \left| \sum_{\nu=N_k+1}^{N} f(n_\nu x) \right| \leqq C(N-N_k) \leqq C(N_{k+1}-N_k) =$$
$$= C([(k+1)^{1/2}]+[(k+1)^{1/4}]) \sim Ck^{1/2} \leqq CN_k^{1/2-\eta}.$$

To see (4.9) let us note that

$$\tau_{N_k} = V_k = O(k^{3/2}) \quad \text{a.s.}$$

and

$$\max_{N_k \leqq N \leqq N_{k+1}} |\tau_N - \tau_{N_k}| = \tau_{N_{k+1}} - \tau_{N_k} = V_{k+1} - V_k = O(k^{1/2}) \quad \text{a.s.}$$

by (3.37) and thus Lemma (3.6) of [1] (with $r=3/2$, $s=1/2$) shows the left side of (4.9) is $O(k^{1/4} \log k) = O(N_k^{1/2-\eta})$ a.s. Hence (4.7) is proved and it remains to show that $\tau_N/b_N \to 1$ a.s. where b_N is defined in (1.6). To this end we notice that $\tau_N/e_N \to 1$ a.s. where e_N is the numerical sequence defined (in analogy with (4.6)) by

$$\begin{cases} e_0 = 0 \\ e_{N_k} = d_k \quad \text{for} \quad k = 1, 2, \dots \\ e_n \quad \text{is linear in the intervals } N_k \leqq n \leqq N_{k+1} \text{ for } k = 0, 1, \dots . \end{cases}$$

(This follows trivially from the piecewise linearity of τ_n and e_n and the relation $V_k/d_k \to 1$ a.s. which is identical with (3.38).) Since $N_k \sim \frac{2}{3} k^{3/2}$ and $e_{N_k} = d_k \geqq Ck^{3/2}$ we have $e_n \geqq Cn$ and thus the remainder term in (4.7) can also be written as $o(e_N^{1/2-\eta})$. Hence (4.7) and $\tau_n \sim e_n$ a.s. imply that the distribution of $S_N/\sqrt{e_N}$ tends to the standard normal distribution. Since the L_4 norm of $S_N/\sqrt{e_N}$ remains bounded (this follows from the second relation of Lemma (2.2) and $e_n \geqq Cn$), the second moment of $S_N/\sqrt{e_N}$ converges to the second moment of the standard normal distribution, i.e. to 1. In other words, $ES_N^2/e_N \to 1$ and since here $ES_N^2 = b_N$ (see (1.6)), we see that the sequences e_N and b_N are asymptotically equal. Thus $\tau_N/e_N \to 1$ a.s. implies $\tau_N/b_N \to 1$ a.s. and this completes the proof of Theorem 1.

References

[1] I. Berkes, On the asymptotic behaviour of $\sum f(n_k x)$ I. Main theorems, *Zeitschrift für Wahrscheinlichkeitstheorie verw. Gebiete*, **34** (1976), 319—345.

[2] I. Berkes, On the asymptotic behaviour of $\sum f(n_k x)$ II. Applications, *Zeitschrift für Wahrscheinlichkeitstheorie verw. Gebiete*, **34** (1976), 347—365.

[3] V. F. Gaposhkin, On the central limit theorem for some weakly dependent sequences (in Russian), *Teor. Verojatn. i Primenen*, **15** (1970), 666—684. English translation: *Theory Prob. and Appl.*, **15** (1970), 649—666.

[4] M. KAC, On the distribution of values of sums of the type $\sum f(2^k t)$, *Ann. Math.*, **47** (1946), 33—49.
[5] M. KAC, Probability methods in some problems of analysis and number theory, *Bull. Amer. Math. Soc.*, **55** (1949), 641—665.
[6] W. PHILIPP and W. F. STOUT, Almost sure invariance principles for partial sums of weakly dependent random variables. *Memoirs of the AMS*, no. 161.
[7] V. STRASSEN, Almost sure behaviour of sums of independent random variables and martingales, *Proc. 5th Berkeley Sympos. Math. Statist. and Probab.* Vol. II. (Part I), 315—343, Univ. of California Press (Berkeley, 1967).
[8] S. TAKAHASHI, The law of the iterated logarithm for a gap sequence with infinite gaps, *Tohoku Math. Journ.*, **15** (1963), 281—288.
[9] A. ZYGMUND, *Trigonometric series*, Vol. I., Cambridge University Press (1959).

(Received September 20, 1977)

MATHEMATICAL INSTITUTE OF THE
HUNGARIAN ACADEMY OF SCIENCES
1053 BUDAPEST, REÁLTANODA U. 13—15
HUNGARY

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS 61801
USA

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 157—163.

О ПОВЕДЕНИИ ПРИ $t\to\infty$ РЕШЕНИЙ ВЫРОЖДАЮЩИХСЯ КВАЗИЛИНЕЙНЫХ ПАРАБОЛИЧЕСКИХ УРАВНЕНИЙ

Р. КЕРШНЕР (БУДАПЕШТ)

§ 1. Введение

Мы будем рассматривать задачу Коши для уравнения

$$Lu \equiv -u_t + (u^\mu)_{xx} - c\cdot u^\nu = 0 \tag{1}$$

в полуплоскости $R_+^2 = \{(t, x)\colon\ 0 \leqq t < \infty,\ -\infty < x < \infty\}$ с начальным условием

$$u(0, x) = u_0(x), \quad -\infty < x < \infty. \tag{2}$$

Здесь $\mu > 1$, $\nu > 0$ и $c > 0$ — постоянные. Относительно функции $u_0(x)$ мы будем предполагать следующее: $u_0(x) = 0$ при $|x| \geqq l$, $u_0(x) > 0$ при $|x| < l$, $u_0(x)$ удовлетворяет условию Гельдера с показателем $\min\{1, (\mu-1)^{-1}\}$ (условие Q).

Уравнение (1) появляется во многих физических задачах. Оно описывает, например, процесс теплопередачи при наличии поглощения, когда коэффициенты теплопроводности и поглощения степенным образом зависят от температуры.

Уравнения вида (1) называются «вырождающимися квазилинейными параболическими уравнениями второго порядка», так как они являются параболическими при $u > 0$ и вырождаются в уравнение первого порядка при $u = 0$.

Известно (см. [1]), что задача (1)—(2) может не иметь классического (т. е. имеющего одну непрерывную производную по t и две по x) решения. Поэтому нужно рассматривать обобщенные решения.

Определение 1. Пусть G — замкнутая подобласть R_+^2. Неотрицательная и ограниченная в G функция $u(t, x)$, удовлетворяющая условию Гельдера, называется обобщённым решением уравнения (1) в G, если для $u(t, x)$ выполняется интегральное тождество

$$I(u, f;\ t_0, t_1, x_0, x_1) \equiv \int_{t_0}^{t_1} \int_{x_2}^{x_1} (uf_t + u^\mu f_{xx} - cu^\nu f)\, dx\, dt - \tag{3}$$

$$- \int_{x_0}^{x_1} uf\, dx \Big|_{t_0}^{t_1} - \int_{t_0}^{t_1} u^\mu f_x\, dt \Big|_{x_0}^{x_1} = 0,$$

каковы бы ни были числа $t_0 < t_1$, $x_0 < x_1$, такие, что прямоугольник $P = [t_0, t_1] \times [x_0, x_1]$ содержится в G, и функция $f(t, x) \in C_{t,x}^{1,2}(P)$, равная нулю при $x = x_0$, $x = x_1$.

Определение 1'. Обобщенным решением задачи (1)—(2) называется обобщённое решение уравнения (1) в R^2_+, удовлетворяющее начальному условию (2).

Однозначная обобщенная разрешимость задачи (1)—(2) доказана в [2].

Как известно (см. [3]), при $\nu<1$ имеет место явление «полного остывания за конечное время»: существует такое $T_0>0$, что $u(t,x)=0$ при всех $t\geqq T_0$ и $x\in R^1$. С нашей точки зрения этот случай неинтересен и всюду в дальнейщем будем предполагать, что

$$\nu\geqq 1. \tag{4}$$

Если выполнено это условие, то $u(t,x_0)>0$ при $t<0$, если только $u(0,x_0)>0$. В работе [4] доказано, что в этом случае $u(t,x)$ удовлетворяет условию Гельдера по x с показателем $\min\{1,(\mu-1)^{-1}\}$, причём этот показатель точен. Гельдеровость по t следует из работы [7].

Известно, что в случае задачи (1)—(2) «возмущение распространяется с конечной скоростью»: при любом $t_0>0$ функция $u(t_0,x)$ финитна по x (см. напр. [3]; впрочем, он является простым следствием результатов настоящей работы).

Существуют две непрерывные кривые $x=\zeta_i(t)$ $(i=1,2)$, такие, что $\zeta_1(t)<0$, $\zeta_2(t)>0$, $\zeta_1(t)$ не возрастает, $\zeta_2(t)$ не убывает и

$$\operatorname{supp} u(t,x)=\{(t,x)\colon\ t\geqq 0,\ \zeta_1(t)\leqq x\leqq\zeta_2(t)\}.$$

Сформулируем два утверждения, нужных нам в дальнейщем. Сначала дадим

Определение 2. Ограниченная в G функция $y(t,x)$, удовлетворяющая условию Гельдера, называется обобщенным суперрешением уравнения (1) в G, если (см. (3)) $I(y,f;t_0,t_1,x_0,x_1)\leqq 0$ для любого $P=[t_0,t_1]\times[x_0,x_1]\subset G$ и любой неотрицательной функции $f(t,x)\in C^{1,2}_{t,x}(P)$, равной нулю при $x=x_0$, $x=x_1$.

Замечание 1. Если ограниченная функция $y(t,x)\geqq 0$ удовлетворяет в G условию Гельдера, является гладкой вне конечного числа непрерывных кривых вида $x=\xi(t)$, удовлетворяет там неравенству $Ly\geqq 0$ и производная $\partial y^\mu/\partial x$ непрерывна при $x=\xi(t)$, то с помощью интегрирования по частям легко убедиться, что $y(t,x)$ является обобщенным суперрешением уравнения (I) в G

Лемма 1. *Пусть* $u(t,x)$ — *обобщенное решение задачи* (1)—(2), *а* $y(t,x)$ *обобщенное суперрешение уравнения* (1) *в* R^2_+, *причем* $u_0(x)\leqq y(0,x)$ *для* $-\infty<x<\infty$. *Тогда* $u(t,x)\leqq y(t,x)$ *всюду в* R^2_+.

Обозначим через H криволинейную трапецию

$$\{(t,x)\colon\ 0\leqq t_0<t\leqq t_1<\infty,\quad \xi_0(t)<x<\xi_1(t);\ \xi_i(t)\in C([t_0,t_1)\}.$$

Пустъ Γ — ее параболическая граница.

Лемма 2. *Пусть* $u(t,x)$ — *обобщенное решение задачи* (1)—(2), *а* $z(t,x)\in C^{1,2}_{t,x}(\overline{H}\setminus\Gamma)\cap C(\overline{H})$, *причем* $u(t,x)\geqq z(t,x)$ *на* Γ, *а в* $\overline{H}\setminus\Gamma$ *выполнены неравенства* $z>0$, $Lz>0$. *Тогда* $u(t,x)\geqq z(t,x)$ *всюду в* H.

Доказательства этих утверждений можно найти в [2].

Целью настоящей работы является получение двусторонних оценок при $t>1$ для обобщенного решения $u(t, x)$ задачи (1)—(2). Параллельно с этими результатами получаются и оценки для функций $\zeta_i(t)$.

§ 2. Случай $\nu>\mu$

Теорема 1. *Пусть $u(t, x)$ — обобщенное решение задачи* (1)—(2). *Тогда существуют такие положительные постоянные $a_1, a_2, \ldots, a_6$, зависящие лишь от данных задачи, что при $t>1$ справедливы неравенства*

(i) $u(t, x)\leqq a_1 t^{-1/(\mu+1)}$,

(ii) $|\zeta_i(t)|\leqq a_2 t^{1/(\mu+1)}$, $i=1, 2$.

Если $\nu<\mu+2$, *то*

(iii) $u(t, x)\geqq a_3 t^{-1/(\nu-1)}$ *при достаточно малом* $|x|$,

(iv) $|\zeta_i(t)|\geqq a_4 t^{(\nu-\mu)/2(\nu-1)}$.

Если $\nu\geqq\mu+2$, *то*

(v) $u(t, x)\geqq a_5 t^{-1/(\mu+1)}$ *при достамочно малом* $|x|$,

(vi) $|\zeta_i(t)|\geqq a_6 t^{1/(\mu+1)}$.

Замечание 2. Оценки, касающиеся $\zeta_i(t)$, впервые были доказаны Б. Ф. Кнерром [6] другим способом, чем это сделано ниже.

Замечание 3. Нетрудно заметить, что при $\nu\to\mu+2-0$ разности между показателями в правых частях (i), и (iii), а также (ii) и (iv), стремятся к нулю.

Доказательство (i) *и* (ii). Положим $A=L^2-x^2(t+1)^{-\beta}$ и рассмотрим в R_+^2 функцию

$$y(t, x)=\begin{cases}\sigma(t+1)^{-\alpha}A^{\omega} & \text{при } A>0\\ 0 & \text{при } A\leqq 0,\end{cases}$$

где положительные постоянные α, β, σ и L будут выбраны позже, а

$$\omega=\frac{1}{\mu-1}. \tag{5}$$

(Это обозначение сохраняется на протяжении всей работы). При $A>0$ находим:

$$\begin{aligned} Ly &= \sigma\omega x^2(t+1)^{-\alpha-\beta-1}A^{\omega-1}[-\beta+4\sigma^{\mu-1}\mu(\omega\mu-1)(t+1)^{-\alpha(\mu-1)-\beta+1}]+\\ &+\sigma(t+1)^{-\alpha\mu-\beta}A^{\omega\mu-1}[-2\omega\mu\sigma^{\mu-1}+\alpha(t+1)^{\alpha(\mu-1)+\beta-1}]-c\sigma^{\nu}(t+1)^{-\alpha\nu}A^{\omega\nu}. \end{aligned} \tag{6}$$

Пусть $\alpha=(\mu+1)^{-1}$, $\beta=2(\mu+1)^{-1}$ и $\sigma=\left(\dfrac{\mu-1}{2\mu(\mu+1)}\right)^{\omega}$. Тогда из (6) получаем $Ly=-c\sigma^{\nu}(t+1)^{-\alpha\nu}A^{\omega\nu}<0$ при $A>0$ и любом L. Так как $Ly=0$ при $A<0$, то $Ly \leqq 0$ в $R^2_+ \setminus \{(t, x): A=0\}$. Из неравенства $\omega\mu > 1$ следует, что $\partial y^\mu/\partial x$ непрерывна при $A=0$. По Замечанию 1, $y(t, x)$ является обобщенным суперрешением уравнения (1) в R^2_+. Далее,

$$y(0, x) = \begin{cases} \sigma(L^2 - x^2)|^\omega & \text{при } |x| < L \\ 0 & \text{при } |x| \geqq L. \end{cases}$$

Поэтому $y(0, x) \geqq u_0(x)$ при $x \in R^1$, если $\sigma(L^2 - l^2)^\omega > M = \sup u_0(x)$, то есть если $L^2 > (M\sigma^{-1})^{\mu-1} + l^2$. При таком выборе параметров, входящих в $y(t, x)$, из Леммы 1 следует, что $u(t, x) \leqq y(t, x)$ всюду в R^2_+. Этим доказаны неравенства (i) и (ii).

Доказательство (iii) *и* (iv). Рассмотрим в области

$$H_1 = \{(t, x): t > 0, \ A_1 \equiv \varrho - x^2(t+\tau)^{-\beta} > 0\}$$

вспомогательную функцию $z_1 = \sigma(t+\tau)^{-\alpha} A_1^\omega$, где положительные постоянные $\alpha, \beta, \sigma, \tau$ и ϱ будут выбраны позже, Пологая $\theta = t + \tau$ имеем:

$$Lz_1 = \omega\sigma x^2 \theta^{-\alpha\mu - 2\beta} A_1^{\omega\mu - 2}[4\omega\mu\sigma^{\mu-1} - \beta\theta^{\alpha(\mu-1)+\beta-1}] + \tag{7}$$
$$+ \sigma\theta^{-\alpha-1} A_1^\omega[\alpha - 2\omega\mu\sigma^{\mu-1}\theta^{-\alpha(\mu-1)-\beta+1} - c\sigma^{\nu-1}\theta^{-\alpha(\nu-1)+1} A_1^{\omega(\nu-1)}].$$

Положим $\alpha(\mu-1) = 1 - \beta$. Для того, чтобы из (7) следовало неравенство $Lz_s > 0$ в H_1 при достаточно большом τ и достаточно малом ϱ, мы должны требовать, чтобы было

$$\alpha \geqq \frac{1}{\nu-1}, \quad \alpha > 2\omega\mu\sigma^{\mu-1}, \ \beta \leqq 4\omega\mu\sigma^{\mu-1}. \tag{8}$$

Пусть $\alpha = (\nu-1)^{-1}$. Тогда $\beta = \dfrac{\nu-\mu}{\nu-1}$. Для совместности неравенств (8) относительно σ необходимо и достаточно, чтобы выполнялось условие $\nu < \mu + 2$.

При указанном выборе α, β и σ, взяв достаточно малое ϱ и достаточно большое τ, получим $Lz_1 > 0$ в H_1 и $u_0(x) \geqq z_1(0, x)$ на основании H_1. После этого утверждения (iii) и (iv) следуют из Леммы 2.

Доказательство (v) и (vi). Рассмотрим в области H_1 функцию $z_1(t, x)$, где $\alpha = \dfrac{1}{\mu+1} + \varepsilon$ и $\beta = \dfrac{2}{\mu+1} - \varepsilon(\mu-1)$, $\varepsilon > 0$ — произвольное. Первое из неравенств (8) выполняется всилу предположения $\nu \geqq \mu + 2$. Второе и третее неравенства (8) относительно σ совместны, так как $2\alpha > \beta$. Аналогично тому, как выше, получаем $u(t, x) \geqq a_5 t^{-1/\mu+1-\varepsilon}$ при достаточно малом $|x|$,

$$|\zeta_i(t)| \geqq a_6 t^{1/\mu+1-\varepsilon(\mu-1)/2},$$

где a_5 и a_6 не зависят от ε. Устремив ε к нулю, приходим к оценкам (v), (vi).

§ 3. Случай $\mu=\nu$

Теорема 2. *Пусть $u(t, x)$ — обобщенное решение задачи* (1)—(2). *Тогда существуют такие положительные постоянные a_7, a_8, a_9 и a_{10}, зависящие лишь от данных задачи, что при $t>1$ справедливы неравенства*

(i) $u(t, x)\geqq a_7 t^{-1/(\mu-1)}$ *при достаточно малом* $|x|$,

(ii) $|\zeta_i(t)|\geqq a_8\sqrt{\ln t},\quad i=1, 2,$

(iii) $u(t, x)\leqq a_9 t^{-1/(\mu-1)}$,

(iv) $|\zeta_i(t)|\leqq a_{10} t^{1/(\mu-1)},\quad i=1, 2.$

Замечание 4. Неравенство (iv) доказывается так же, как неравенство (ii) Теоремы 1. Неравенство (ii) впервые доказал Б. Ф. Кнерр [6] другим способом, чем это сделано ниже.

Доказательство (i) и (ii). Рассмотрим в области $H_2=\{(t, x)\colon t>0, A_2\equiv\varrho-x^2\ln^{-\beta}(t+\tau)>0\}$ вспомогательную функцию $z_2(t, x)=(t+\tau)^{-\alpha}[\varrho-$ $-x^2\ln^{-\beta}(t+\tau)]^{\omega}$, где положительные постоянные α, β, ϱ и τ будут выбраны ниже.

Полагая $\theta=t+\tau$, при $A_2>0$ находим

$$Lz_2=\alpha\theta^{-\alpha-1}A_2^{\omega}-\beta\omega x^{2-\alpha-1}\ln_0^{-\beta-1}\theta A_2^{\omega-1}-2\omega\mu\theta^{-\alpha\mu}\ln_0^{-\beta}A_2^{\omega\mu-1}+$$

$$+4\omega^2\mu x^2\theta^{-\alpha\mu}\ln_0^{-2\beta}A_2^{\omega\mu-2}-c\theta^{-\alpha\mu}A_2^{\omega\mu}\equiv I_1+\ldots+I_5.$$

далее,

$$I_2+I_4=\omega x^2\theta^{-\alpha\mu}\ln_0^{-2\beta}A_2^{\omega-1}[4\omega\mu-\beta\theta^{\alpha(\mu-1)-1}\ln_0^{\beta-1}],$$

$$I_1+I_3+I_5=\theta^{-\alpha-1}A_2^{\omega}[\alpha-2\omega\mu\theta^{-\alpha(\mu-1)+1}\ln^{-\beta}\theta-c\theta^{-\alpha(\mu-1)+1}A_2].$$

Если мы положим $\alpha=\omega$ и $\beta=1$, то получим, что $I_2+I_4>0$ в H_2. По условию Q мы можем предполагать, что $u_0(x)\geqq\varepsilon>0$ при $|x|\leqq\delta$. Значит, $z_2(0, x)\leqq$ $\leqq u_0(x)$ на основании области H_2, если $\varrho\ln\tau\leqq\delta^2$ и $\varrho^{\omega}\leqq\varepsilon\tau^{\omega}$. Пусть $\tau>1$ и $\varrho=\tau^{-1}$. Тогда $z_1(0, x)\leqq u_0(x)$, если τ настолько велико, что $\tau^{-1}\ln\tau\leqq\delta^2$ и $\tau^{-2\omega}\leqq\varepsilon$. На боковой странице H_2 имеем $z_2(t, x)=0\leqq u(t, x)$. За счет возможного увеличения τ можно достичь того, чтобы одновременно с неравенством $z_2(t, x)\leqq u(t, x)$ на параболической границе H_2, выполнялось и неравенство $I_1+I_3+I_5>0$ в H_2. По Лемме 2 получаем, что $u(t, x)\geqq z_2(t, x)$ в H_2. Отсюда следуют оценки (i), (ii).

Доказательство (iii). Пусть $M=\sup u_0(x)$. Рассмотрим в R_+^2 вспомогательную функцию

$$y_2(t, x)=[c(\mu-1)t+M^{-\mu+1}]^{-\omega}.$$

Она является обобщенным решением уравнения (1) и $y_2(0, x)=M\geqq u_0(x)$. Поэтому неравенство $y_2(t, x)\geqq u(t, x)$ следует из Леммы 1. Теорема доказана.

4. Случай $\nu > \mu$

По причинам, объясненным во введении, мы предположим, что $\nu \geqq 1$, причем случаи $\nu > 1$ и $\nu = 1$ будут рассматриваться отдельно.

Теорема 3. *Пусть $u(t, x)$ — обобщенное решение задачи* (1)—(2) *и* $\nu > 1$. *Тогда существуют такие положительные постоянные a_{11}, a_{12} и a_{13}, что при $t > 1$ выполнены следующие неравенства:*

(i) $u(t, 0) \geqq a_{11} t^{-1/(\nu-1)}$,

(ii) $u(t, x) \leqq a_{12} t^{-1/(\nu-1)}$,

(iii) $|\zeta_i(t)| \leqq a_{13} \quad (i = 1, 2)$.

Доказательство. Неравенство (iii) утверждает, по определению, что в данном случае имеет место локализация возмущений — факт, известный из [2]. Доказательство неравенства (ii) аналогично доказательству последнего утверждения Теоремы 2. Вместо $y_2(t, x)$ нужно рассмотреть функцию

$$y_3(t, x) = [c(\nu-1)t + M^{-(\nu-1)}]^{-1/(\nu-1)}.$$

Для доказательства пункта (i) рассмотрим в области $H = \{(t, x): t > 0,\ A \equiv \varrho - x^2 \ln (t+\tau) > 0\}$ вспомогательную функцию $z(t, x) = (t+\tau)^{-\alpha} [\varrho - x^2 \ln (t+\tau)]^{\omega}$, где положительные постоянные α, ϱ и τ будут выбраны ниже. Сначала выберем $\varrho > 0$ и $\tau > 1$ таким образом ($\varrho \leqq \varrho_1$, $\tau \geqq \tau_1$), чтобы выполнялось неравенство $z(0, x) \leqq u_0(x)$ на основании области H. Полагая $\theta = t + \tau$ при $A > 0$ находим

$$Lz = \alpha\theta^{-\alpha-1}A^{\omega} + \omega x^2 \theta^{-\alpha-1} A^{\omega-1} - 2\omega\mu\theta^{-\alpha\mu} \ln \theta A^{\omega\mu-1} +$$
$$+ 4\omega^2 \mu x^2 \theta^{-\alpha\mu} \ln^2 \theta A^{\omega\mu-1} - c\theta^{-\alpha\nu} A^{\omega\nu} \equiv I_1 + \ldots + I_5.$$

Ясно, что $I_2 + I_4 > 0$ в H. Далее,

$$I_1 + I_3 + I_5 = \theta^{-\alpha-1} A^{\omega} [\alpha - 2\omega\mu\theta^{-\alpha(\mu-1)+1} \ln \theta - c\theta^{-\alpha(\nu-1)+1} A^{\omega(\mu-1)}].$$

Положим $\alpha = (\nu+1)^{-1}$. При этом $I_1 + I_3 + I_5 > 0$ в H, если

$$\frac{1}{\nu-1} > 2\omega\mu\theta^{-\mu-\nu/\nu-1} \ln \theta + c\varrho^{\omega(\nu-1)}. \tag{9}$$

Взяв достаточно малое ϱ ($\varrho \leqq \varrho_2$) и достаточно большое τ ($\tau \geqq \tau_2$), мы можем достичь выполнения неравенства (9). Поэтому при $\varrho < \min (\varrho_1, \varrho_2)$ и $\tau > \max (\tau_1, \tau_2)$ $Lz > 0$ в H, что доказывает (i) ввиду Леммы 2. Теорема доказана.

Перейдем к случаю $\nu = 1$, который является особым с точки зрения поведения решений на бесконечности. А именно, имеет место следующая

Теорема 4. *Пусть $u(t, x)$ — обобщенное решение задачи* (1)—(2), *где* $\nu = 1$. *Тогда существуют такие положительные постоянные a_{14} и a_{15}, что при $t > 0$ имеют место неравенства*

(i) $u(t, x) \geqq a_{14} e^{-ct}$ *при достаточно малом* $|x|$,

(ii) $u(t, x) \leqq a_{15} e^{-ct}$.

Доказательство. Оба неравенства следуют из следующего замечания (см. [5], [8]): обобщенным решением задачи

$$v_t = (v^\mu)_{xx} - cv, \quad t > 0, \quad -\infty < x < \infty,$$

$$v(0, x) = \begin{cases} Dh^{-\omega}(\delta^2 - x^2)^\omega & \text{при} \quad |x| < \delta \\ 0 & \text{при} \quad |x| \geqq \delta, \end{cases}$$

где $D = [2\omega\mu(\mu+1)^{-\omega}]$, $h > 0$, $\delta > 0$ — любые, является функция

$$v(t, x) = \begin{cases} De^{-ct}[g(t, h)]^{-1/\mu+1}\{\delta^2 h^{-2/\mu+1} - x^2[g(t, h)]^{-2/\mu+1}\}^\omega & \text{при} \quad \{\dots\} > 0 \\ 0 & \text{при} \quad \{\dots\} \leqq 0, \end{cases}$$

где

$$g(t, h) = \frac{1 - e^{-c(\mu-1)t}}{c(\mu-1)} + h.$$

С помощью Лемм 1 и 2 и этой функции можно оценить и снизу, и сверху решение $u(t, x)$. Для доказательства (i) нужно выбрать $\delta > 0$ достаточно малым, а $h > 0$ достаточно большим. Для доказательства неравенства (ii) — наоборот.

Литература

[1] Г. И. Баренблатт, Об автомодельных решениях задачи Коши для нелинейного параболического уравнения нестационарной фильтрации газа в пористой среде, *Прикл. матем. и механ.*, **20(6)** *(1956)*.

[2] А. С. Калашников, О характере распространения возмущений в процессах, описываемых квазилинейными вырождающимися параболическими уравнениями, *Труды семинара им. И. Г. Петровского*, **1** *(1975)*.

[3] А. С. Калашников, О характере распространения возмущений в задачах нелинейной теплопроводности с поглощением, *Ж. выч. мат. и мат. физ.*, **14(4)** *(1974)*.

[4] А. С. Калашников, О дифференциальных свойствах обобщенных решений уравнений типа нестационарной фильтрации, *Вестник МГУ, мат.*, **1** (1974).

[5] А. С. Калашников, О влиянии поглощения на распространение тепла в среде с теплопроводностью, зависящей от температуры, *Журн. вычисл. мат. и матем. физ.* **16(3)** (1976).

[6] B. F. Knerr, *The behavior of the support of solutions of the equation of nonlinear heat conduction with absorption in one dimension*. Preprint, 1977.

[7] С. Н. Кружков, Результаты о характере непрерывности решений параболических уравнений и некоторые их применения, *Мат. зам.*, **6(1)** (1969).

[8] Л. К. Мартинсон, К. Б. Павлов, К вопросу о пространственной локализации тепловых возмущений в теории нелинейной теплопроводности, *Журн. вычисл. мат. и мат. физ.*, **12(4)** (1972).

(Поступило 9.12.1977.)

R. KERSNER
MTA SZTAKI
NUMERIKUS MÓDSZEREK OSZTÁLYA
1014 BUDAPEST, ÚRI U. 49.

Acta Mathematica Academiae Scientiarum Hungarica
Tomus 34 (1—2), (1979), 165—176.

MULTIPLE PACKING AND COVERING OF SPHERES

By

G. FEJES TÓTH (Budapest)

Dedicated to Professor A. Florian on his 50th birthday

A system of equal spheres is said to form a *k-fold packing* if each point of the space belongs to at most k spheres. Analogously, a system of equal spheres is said to form a *k-fold covering* if each point of the space belongs to at least k spheres. Let δ_k^n be the supremum of the densities of all k-fold packings of equal spheres in Euclidean n-space. Similarly, let Δ_k^n be the infimum of the densities of all k-fold coverings of the Euclidean n-space with equal spheres. Obviously, we have the trivial bounds

$$\delta_k^n \leqq k \leqq \Delta_k^n.$$

In a series of papers [5, 6, 7, 8], L. FEW obtained rather good upper bounds for δ_k^n when n is large compared with k. But his bounds are worse than the trivial bound k if $k > \sqrt{n}$ as well as for small values of k and n. In his survey on multiple packing of spheres held on the Colloquium on Convexity in Copenhagen in 1964, L. FEW [7] posed the problem to find a non-trivial upper bound for δ_k^n valid for all values of k and n. As far as I know, no non-trivial lower bound for Δ_k^n was known for $n>2$ and $k>1$. A. FLORIAN [10] proved that

$$\delta_k^n < k < \Delta_k^n,$$

but his method does not enable him to deduce explicit non-trivial bounds for δ_k^n and Δ_k^n. In a recent paper [4] I showed that δ_k^2 is at most k times the density of one circle in the circumscribed regular $6k$-gon and Δ_k^2 is at least k times the density of one circle with respect to the inscribed regular $6k$-gon. The aim of this paper is to give a non-trivial upper bound for δ_k^n and a non-trivial lower bound for Δ_k^n for all values of k and n.

In order to formulate our results we have to introduce some notations. If nothing else is explicitly stated, we shall work in Euclidean n-space. Let S be a unit sphere and $S(r)$ the sphere of radius r concentric with S. Let $m_k(r)$ be the maximum of the number of unit spheres which can be placed in $S(r)$ so as to form together with S a k-fold packing. Let $M_k(r)$ be the minimum of the number of unit spheres which, together with S, cover the sphere $S(r)$ k times. Let $v(f)$ be the minimal number such that the boundary of any polyhedral region whose faces are contained in at most f hyperplanes can be decomposed into $v(f)$ convex sets. We define the quantity $w(f, k)$ by

$$w(f, k) = \min\left\{v(f), k^2\binom{f}{k}\right\}.$$

Let J_n and ω_n be the volume and the surface-volume of the n-dimensional unit sphere, respectively. Let $V(\vartheta)$ be the volume of a right spherical cone of altitude 1 and solid angle ϑ $(0<\vartheta<\omega_n/2)$.

Now we are in the position to phrase our main results:

THEOREM 1. *We have*

$$\delta_k^n \leqq kJ_n\Bigg[\max_{\substack{r>1\\ l=w(m_k(2r+1),k)}}\ \min_{0\leqq\vartheta_0\leqq\omega_n}\left\{\vartheta_0\frac{r^n}{n}+lV\left(\frac{\omega_n-\vartheta_0}{l}\right)\right\}\Bigg]^{-1}. \tag{1}$$

THEOREM 2. *We have*

$$\Delta_k^n \geqq k\int_0^1 \frac{dt}{[1-C^\delta(1-t^\delta)]^{n/2}}, \tag{2}$$

where
$$\delta=\frac{2}{n-1}\quad \textit{and}\quad C=\frac{nJ_n}{2w(M_k(3),k)J_{n-1}}.$$

The fact that these bounds are non-trivial, which is not difficult to show directly, will be obvious from the proofs.

The proof of both theorems rests on the investigation of certain polytopes associated with the spheres. Let $\mathcal{S}$ be a system of different unit spheres. We associate with each sphere $S\in\mathcal{S}$ the *k'th Dirichlet cell* D_S^k of S consisting of those points of the space to which there are at most $k-1$ centres of the spheres of $\mathcal{S}$ nearer than the centre of S. (Of course the k'th Dirichlet cell depends also on $\mathcal{S}$ but we shall not indicate this in our notations.) For $k=1$ we obtain the ordinary Dirichlet cells. We emphasise the following important properties of the k'th Dirichlet cells:

(i) *If $\mathcal{S}$ contains at least k spheres then the k'th Dirichlet cells cover the space exactly k times.* More precisely, each point of the space which is not a boundary-point of a k'th Dirichlet cell is an interior point of exactly k k'th Dirichlet cells. For, let P be an arbitrary point and let $O_1, O_2, \ldots$ be the centres of the spheres, choosing the notations so that $PO_1\leqq PO_2\leqq\ldots\leqq PO_k\leqq PO_{k+1}\leqq\ldots$. Then P lies in the interior of the k'th Dirichlet cells of the spheres with centres $O_1, \ldots, O_k$ except when $PO_k=PO_{k+1}$. In this latter case P lies on the common boundary of the k'th Dirichlet cells of the spheres centred at O_k and O_{k+1}.

(ii) *If $\mathcal{S}$ forms a k-fold packing then each sphere of $\mathcal{S}$ is contained in its k'th Dirichlet cell. If $\mathcal{S}$ forms a k-fold covering then each sphere of $\mathcal{S}$ contains its k'th Dirichlet cell.* For, suppose that $\mathcal{S}$ forms a k-fold packing and P is an inner point of the sphere $S\in\mathcal{S}$ which does not belong to D_S^k. Then, by definition, there are at least k spheres of $\mathcal{S}$, whose centres lie nearer to P than the centre of S. But this means that P belongs to more than k spheres of $\mathcal{S}$, which is a contradiction. This proves the statement for the k-fold packings. The case when $\mathcal{S}$ forms a k-fold covering can be settled in a similar way.

(iii) *D_S^k is a polyhedral region which is star-convex with respect to the centre O of S.* In order to see this we have only to observe that D_S^k can be constructed in the following way: Let us draw for each sphere $S'\in\mathcal{S}$, $S'\neq S$ the radical hyperplane of S and S'. Then D_S^k consists of all points P of the space with the property that the segment OP intersects at most $k-1$ of these hyperplanes.

Now we turn to the proof of Theorem 1. Let $\mathscr{S}$ be a k-fold packing of unit spheres. Without loss of generality we may assume that $\mathscr{S}$ does not contain coincident spheres. Then we can construct to each sphere $S\in\mathscr{S}$ its k'th Dirichlet cell D_S^k. In order to give a non-trivial upper bound for the density of $\mathscr{S}$, it sufficies, by property (i), to give a uniform non-trivial lower bound for the volume of D_S^k. This can be done in the following way:

Instead of D_S^k we consider the set $S(r)\cap D_S^k$, where, in accordance with our previous notations, $S(r)$ denotes the sphere of radius r concentric with S. We shall assume that $r>1$. Each face of D_S^k lies in the radical hyperplane of S and some other sphere of $\mathscr{S}$. Hence it is obvious that from the point of view of the construction of the set $S(r)\cap D_S^k$ those spheres whose centres lie outside of the sphere $S(2r)$ are irrelevant. Thus if $\bar{\mathscr{S}}$ is the set of those spheres of $\mathscr{S}$ which are contained in $S(2r+1)$ and $\bar{D}_S^k$ is the k'th Dirichlet cell of S with respect to $\bar{\mathscr{S}}$ then we have

$$S(r)\cap D_S^k = S(r)\cap \bar{D}_S^k.$$

We divide the boundary of $\bar{D}_S^k$ into a minimal number of convex sets and consider the non-empty intersections $F_1, F_2, \dots, F_l$ of these sets with $S(r)$. Let F_0 be the part of the boundary of $S(r)$ which lies in $\bar{D}_S^k$. Let C_i be the convex cone with apex at the centre O of S based on F_i $(i=1, \dots, l)$. Let C_0 be the "cone-like" body based on F_0 with apex O. Let ϑ_i be the solid angle of C_i at O $(i=0, \dots, l)$. Let $|X|$ denote the volume of the body X. Then we have

$$|C_0| = \vartheta_0\frac{r^n}{n} \quad\text{and}\quad |C_i|\geqq V(\vartheta_i) \quad\text{for}\quad i=1, \dots, l. \tag{3}$$

The first relation is obvious. To see the validity of the inequalities $|C_i|\geqq V(\vartheta_i)$, $i=1, \dots, l$, we consider the right spherical cone R_i with altitude 1, solid angle ϑ_i and base G_i such that the centre O_i of G_i lies on the half-line emanating from O in the direction of the point of F_i lying nearest to O. Let F_i' be the intersection of the cone C_i with the hyperplane containing G_i. Let C_i' be the cone with apex O based on F_i'. In view of property (ii) we have $|C_i'|\leqq|C_i|$. By moving small conical elements of C_i' from outside R_i to inside R_i we see intuitively that $|C_i'|\geqq$ $\geqq|R_i'|=V(\vartheta_i)$. Formally this can be seen as follows: Let $\bar{F}_i'$ and $\bar{G}_i$ denote the projections of the sets F_i' and G_i onto the boundary of S. Then we have

$$n|C_i'| = \int_{\bar{F}_i'} \operatorname{cosec}^n \sphericalangle O_iOX\,ds = \int_{\bar{F}_i'\cap\bar{G}_i} \operatorname{cosec}^n \sphericalangle O_iOX\,ds + \int_{\bar{F}_i'-\bar{G}_i} \operatorname{cosec}^n \sphericalangle O_iOX\,ds$$

and

$$n|R_i| = \int_{\bar{G}_i} \operatorname{cosec}^n \sphericalangle O_iOX\,ds = \int_{\bar{G}_i\cap\bar{F}_i'} \operatorname{cosec}^n \sphericalangle O_iOX\,ds + \int_{\bar{G}_i-\bar{F}_i'} \operatorname{cosec}^n \sphericalangle O_iOX\,ds,$$

where ds denotes the $(n-1)$-dimensional surface-element of the unit sphere at the variable point X. We have, on the one hand,

$$\int_{\bar{F}_i'-\bar{G}_i} ds = \int_{\bar{G}_i-\bar{F}_i'} ds.$$

On the other hand, the integrand $\operatorname{cosec}^n \sphericalangle O_i O X$ is at any point of the set $\bar{F}_i' - \bar{G}_i$ greater than at any point of the set $\bar{G}_i - \bar{F}_i'$. Thus we have, indeed, $|C_i| \geqq |C_i'| \geqq \geqq |R_i| = V(\vartheta_i)$.

We continue to show that $V(\vartheta)$ is a convex function of ϑ. Consider a right spherical cone R with altitude 1 and solid angle ϑ. Let $g(\vartheta)$ be the length of the generator of R. Let R' be the right spherical cone with altitude 1 and solid angle $\vartheta + d\vartheta$ concentric with R. Then we have $|R'| = |R| + \frac{g(\vartheta)^n}{n} d\vartheta + o(g(\vartheta)^n)\, d\vartheta$, showing that

$$\frac{d}{d\vartheta} V(\vartheta) = \frac{g(\vartheta)^n}{n}. \tag{4}$$

The convexity of $V(\vartheta)$ follows immediately form the obvious fact that $g(\vartheta)$ is an increasing function of ϑ.

Now we are in the position to give a non-trivial lower bound for the volume of D_S^k. Using property (iii), the relations (3), the convexity of $V(\vartheta)$ and Jensen's inequality, we obtain

$$|D_S^k| \geqq |D_S^k \cap S(r)| = \sum_{i=0}^{l} |C_i| \vartheta_0 \frac{r^n}{n} + \sum_{i=1}^{l} V(\vartheta_i) \geqq \vartheta_0 \frac{r^n}{n} + lV\left(\frac{\omega_n - \vartheta_0}{l}\right). \tag{5}$$

We claim that $l \leqq w(m_k(2r+1), k)$. As an immediate consequence of the definitions we have $l \leqq v(m_k(2r+1))$. It remains to prove that $l \leqq k^2 \binom{m_k(2r+1)}{k}$. Observe that $\bar{D}_S^k$ can be constructed also in the following way: Consider any subset $\mathscr{A}$ of $\bar{\mathscr{S}}$. We associate with $\mathscr{A}$ the set $D_{\mathscr{A}}$ of those points of the space whose distance from the centre of any sphere of $\mathscr{A}$ is less than or equal to their distance from the centre of any sphere of $\bar{\mathscr{S}} - \mathscr{A}$. Let $\|\mathscr{A}\|$ denote the cardinality of $\mathscr{A}$. The definitions of D_S^k and $D_{\mathscr{A}}$ imply that

$$D_S^k = \bigcup_{\mathscr{A} \ni S,\, \|\mathscr{A}\| = k} D_{\mathscr{A}}.$$

We are going to show that $D_{\mathscr{A}}$ is a convex polyhedral region with at most $\|\mathscr{A}\| \left(m_k(2r+1) - \|\mathscr{A}\|\right)$ faces. Let us draw to each pair of spheres $S_1 \in \mathscr{A}$ and $S_2 \in \bar{\mathscr{S}} - \mathscr{A}$ the half-space consisting of those points which lie nearer to the centre of S_1 than to the centre of S_2. Then the region $D_{\mathscr{A}}$ is the intersection of all these half-spaces. The number of these half-spaces is $\|\mathscr{A}\| \, \|\bar{\mathscr{S}} - \mathscr{A}\|$ which is at most $\|\mathscr{A}\| \left(m_k(2r+1) - \|\mathscr{A}\|\right)$. It is easily seen that the sets $D_{\mathscr{A}}$ and $D_{\mathscr{A}'}$ with $\|\mathscr{A}\| = \|\mathscr{A}'\|$, $\mathscr{A} \neq \mathscr{A}'$ are disjoint. Thus $\bar{D}_S^k$ can be decomposed into at most $\binom{m_k(2r+1)}{k-1}$ convex polyhedral regions each of which has at most $k\left(m_k(2r+1) - k\right)$ faces. It follows immediately that the boundary of $\bar{D}_S^k$ can be decomposed into at most $k^2 \binom{m_k(2r+1)}{k}$ convex sets. Thus we have indeed $l \leqq w(m_k(2r+1), k)$.

Since $V(\vartheta)$ is a strictly convex function, the quantity on the right hand side of the inequality (5) decreases if we replace l by a greater value. Thus

$$|D_S^k| \geqq \vartheta_0 \frac{r^n}{n} + w(m_k(2r+1), k) V\left(\frac{\omega_n - \vartheta_0}{w(m_k(r2+1), k)}\right),$$

whence

$$|D_S^k| \geqq \min_{0 \leqq \vartheta_0 \leqq \omega_n} \left\{ \vartheta_0 \frac{r^n}{n} + w(m_k(2r+1), k) V\left(\frac{\omega_n - \vartheta_0}{w(m_k(2r+1), k)}\right)\right\}.$$

Since this last bound is valid for all values of r greater than 1, we have

$$|D_S^k| \geqq \max_{\substack{r>1 \\ l=w(m_k(2r+1),k)}} \min_{0 \leqq \vartheta_0 \leqq \omega_n} \left\{ \vartheta_0 \frac{r^n}{n} + lV\left(\frac{\omega_n - \vartheta_0}{l}\right)\right\}.$$

Now Theorem 1 follows immediately by property (i).

The proof of Theorem 2 goes analogously. We apply an idea of R. P. BAMBAH and H. DAVENPORT [1] and P. ERDŐS and C. A. ROGERS [3] using the k'th Dirichlet cells instead of the ordinary Dirichlet cells.

Let now $\mathcal{S}$ be a system of unit spheres forming a k-fold covering of the space. Without loss of generality we assume, on the one hand, that the spheres of $\mathcal{S}$ are all different, on the other hand, that there is no finite subset of $\mathcal{S}$ which can be replaced by a smaller number of spheres so that the spheres continue to form a k-fold covering. We shall show that under these assuptions the density of each sphere $S \in \mathcal{S}$ with respect to the k'th Dirichlet cell D_S^k is at least

$$\int_0^1 \frac{dt}{[1-C^\delta(1-t^\delta)]^{n/2}},$$

where

$$\delta = \frac{2}{n-1} \quad \text{and} \quad C = \frac{nJ_n}{2w(M_k(3), k) J_{n-1}}.$$

The proof rests on a result of C. A. ROGERS [13] which he phrases only for convex polytopes, but his proof immediately implies the following

LEMMA. *Let P be a polytope contained in a unit sphere S with centre O. Suppose that P can be decomposed into l convex piramids with the common apex O. Then*

$$\frac{|S|}{|P|} \geqq \int_0^1 \frac{dt}{[1-\bar{C}^\delta(1-t^\delta)]^{n/2}},$$

where $\delta = \dfrac{2}{n-1}$ *and* $\bar{C} = \dfrac{n|P|}{J_{n-1}}$.

The number of faces of D_S^k cannot be greater than the number of spheres of $\mathcal{S}$ lying in $S(3)$. There are at most $M_r(3)$ spheres of $\mathcal{S}$ lying in $S(3)$, since otherwise we could replace the spheres of $\mathcal{S}$ contained in $S(3)$ by a smaller number of spheres so that the spheres continue to form a k-fold covering. Thus D_S^k has at most $M_k(3)$ faces. Hence we can see, in the same way as in the case of a k-fold packing, that D_S^k can be decomposed into $w(M_k(3), k)$ convex piramides with the common apex Q. Using the lemma we obtain

$$\frac{|S|}{|D_S^k|} \geqq \int_0^1 \frac{dt}{[1-\bar{C}^\delta(1-t^\delta)]^{n/2}},$$

where

$$\delta = \frac{2}{n-1} \quad \text{and} \quad \bar{C} = \frac{n|D_S^k|}{w(M_k(3), k) J_{n-1}}.$$

If $|S|/|D_S^k| < 2$ then $|\bar{C}| > |C|$ and we have

$$\frac{|S|}{|D_S^k|} \geqq \int_0^1 \frac{dt}{[1-\bar{C}^\delta(1-t^\delta)]^{n/2}} \geqq \int_0^1 \frac{dt}{[1-C^\delta(1-t^\delta)]^{n/2}}$$

as claimed. Consider now the case when $|S|/|D_S^k| \geqq 2$. Using the obvious inequality $w(M_k(3), k) \geqq M_k(3) \geqq 3^n$ it can be easily checked that $C^\delta < 2^{1-n}$, i.e. $C^\delta < 1/4$. Hence

$$\int_0^1 \frac{dt}{[1-C^\delta(1-t^\delta)]^{n/2}} < \int_0^1 \left(\frac{4}{3+t^\delta}\right)^{1/\delta} dt.$$

But it is not difficult to show that $\left(\frac{4}{3+t^\delta}\right)^{1/\delta} \leqq t^{-1/2}$ for $0<\delta$, $0 \leqq t \leqq 1$. Thus we have again

$$\frac{|S|}{|D_S^k|} \geqq 2 = \int_0^1 t^{-1/2} dt \geqq \int_0^1 \left(\frac{4}{3+t^\delta}\right)^{1/\delta} dt \geqq \int_0^1 \frac{dt}{[1-C^\delta(1-t^\delta)]^{n/2}}.$$

Referring to property (i) the proof of Theorem 2 is finished.

It seems to be extremely difficult to determine the exact values of the bounds (1) and (2) even for small values of k and n. We shall try to get some information about the asymptotic behaviour of our bounds for great values of k and n. We start to give estimates for the quantities $m_k(r)$, $M_k(r)$ and $v(f)$.

Obviously, we have

(6) $$m_k(r) \leqq kr^n.$$

For $1 \leqq r < 4$ we give an other upper bound for $m_k(r)$. Let S, as above, be a unit sphere with centre O and $S(r)$ the concentric sphere of radius r. Let $\mathcal{S}$ be a system of unit spheres lying in $S(r)$ and forming, together with S, a k-fold packing. There are at most $k-1$ spheres of $\mathcal{S}$ which contain O. On the other hand, it is easy to show that for $1 \leqq r < 4$ any unit sphere which lies in $S(r)$ and does not contain O intersects the boundary of $S(\sqrt{r})$ in a spherical cap whose angular radius is at least $\arcsin \frac{\sqrt{4-r}}{2}$. Obviously, the spherical caps cut out by the spheres of $\mathcal{S}$ from the boundary of $S(\sqrt{r})$ form a k-fold packing. Thus, if we denote by $N_k(\alpha)$ the maximum number of spherical caps with angular radius α forming a k-fold packing, then

(7) $$m_k(r) \leqq k-1+N_k\left(\arcsin \frac{\sqrt{4-r}}{2}\right), \quad 1 \leqq r < 4.$$

For $N_k(\alpha)$ we have the trivial bound

(8) $$N_k(\alpha) \leqq kC(\pi)/C(\alpha),$$

where $C(\alpha)$ denotes the $((n-1)$-dimensional$)$ volume of a cap of radius α. Also we have

$$(9)\qquad \frac{C(\pi)}{C(\alpha)}=\frac{n\pi^{1/2}\Gamma(n/2)\cos}{\Gamma\left(\frac{n+1}{2}\right)\sin^n}\left\{1-\frac{\tan^2\alpha}{n+2}+\frac{3\vartheta\tan^4\alpha}{(n+2)(n+4)}\right\}^{-1},$$

where $0\leqq\vartheta\leqq 1$ (see [11], Lemma 4). It follows immediately that

$$\lim_{n\to\infty}[N_k(\alpha)]^{1/n}\leqq\operatorname{cosec}\alpha.$$

Hence by (7)

$$\lim_{n\to\infty}[m_k(3)]^{1/n}\leqq 2,$$

showing that the bound (7) is, for great values of n in a neighbourhood of $r=3$, better than the trivial bound (6). Numerical computations suggest that the same is true for any $n>1$. For $(k+1)^2<n/(2\log n)$ a non-trivial upper bound for $N_k(\alpha)$ was given by N. M. BLACHMAN and L. FEW [2]. Their result implies that

$$\lim_{n\to\infty}[m_k(3)]^{1/n}\leqq 2\sqrt{\frac{k}{k+1}}.$$

We continue to show that

$$(10)\qquad m_k(r)>\left[\frac{k(r-1)^2}{2(k+1)}\right]^{n/2}-1.$$

Let $\bar{\mathscr{S}}$ be a system of spheres of radii $\frac{k+1}{2k}$ with the following properties: a) for any sphere $\bar{S}\in\bar{\mathscr{S}}$ we have $\bar{S}\subset S\left(r-1+\sqrt{\frac{k+1}{2k}}\right)$, b) $\left\{S\left(\sqrt{\frac{k+1}{2k}}\right)\right\}\cup\bar{\mathscr{S}}$ is a packing and c) any sphere of radius $\frac{k+1}{2k}$ which is contained in $S\left(r-1+\sqrt{\frac{k+1}{2k}}\right)$ intersects one of the spheres of $\left\{S\left(\sqrt{\frac{k+1}{2k}}\right)\right\}\cup\bar{\mathscr{S}}$. Let $\mathscr{S}$ be the system of unit spheres concentric with the spheres of $\bar{\mathscr{S}}$. Then, the spheres of $\mathscr{S}$ are contained in $S(r)$ and it follows from a result of L. FEW [6] that $\{S\}\cup\mathscr{S}$ is a k-fold packing. On the other hand, it is easily seen that the spheres of radii $2\sqrt{\frac{k+1}{2k}}$ concentric with the spheres of $\bar{\mathscr{S}}$, together with the sphere $S\left(\sqrt{\frac{2(k+1)}{k}}\right)$ cover the sphere $S(r-1)$. Thus we have

$$\|\mathscr{S}\|+1\geqq\frac{(r-1)^n}{\left[\sqrt{\frac{2(k+1)}{k}}\right]^n}$$

in accordance with (10).

C. A. ROGERS [14] proved that there is an absolut constant c such that any sphere of radius r can be covered by less than $cn^{5/2}r^n$ unit spheres. It follows immediately that

$$M_k(r) \leqq kcn^{5/2}r^n. \tag{11}$$

On the other hand, we have the trivial bound

$$kr^n \leqq M_k(r). \tag{12}$$

A better upper bound for $M_k(r)$ can be obtained by means of recent results of L. FEW [9].

Let $f(n, l)$ be the maximal number of cells determined by l hyperplanes in Euclidean n-space. It is easily seen that $f(n, l)$ is defined by the following recursive formula:

$$f(1, l) = 1+l \quad \text{for} \quad l = 0, 1, \dots, \quad f(n, l) = 1+\sum_{j=0}^{l-1} f(n-1, j).$$

Using the wellknown fact that

$$\sum_{j=0}^{l} j^n \leqq \frac{(l+1)^{n+1}}{n+1}$$

it easily follows that

$$f(n, l) < \frac{(l+1)^{n+1}}{(n+1)!}.$$

If Π is a polyhedral region whose faces lie in at most f hyperplanes then each face can be divided into at most $\frac{f^{n+1}}{(n+1)!}$ convex sets. Thus we have

$$v(f) \leqq \frac{f^{n+2}}{(n+1)!}. \tag{13}$$

Now consider the quantity

$$\vartheta_0 \frac{r^n}{n} + lV\left(\frac{\omega_n - \vartheta_0}{l}\right)$$

occuring in (1). We have by (4)

$$\frac{d}{d\vartheta_0}\left[\vartheta_0 \frac{r^n}{n} + lV\left(\frac{\omega_n - \vartheta_0}{l}\right)\right] = \frac{r^n}{n} - \frac{g\left(\frac{\omega_n - \vartheta_0}{l}\right)^n}{n}.$$

It follows that

$$\vartheta_0 \frac{r^n}{n} + lV\left(\frac{\omega_n - \vartheta_0}{l}\right)$$

is a convex function of ϑ_0 which attains its minimum at the value $\bar{\vartheta}_0$ for which $g\left(\frac{\omega_n - \bar{\vartheta}_0}{l}\right) = r$, or if $g\left(\frac{\omega_n}{l}\right) < r$ at $\vartheta_0 = 0$.

It is obvious from the relations (9) and (10) that there is a minimal constant $\bar{r}=\bar{r}(k,n)$ such that

$$g\left(\frac{\omega_n}{w\left(m_k(2\bar{r}+1),k\right)}\right)\leqq \bar{r},$$

and we have $\lim_{k,n\to\infty}\bar{r}(k,n)=1$. Numerical computations show that $\bar{r}<3/2$ for any $k>0$ and $n>1$. Thus we have

$$\delta_k^n\leqq kJ_n/\left[m_k(4)V\left(\frac{\omega_n}{w\left(m_k(4)\right)}\right)\right].$$

Let $d(\vartheta)$ denote the density of a unit sphere centered at the apex of a right spherical cone of altitude 1 and solid angle ϑ with respect to the cone. Then the quantity on the right hand side of the last inequality is nothing else but $kd\left(\frac{\omega_n}{w\left(m_k(4)\right)}\right)$. Thus we have

(14) $$\delta_k^n\leqq kd\left(\frac{\omega_n}{w\left(m_k(4)\right)}\right).$$

It easily follows from (9) that for great values of n

$$[d(\vartheta)]^{1/n}\approx[g(\vartheta)]^{-1}.$$

Comparing this with (9) and (14) we see that for any fixed k, δ_k^n tends exponentially to zero when n tends to infinity. However our bound tends to zero at a slower rate than the bound of L. FEW [6]. On the other hand, it can be expected that for fixed n even the difference of k and our bound tensd to zero as k tends to infinity. Thus (1) seems to give a very weak upper bound for δ_k^n if k is large compared with n.

Consider now the quantity

$$\int_0^1\frac{dt}{[1-C^\delta(1-t^\delta)]^{n/2}}$$

occuring in (2), where

$$\delta=\frac{2}{n-1}\quad\text{and}\quad C=\frac{nJ_n}{2w\left(M_k(3),k\right)J_{n-1}}.$$

It follows by (11) and (12) that for fixed k, C^δ converges to a positive constant λ_k as n tends to infinity. Consequently

$$\int_0^1\frac{dt}{[1-C^\delta(1-t^\delta)]^{n/2}}\to\int_0^1 t^{-\lambda_k^{-2}}\,dt=\frac{1}{1-\lambda_k^2}.$$

Thus our lower bound for Δ_k^n is hardly better than the trivial bound.

In 3-space we obtain better bounds than (1) and (2) by a slight modification of the original proofs. Let $\bar{V}(\vartheta)$ denote the volume of a right spherical cone with generator 1 and solid angle ϑ. We shall skech the proofs of the following theorems:

THEOREM 3. *We have*

$$\delta_k^3 \leqq \frac{4k\pi}{3}\left[\max_{r>0}\min_{0\leqq\vartheta_0\leqq 4\pi}\left\{\vartheta_0\frac{r^3}{3}+m_k(2r+1)V\left(\frac{4\pi-\vartheta_0}{m_k(2r+1)}\right)\right\}\right]^{-1}.$$

THEOREM 4. *We have*

$$\Delta_k^3 \geqq \frac{4k\pi}{3} M_k(3)V\left(\frac{4}{M_k(3)}\right).$$

Let $\mathcal{S}$ be a system of different unit spheres. For $S\in\mathcal{S}$ let l be the number of the face-planes of D_S^k which intersect $S(r)$ $(r>1)$. We divide $S(r)\cap D_S^k$ into certain parts $C_0, \dots, C_l$ defined as follows: The part C_0 is the point-set-union of the segments joining the centre O of S with the boundary-points of $S(r)$ lying in D_S^k. The part C_i $(0<i\leqq l)$ is the point-set-union of the segments joining O with the boundary-points of D_S^k lying in $S(r)$ and one particular face-plane of D_S^k. Let x_i be the altitude of the cone C_i $(i=1, \dots, l)$ and ϑ_i the solid angle of C_i at O $(i=0, \dots, l)$. Let $V(\vartheta, x)$ and $\overline{V}(\vartheta, x)$ be the volume of a right cone of altitude x and solid angle ϑ based on a circular ring such that the inner and outer generator is equal to 1, respectively. An argument analogous to that used in the proof of (3) shows that

$$|C_0| = \frac{\vartheta_0 r^3}{3}, \quad |C_i| \geqq V(\vartheta_i, x_i) \quad \text{for} \quad x_i < 1$$

and

$$|C_i| \geqq V(\vartheta_i) \quad \text{for} \quad x_i \geqq 1 \quad (i = 1, \dots, l)$$

when $\mathcal{S}$ is a k-fold packing and

$$|C_0| = 0, \quad |C_i| \leqq \overline{V}(\vartheta_i, x_i) \quad (i = 1, \dots, l)$$

when $\mathcal{S}$ is a k-fold covering.

We obtain by an elementary computation that

$$V(\vartheta, x) = \frac{\pi x\vartheta(4\pi x-\vartheta)}{3(2\pi x-\vartheta)^2} \quad \left(0 \leqq \frac{\vartheta}{2\pi} \leqq x \leqq 1\right), \quad \frac{\partial V}{\partial x} = \frac{\pi\vartheta^2(\vartheta-6\pi x)}{3(2\pi x-\vartheta)^3} \leqq 0,$$

$$\overline{V}(\vartheta, x) = \frac{\pi\vartheta x(4\pi x+\vartheta)}{2\pi x+\vartheta} \quad \left(0 \leqq \frac{\vartheta}{2\pi} \leqq x-1 \leqq 1\right), \quad \frac{\partial \overline{V}}{\partial x} = \frac{\pi\vartheta^2(12\pi+\vartheta)}{3(2\pi+\vartheta)^3} \geqq 0.$$

Thus we have $V(\vartheta, x) \geqq V(\vartheta, 1) = V(\vartheta)$ and $\overline{V}(\vartheta, x) \leqq \overline{V}(\vartheta, 1) = \overline{V}(\vartheta)$. Hence

$$|D_S^k| \geqq |D_S^k \cap S(r)| = \sum_{i=1}^{l} |C_i| \geqq \frac{\vartheta_0 r^3}{3} + \sum_{i=1} V(\vartheta_i)$$

if $\mathcal{S}$ is a k-fold packing and

$$|D_S^k| = \sum_{i=1} |C_i| \geqq \sum_{i=1}^{l} \overline{V}(\vartheta_i)$$

if $\mathcal{S}$ is a k-fold covering.

We have already seen that $V(\vartheta)$ is convex. On the other hand, it is easy to show that

$$\bar{V}(\vartheta) = \frac{\pi\vartheta(4\pi+\vartheta)}{2\pi+\vartheta}$$

is concave. Using Jensen's inequality we obtain

$$|D_S^k| \geqq \frac{\vartheta_0 r^3}{3} + lV\left(\frac{4\pi-\vartheta_0}{l}\right)$$

for the case of a k-fold packing and

$$|D_S^k| \leqq l\bar{V}\left(\frac{4\pi}{l}\right)$$

for the case of a k-fold covering.

Now we have $l \leqq m_k(2r+1)$ if the spheres form a k-fold packing and $l \leqq M_k(3)$ if the spheres constitute a k-fold covering such that there is no finite subset of $\mathscr{S}$ which can be replaced by a smaller number of spheres so that the spheres continue to form a k-fold covering. Thus we have

$$|D_S^k| \geqq \max_{r>1} \min_{4 \leqq \vartheta_0 \leqq 4\pi} \left\{\frac{\vartheta_0 r^3}{3} + m_k(2r+1)V\frac{4\pi-\vartheta_0}{m_k(2r+1)}\right\}$$

and

$$|D_S^k| \leqq M_k(3)\bar{V}\left(\frac{4\pi}{M_k(3)}\right),$$

respectively.

Theorems 3 and 4 follow now by property (i).

Theorem 3 gives a rather good upper bound for δ_2^3. From (7) and (8) we obtain $m_2(2\cdot 1.057+1)=m_2(3.114)\leqq 34$. The quantity

$$\frac{\vartheta_0 1.057^3}{3} + 34\, V\left(\frac{4\pi-\vartheta_0}{34}\right)$$

attains its minimum 4.5864... at $\vartheta_0=1.0462\ldots$. Thus we have

$$\delta_2^3 \leqq \frac{8\pi}{3\cdot 4.5864\ldots} = 1.826\ldots.$$

The best known lower bound for δ_2^3 is [10]

$$\delta_2^3 \geqq 1.612\ldots$$

References

[1] R. P. Bambah and H. Davenport, The covering of n-dimensional space by spheres, *J. London Math. Soc.*, **27** (1952), 224—229.
[2] N. M. Blachman and L. Few, Multiple packing of spherical caps, *Mathematika*, **10** (1963), 84—88.

[3] P. ERDŐS and C. A. ROGERS, The covering of n-dimensional space by spheres, *J. London Math. Soc.,* **28** (1953), 287—293.
[4] G. FEJES TÓTH, Multiple packing and covering of the plane with circles, *Acta Math. Acad. Sci. Hungar.,* **27** (1976), 135—140.
[5] L. FEW, The double packing of spheres, *J. London Math. Soc.,* **28** (1953), 297—304.
[6] L. FEW, Multiple packing of spheres, *J. London Math. Soc.,* **39** (1964), 51—54.
[7] L. FEW, Multiple packing of spheres: a survey, *Proc. Coll. Convexity* (Copenhagen, 1965), 88—93.
[8] L. FEW, Double packing of spheres: A new upper bound, *Mathematika,* **15** (1968), 88—92.
[9] L. FEW, Multiple covering by convex bodies, lecture held on the *Colloquium on Discret Geometry* in Salzburg, 1975.
[10] L. FEW and P. KANAGASABATHY, The double packing of spheres, *J. London Math. Soc.,* **44** (1969), 141—146.
[11] A. FLORIAN, Mehrfache Unterdeckungen und Überdeckungen von konvexen Körper.
[12] A. RANKIN, The closest packing of spherical caps in n dimensions, *Proc. Glasgow Math. Ass.,* **2** (1955), 139—144.
[13] C. A. ROGERS, The volume of a polyhedron insribed in a sphere, *J. London Math. Soc.,* **28** (1953), 410—416.
[14] C. A. ROGERS, Covering a sphere with spheres, *Mathematika,* **10** (1963), 157—164.

(Received October 31, 1977)

MATHEMATICAL INSTITUTE
OF THE HUNGARIAN ACADEMY OF SCIENCES
1053 BUDAPEST, REÁLTANODA U. 13—15.

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 177—183.

ISOMORPHISMS OF CAYLEY GRAPHS. II

By

L. BABAI and P. FRANKL (Budapest)

1. Introduction. By groups we mean finite ones throughout this paper. For G a group, let G_L denote the left regular representation (group of left translations) of G. A *Cayley graph* of G is a graph X on the vertex set G such that its automorphism group Aut X contains G_L. Denoting by R the set of neighbours of the unity $e \in G$ in X, we have $R = R^{-1}$, $e \notin R$. X is uniquely determined by R; a and b are adjacent in X iff $a^{-1}b \in R$. Set therefore $X = X(G; R)$. If α is an automorphism of the *group* G, then $X^\alpha = X(G; R^\alpha)$ is another Cayley-graph of G, isomorphic to X.

X is called a *CI-object* of G, if all Cayley graphs of G isomorphic to X, are of the form X^α ($\alpha \in \mathrm{Aut}\, G$). G is a $\mathcal{G}-CI$ group, if all Cayley graphs of G are *CI*-objects. ($\mathcal{G}$ stands for graphs; the analogous concepts can be introduced for other classes of structures — digraphs, relational systems, etc. —, too.) We are concerned with the following problem:

Which are the $\mathcal{G}-CI$ groups?

This problem has been raised in [3], generalizing a question of A. ÁDÁM [1].

Let Z_n denote the cyclic group of order n, and $G^m = G \times \ldots \times G$ (direct power). The following groups are known to be $\mathcal{G}-CI$ (p, q are primes, $p \neq q$); Z_p (DJOKOVIĆ [7]; TURNER [18]), the groups of order $2p$ (BABAI [3]), and $3p$ [4], Z_{4p} (GODSIL [8]), Z_{pq} (GODSIL [8], KLIN—KALOUJNINE—PÖSCHEL, see [11, 12, 13, 15], ALSPACH—PARIONS [2], Z_p^2 ([4, 8]), and the only known result involving an arbitrarily large number of prime factors: Z_n is $\mathcal{G}-CI$ (in fact, $\mathcal{C}-CI$ for any class $\mathcal{C}$ of structures), provided $n = p_1 \ldots p_r$ where $p_{i+1} > p_1 \ldots p_i$ ($i = 1, \ldots, r-1$), and *g.c.d.* $(n, \varphi(n)) = 1$(PÁLFY [14]).

On the other hand, constraints on the structure of $\mathcal{G}-CI$ groups were obtained by the present authors [5], principally for the case of odd $|G|$ (cf. Section 2). The main objective of this note is to prove

THEOREM 1.1. *Let G be a non-solvable $\mathcal{G}-CI$ group. Then $G = L \times N$, where*

(i) *L is isomorphic to one of $PSL(2, 5)$, $SL(2, 5)$ $PSL(2, 13)$, and $SL(2, 13)$;*

(ii) *N is a direct product of elementary abelian groups*;

(iii) *g.c.d.* $(|N|, |L|) = 1$.

The proof depends on deep results of group theory, in particular on GORENSTEIN—WALTER [9] and BRAUER—SUZUKI [6].

In a later paper, we shall return to the problem of restricting the structure of solvable $\mathcal{G}-CI$ groups of even order. We should stress here, that we do not know if there are any nonsolvable $\mathcal{G}-CI$ groups. (This note might eventually serve as a step towards a proof of their nonexistence.) For $\mathcal{G}-CI$ groups of odd order, the following holds:

THEOREM 1.2 [5]. *Let G be a $\mathcal{G}-CI$ group of odd order, and G_p its Sylow p-subgroup. Then*

(i) *G_p is either elementary abelian or Z_9 or Z_{27};*

(ii) *G is either abelian or it has an abelian normal subgroup M of index 3 and an element b of order 3^l in $G-M$ and there exists an r such that $b^{-1}xb=x^r$ for each $x\in M$ ($r^3\equiv 1 \bmod |G|$).*

In particular, every subgroup of M is normal in G, hence G is supersolvable. Moreover, G is abelian, in fact, a product of elementary abelian groups, provided $|G|$ is prime to 6.

2. Preliminaries

θ will denote *the class of factor groups of $\mathcal{G}-CI$ groups* throughout. We shall require a few lemmas from [5].

LEMMA 2.1. *θ is closed under subgroups and factor groups.* [5, 3.2.]

Note, that

LEMMA 2.2. *Subgroups of $\mathcal{G}-CI$ groups are $\mathcal{G}-CI$. If N is a characteristic subgroup of the $\mathcal{G}-CI$ group G, then G/N is $\mathcal{G}-CI$ again.* [5, 3.2, 3.5.]

We do not know if all factor-groups of a $\mathcal{G}-CI$ group G are $\mathcal{G}-CI$. (This holds if G has odd order and also if G is not solvable, since by 1.1 and 1.2, any factor-group of G can in these cases be obtained by taking a factor-group by a characteristic subgroup, of a subgroup of G.)

LEMMA 2.3. *If a p-group P belongs to θ, then P is either an elementary abelian group, or a cyclic group of order p^l, $l\leqq 3$, $p\leqq 3$, or the quaternion group Q of order* 8. [5, 5.1.]

LEMMA 2.4. *If $G\in\theta$, $H_1, H_2\leqq G$ and $|H_1|=|H_2|$ then $H_1^\alpha=H_2$ for some* $\alpha\in \operatorname{Aut} G$. [5, 5.1.]

LEMMA 2.5. *Let m, n, k denote positive integers, g.c.d. $(m, k)=1$, and assume that n is the order of k* mod m. *If the metacyclic group*

$$G = \langle a, b | a^m = b^n = e,\ b^{-1}ab = a^k\rangle$$

belongs to θ, then $n\leqq 6$, $n\neq 5$. [5, 4.5.]

LEMMA 2.6. *Let $G\in\theta$ be a nonabelian group, P its Sylow p-subgroup. Assume that $|P|=p^k$, P is a minimal normal subgroup of G, and G has a cyclic Hall p'-subgroup $\langle a\rangle$. Then either $k=1$ and $|a|\in\{2, 3, 4, 6\}$, or $k=2$ and either $p=2$, $|a|=3$, or $p=3$, $|a|\in\{4, 8\}$.* [5, 4.6.]

Several times we shall refer to the following result from group theory:

LEMMA 2.7 (BURNSIDE, see [10, p. 419].) *Let P be a Sylow p-subgroup of G satisfying $P\leqq Z(N_G(P))$ (i.e. $C_G(P)=N_G(P)$). Then P has a normal complement in G.*

(As usual, $N_G(P)$ and $C_G(P)$ denote the normalizer and the centralizer of P in G, $Z(G)$ is the centre of G, and a normal complement of P is a normal subgroup $N \lhd G$ such that $NP=G$, $|N \cap P|=1$.)

Our standard reference book on group theory is HUPPERT [10]. All the results used here are proved or at least quoted there.

3. The proof

LEMMA 3.1. *Let P denote the Sylow 2-subgroup of $G \in \theta$. Then either*

(a) P has a normal complement in G; or

(b) P is isomorphic to the quaternion group Q of order 8; or

(c) $P \cong Z_2 \times Z_2$ and $N_G(P)/C_G(P) \cong Z_3$.

PROOF. If P is cyclic then (a) holds by [10, p. 32, Aufgabe 21]. If $C_G(P)=N_G(P)$ then again (a) holds by a theorem of BURNSIDE (Lemma 2.7). In the remaining cases, $P \cong Z_k^2$ for some $k \geqq 2$ (by Lemmas 2.1, 2.3) and $C_G(P) \neq N_G(P)$. By the mentioned theorem of BURNSIDE (2.7), P has a normal complement in $C_G(P)$, say M. Clearly, $C_G(P)=P \times M$ and both M and P are characteristic subgroups of $N_G(P)$. Let $\varphi\colon N_G(P) \to N_G(P)/M$ denote the natural epimorphism. Set $\varphi x=\bar{x}$, $\varphi(H)=\bar{H}$ for $x \in N_G(P)$, $H \leqq N_G(P)$. Let u be an arbitrary member of $N_G(P)-C_G(P)$. Consider the group $\langle \bar{u}, \bar{P} \rangle = L$. Clearly, $P \cong \bar{P} \lhd L$, and $C_L(\bar{P})=\bar{P}$. $\bar{P}$ contains a minimal normal subgroup P_1 of L. Let $|P_1|=2^l$. By Lemma 2.1, $L \in \theta$. This implies that each subgroup of P having order 2^l is a minimal normal subgroup of L by Lemma 2.4. Hence either $l=1$ or $l=k$. The first case is impossible since then $\bar{P} \leqq Z(L)$, contradicting with $C_L(\bar{P})=\bar{P}$. Hence $\bar{P}$ itself is a minimal normal subgroup. As $\bar{u}^2$ generates a Hall $2'$-subgroup of L, all conditions of Lemma 2.6 are fulfilled, and we infer that $k=2$ and $|\bar{u}^2|=3$. Hence $P=Z_2 \times Z_2$, and every member of $N_G(P)-C_G(P)$ induces a non-trivial automorphism of order 3 of P. This proves that (c) holds.

LEMMA 3.2. *If $G \in \theta$ is a non-cyclic simple group then $G \cong PSL(2, p)$ with $p=5$ or 13.*

PROOF. Let P be the Sylow 2-subgroup of G. BRAUER and SUZUKI [6, cf. 10, p. 624] have shown that a group with quaternion Sylow 2-subgroup is never simple.

Hence, by the previous lemma $P \cong Z_2 \times Z_2 = D_2$, the dihedral group of degree 2. All simple groups with dihedral Sylow 2-subgroups have been determined by GORENSTEIN and WALTER [9]. These groups are $PSL(2, q)$ ($q=p^f$, p prime $\neq 2$, $p^f>3$) and the alternating group A_7.

The Sylow 2-subgroup of A_7 has order 8, hence $G \not\cong A_7$. Assume $G \cong PSL(2, q)$. As $|G|=q(q^2-1)/2$, we have $q \equiv \pm 3 \bmod 8$.

$PSL(2, q)$ (viewed as a group of fractional linear transformations over $GF(q)$) contains a subgroup, consisting of the linear functions $x \mapsto ax+b$ ($a, b \in GF(q)$, $a \neq 0$ is a square in $GF(q)$). This group satisfies the conditions of Lemma 2.6, hence $(q-1)/2 \in \{2, 3, 4, 6, 8\}$. Of these, only $q=5$ and 13 satisfy $q \equiv \pm 3 \bmod 8$.

PROPOSITION 3.3. *Neither $PGL(2, 5)$ nor $PGL(2, 13)$ belongs to θ.*

PROOF. Both groups have dihedral Sylow 2-subgroups of order 8, excluded by Lemma 2.3.

The following is folklore:

LEMMA 3.4. Aut $(PSL(2,p))\cong$ Aut $(SL(2,p)\cong PGL(2,p)$ $(p\geqq 5)$.

PROOF. Clearly, $PGL(2,p)$ induces a group of automorphisms on $PSL(2,p)$ and $GL(2,p)$ on $SL(2,p)$. The latter action is not faithful; actually $GL(2,p)/$ $/Z(GL(2,p))=PGL(2,p)$ induces this automorphism group, too. $PSL(2,p)$ has no other automorphisms (SUZUKI [17, p. 659]). Clearly, $Z(SL(2,p))$ is invariant under Aut $(SL(2,p))$. Hence, in order to prove $|\text{Aut}(SL(2,p))|=|\text{Aut}(PSL(2,p))|$ we only have to prove that any $\alpha\in\text{Aut}(SL(2,p))$ which induces the identity on $PSL(2,p)$ is the identity itself. But such an α has the form $\alpha(A)=\beta(A)A$ where β is a homomorphism $SL(2,p)\to GF(p)^{\times}$. Having the unique non-trivial normal subgroup of $SL(2,p)$ 2 elements only, we infer that $\text{Ker}\,\beta=SL(2,p)$.

LEMMA 3.5. *Let $G\in\theta$ and $K\lhd G$. Assume that K is one of $PSL(2,p)$, $SL(2,p)$, $p\in\{5, 13\}$. Then $G=K\times N$ where N is the largest normal subgroup of odd order in G.*

PROOF. Let $\varphi(g)$: $x\mapsto x^g$ $(x\in K)$ denote the conjugation of K by $g\in G$. (φ is a homomorphism of G into Aut K.) Clearly, $\varphi(G)$ contains the inner automorphism group $K/Z(K)$. We assert that $\varphi(G)=K/Z(K)$.

For assume $\varphi(G)\not\cong K/Z(K)$. This implies $\varphi(G)\cong PGL(2,p)$ by Lemma 3.4. We infer by Proposition 3.3 that $\varphi(G)\notin\theta$, in contradiction with $G\in\theta$.

We conclude that $\varphi(G)=K/Z(K)=\varphi(K)$. Clearly, $\text{Ker}\,\varphi\cap K=Z(K)$. If $|Z(K)|=1$, this implies $G=K\times\text{Ker}\,\varphi$. Set $L=\text{Ker}\,\varphi$ in this case.

K has no normal 2-complement; G has still less. Consequently, by Lemma 3.1, the Sylow 2-subgroup of G is either $Z_2\times Z_2$ or Q, the quaternion group $(|Q|=8)$. Clearly, the latter is the case when $K=SL(2,p)$, and the former otherwise. In any case, $|G: K|$ is odd. Hence, if $|Z(K)|=2$, then $|\text{Ker}\,\varphi|\equiv 2 \bmod 4$, and so, Ker φ contains a characteristic subgroup of index 2. Call this subgroup N. Again, $|N|$ is odd and $N\lhd G$. Moreover, clearly, $N\cap K=\{e\}$, hence $G=K\times N$.

Obviously, N is the largest normal subgroup of odd order in G.

We shall need the following result of SCHUR (cf. [17, p. 658]):

LEMMA 3.6. (SCHUR [16, p. 120]). *Assume that the Sylow 2-subgroup of the group G is the quaternion group, and $G/F\cong PSL(2,p)$ for some normal subgroup F order 2. Then $G\cong SL(2,p)$.*

LEMMA 3.7. *Assume that the group $G\in\theta$ is not solvable. Let N be the largest normal subgroup of odd order and R the largest solvable normal subgroup of G. Then $G/R\cong PSL(2,p)$ where $p\in\{5, 13\}$, and either*

(i) $|R: N|=2$, $G/N\cong SL(2,p)$; *or*

(ii) $R=N$ *and* $G/N\cong PSL(2,p)$.

PROOF. Let H be a non-cyclic simple group appearing as a composition factor of G. By Lemma 3.2, $H\cong PSL(2,p)$ where $p=5$ or 13. Let M be a maximal normal subgroup such that G/M still has H as a composition factor. Let T be a minimal normal subgroup of G/M. So, T is characteristically simple, whence $T\cong H\times\ldots\times H$. Now $T\in\theta$ and $H\lhd T$, and by Lemma 3.5 this implies that $|T/H|$

is odd, $T \cong H$. Thus $H \lhd G/M$. Again by 3.5, $G/M = H \times L$, hence $H \cong G/R$ for some normal subgroup R of G.

Let P be the Sylow 2-subgroup of G. P has no normal complement in G (since G is not solvable). By Lemma 3.1, either (b) $P \cong Q$, or (c) $P \cong Z_2 \times Z_2$ holds. In the first case $|R| \equiv 2 \bmod 4$; let N denote the characteristic subgroup of index 2 in R. In the second case set $N = R$. Clearly, in both cases, N is the largest odd normal subgroup of G and R is the largest solvable one. In the second case (ii) holds. In the first case we have to apply Lemma 3.6 to obtain (i).

COROLLARY 3.8. *Using the assumptions and notation of Lemma* 3.7, *let* $K = G/N$. *Let* D *be a non-solvable subgroup of* G. *Then* $D/D_1 \cong K$, *where* D_1 *is the largest odd normal subgroup of* D.

PROOF. $G/R = H \cong PSL(2, p)$. By Lemma 3.7, this is the only non-cyclic composition factor of G, hence it is one of the composition factors of D. If $K = H$ we are done. If $|K|/|H| = 2$ we use that the Sylow 2-subgroup of G is the quaternion group. Therefore D cannot have $Z_2 \times Z_2$ as its subgroup. So, applying Lemma 3.7 to D, (i) will hold.

Henceforth G, R, N are defined as in Lemma 3.7, and p denotes the prime for which $G/R \cong PSL(2, p)$. By Theorem 1.2, (and Lemma 2.1 which will not be mentioned henceforth), N contains a characteristic Hall $3'$-subgroup W which is the direct product of elementary abelian groups. Let T denote the Sylow 3-subgroup of G. T is either cyclic of order $\leqq 27$, or elementary abelian (Lemma 2.3). Let $|T| = 3^t$. (We have $t \geqq 1$).

PROPOSITION 3.9. *Let* r *be an integer and* D *be any of* R, N, W. *If* G *has an element of order* r *outside* D, *then all elements of order* r *are outside* D.

PROOF. Clear by Lemma 2.4, since D is a characteristic subgroup of G.

COROLLARY 3.10. *g.c.d.* $(|W|, |G/W|) = 1$ *and* $|N: W| = 3^{t-1}$.

PROOF. Let r be a prime, dividing $|W|$. Note that $r \geqq 5$. By 3.9, W contains all elements of order r. By Lemma 2.3, the Sylow r-subgroup of G is elementary abelian, hence it is contained in W.

As 3 divides $|K|$ but 9 does not $|T: T \cap N| = 3$, and $|N: W| = 3^{t-1}$ follows.

COROLLARY 3.11. *T is cyclic.*

PROOF. If $|T| \geqq 9$ then $|N \cap T| \geqq 3$. It follows by 3.9 that N contains all elements of order 3. But $N \not\geqq T$ (since $3 \,|\, |K|$). We infer that if $|T| \geqq 9$ then T is not elementary abelian.

LEMMA 3.12. $|G: C_G(W)| = 3^s$ *where* $0 \leqq s \leqq t-1$. (*Note that* $1 \leqq t \leqq 3$.)

PROOF. Let z be an element of order p in G. $z \notin N$ by 3.10. $\langle z, W \rangle$ is a group in θ of order $p|W|$ which is prime to 6, whence this group is abelian (Theorem 1.2). This yields $z \in C_G(W)$. Let $F_0 = \langle z : |z| = p \rangle$. F_0 is a characteristic subgroup of G. It is not solvable (since $z \notin R$). This implies, by 3.8, that $|K| \,\big|\, |F_0/F_1|$ where F_1 denotes the largest odd normal subgroup of F_0. As clearly $F_0 \cap W \leqq F_1$ and $F_0 W \leqq C_G(W)$, we obtain $|K|\,|W| \,\big|\, |C_G(W)|$. But $|K|\,|W| = |G|/3^{t-1}$.

LEMMA 3.13. *$C_G(W)=W\times L$ where L is a nonsolvable characteristic subgroup of G.*

PROOF. $C_G(W)$ is nonsolvable since it is not contained in R (3.12). W being in the centre of $C_G(W)$ we obtain by 3.10 and Burnside's theorem (Lemma 2.7), that each Sylow subgroup of W — and hence W itself — has a normal complement. Denote this latter one by L. Clearly, $C_G(W)=W\times L$. As L is a normal Hall-subgroup of $C_G(W)$, it is characteristic in $C_G(W)$, hence in G (since W char G).

LEMMA 3.14. *Let L be as in* 3.13, *and let M denote the largest odd normal subgroup of L. Then $L/M\cong K$ and $|M|=3^{t-1-s}$.*

PROOF. $L/M\cong K$ follows from 3.8. As

$$|G| = 3^s|C_G(W)| = 3^s|W|\,|L| = 3^s|W|\,|K|\,|M|$$

and

$$|G| = |N|\,|K| = 3^{t-1}|W|\,|K|,$$

we infer $|M|=3^{t-1-s}$.

LEMMA 3.15. $M\leqq Z(L)$.

PROOF. Every Sylow 3-subgroup of L contains M since M is a normal 3-subgroup. The Sylow 3-subgroups are cyclic. Therefore it suffices to prove that the subgroup L_1 generated by all Sylow 3-subgroups of L coincides with L. But this easily follows from the fact that $L/M\cong K$ is generated by its Sylow 3-subgroups.

LEMMA 3.16. $|M|=1$.

PROOF. We use the following result, obtained by a simple application of the transfer: *If U is an abelian Hall subgroup of a group B, then* $|B'\cap Z(B)\cap U|=1$. ([10, p. 416, Satz 2.2].)

Let $B=L$, and U the Sylow 3-subgroup of L. We have $|U:M|=3$ (as $L/M\cong K$, 3.14). Now $Z(L)\cap U=Z(L)=M$, whence $|M\cap L'|=1$. This implies $L'\cong K$ by 3.8, and $L=M\times L'$. Both M and L' being characteristic subgroups of $L\in\theta$, it is impossible that both contain members of order 3 (by Lemma 2.4).

CONCLUSION. We have found a subgroup L, $K\cong L\lhd G$. By 3.7, $K=G/N$ is isomorphic to one of $SL(2,p)$, $PSL(2,p)$, $p\in\{5, 13\}$. Hence, by 3.5, $G=L\times N$. As L contains elements of order 3, N does not (3.9). We infer that $N=W$, which implies that *g.c.d.* $(|N|, |K|)=1$ (3.10). Moreover, $N=W$ is the direct product of elementary abelian groups (1.2). This completes the proof of Theorem 1.1.

REMARK. From these results one can easily derive that *every $\mathcal{D}-CI$ group is solvable.* ($\mathcal{D}$ stands for the class of digraphs.)

References

[1] A. ÁDÁM, Problem 2—10, *J. Comb. Theory,* **2** (1967), 393.
[2] B. ALSPACH and T. D. PARSONS, *Isomorphism of circulant graphs,* manuscript.
[3] L. BABAI, Isomorphism problem for a class of point-symmetric structures, *Acta Math. Acad. Sci. Hungar.,* **29** (1977), 329—336.
[4] L. BABAI, *Isomorphisms of Cayley graphs.* III, in preparation.

[5] L. Babai and P. Frankl, Isomorphisms of Cayley graphs I, *Proc. 5th Hungarian Combinat. Colloq.* (Keszthely, 1976), Bolyai—North-Holland, **18** (1978), 35—52.

[6] R. Brauer, Some applications of the theory of blocks of characters of finite groups II, *J. Algebra* **1** (1964), 307—334.

[7] D. Z. Djoković, Isomorphism problem for a special class of graphs, *Acta Math. Acad. Sci. Hungar.*, **21** (1970), 267—270.

[8] C. D. Godsil, *On Cayley graph isomorphisms*, submitted.

[9] D. Gorenstein and J. Walter, The characterization of finite groups with dihedral Sylow 2-subgroups I—II—III, *J. Algebra*, **2** (1965), 85—151, 218—270, 334—393.

[10] B. Huppert, *Endliche Gruppen I.* Springer, 1967.

[11] Л. А. Калужнин, М. Х. Клин, О некоторых числовых инвариантах групп подстановок, *Латвийский матем. ежегодник*, **18** (1976), 81—99.

[12] L. A. Kalužnin and R. Pöschel, Funktionen und Relationenalgebren, in „*Diskrete Mathematik*", Deutscher Verlag der Wissenschaften, 1978 (to appear).

[13] М. Х. Клих, Р. Пешель, Проблема изоморфизма циклических графов с p^2 и pq вершинами (abstract), presented at the Combinatorial Seminar (Odessa, 1975).

[14] P. P. Pálfy, On regular pronormal subgroups of symmetric groups, *Acta Math. Acad. Sci. Hungar.*, to appear.

[15] R. Pöschel, Untersuchungen von S-Ringen, insbesondere im Gruppenring von p-Gruppen, *Math. Nachrichten*, **60** (1974), 1—27.

[16] I. Schur, Untersuchungen über die Darstellung der endlichen Gruppen durch gebrochene lineare Substitutionen, *J. für reine u. angew. Math.*, **132** (1907), 85—137.

[17] M. Suzuki, On finite groups with cyclic Sylow subgroups for all odd primes, *Amer. J. Math.*, **77** (1955), 657—691.

[18] J. Turner, Point-symmetric graphs with a prime number of points, *J. Comb. Theory*, **3** (1967), 136—145.

(Received November 11, 1977)

DEPARTMENT OF ALGEBRA AND NUMBER THEORY
EÖTVÖS L. UNIVERSITY
H—1445 BUDAPEST, PF. 323.

MATHEMATICAL INSTITUTE OF THE
HUNGARIAN ACADEMY OF SCIENCES
H—1053 BUDAPEST, REÁLTANODA U. 13—15.

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 185—203.

ON THE UNIFORM MODULUS OF CONTINUITY OF THE OPERATOR OF BEST APPROXIMATION IN THE SPACE OF PERIODIC FUNCTIONS

By

A. KROÓ (Budapest)

Introducion and preliminary results

In this paper we shall present some results connected with the uniform continuity of best approximations.

In the last fifteen years problems dealing with local continuity of best approximation have been widely investigated. It was proved ([1], [2]) that the metric projection operator onto a finite-dimensional Čebysev subspace of $C[a, b]$ is pointwise Lip 1. An analogous result was obtained for the space L_p, $2<p<\infty$ ([3]). In [4], the local modulus of continuity of the metric projection operator in space L_p, $1\leqq p<2$ was discussed.

The problem of investigation of the uniform modulus of continuity of the operator of best approximation was raised by S. B. Stečkin. This subject was discussed in several papers ([5]—[8]).

Let $C[a, b]$ be the space of real valued functions continuous on $[a, b]$, and let the set of functions $\{\varphi_i\}_{i=0}^n$ be a Čebysev system on $[a, b]$. Further let U_n be the set of polynomials of the system $\{\varphi_i\}_{i=0}^n$, and let $p_n(f)\in U_n$ and $E_n(f)$ denote the polynomial and measure of best Čebysev approximation of the function $f\in C[a, b]$, respectively. Then for $\varepsilon>0$ and $M\subseteqq C[a, b]$ we can introduce the uniform modulus of continuity of the operator of best approximation on M as

$$\Omega_n^{(1)}(M, \varepsilon) = \sup_{f\in M} \sup_{\substack{g\in C[a,b] \\ \|f-g\|\leqq\varepsilon}} \|p_n(f)-p_n(g)\|.$$

Let us define also the uniform modulus of strong unicity of the operator of best approximation on M:

$$\Omega_n^{(2)}(M, \varepsilon) = \sup_{f\in M} \sup_{\substack{q_n\in U_n \\ \|f-q_n\|\leqq E_n(f)+\varepsilon}} \|p_n(f)-q_n\|.$$

As it was shown in [5], if $M=C[a, b]$ then $\Omega_n^{(1)}(M, \varepsilon)$ is even unbounded, so it was reasonable to consider $\Omega_n^{(1)}(M, \varepsilon)$ and $\Omega_n^{(2)}(M, \varepsilon)$ on sets M having certain compactness properties. (Later in [9] some sufficient conditions for $\Omega_n^{(1)}(M, \varepsilon)$ and $\Omega_n^{(2)}(M, \varepsilon)$ to converge to zero as $\varepsilon\to 0$ were obtained in the general case.)

The first result in this direction is due to P. V. Galkin [5]. He proved the following

Theorem. *Let* $r, n\in\mathbf{Z}_+$, $r<n$, $\varphi_i(x)=x^i$, $i=0, 1, \dots, n$, $M=W_r^* H[\omega]=$ $=\{f\in C[a, b]\colon \omega(f^{(r)}, \delta)\leqq\omega(\delta)\}$. *Then*

$$\Omega_n^{(1)}(M, \varepsilon) \overset{r,n}{\asymp} \begin{cases} R_n(\varepsilon), & \text{if} \quad n\geqq 1 \quad \text{and} \quad 0<\varepsilon\leqq\omega(b-a); \\ \varepsilon, & \text{otherwise}, \end{cases}$$

where $R_n(\varepsilon)$ satisfies the equation

$$R_n = \left(\frac{\varepsilon}{R_n}\right)^{r/n} \omega\left(\left(\frac{\varepsilon}{R_n}\right)^{1/n}\right)$$

and the constants involved depend only on r and n.

The same estimations were obtained for $\Omega_n^{(1)}(M, \varepsilon)$ and $\Omega_n^{(2)}(M, \varepsilon)$ in the periodic case [6] without any restriction on r and n and then in [7] an analogous result was proved in the algebraic case.

So the orders of $\Omega_n^{(1)}(\varepsilon)$ and $\Omega_n^{(2)}(\varepsilon)$ as functions of ε were determined and it turned out that they are closely connected with some Kolmogorov-type inequalities.

In the present paper we shall study the order of magnitude of $\Omega_n^{(1)}(\varepsilon)$ and $\Omega_n^{(2)}(\varepsilon)$ in the periodic case with respect to ε and n simultaneously. We shall obtain uniform estimations for $\Omega_n^{(1)}(\varepsilon)$ and $\Omega_n^{(2)}(\varepsilon)$ in terms of ε and n. Moreover, the upper estimation will be uniform even with respect to r.

In the second part of this paper we consider some classes of infinitely differentiable functions and study $\Omega_n^{(1)}(M, \varepsilon)$ and $\Omega_n^{(2)}(M, \varepsilon)$ on these sets.

§ 1

Let $C[-\pi, \pi]$ be the space of 2π-periodic real valued functions continuous on $[-\pi, \pi]$, T_k $(k \in \mathbf{Z}_+)$ the set of trigonometric polynomials of order at most k, $W_r H[\omega] = \{f \in C[-\pi, \pi]: \omega(f^{(r)}, \delta) \leqq \omega(\delta)\}$, where $\omega(\delta)$ is a fixed modulus of continuity, $r \in \mathbf{Z}_+$. Then we can define the uniform moduli $\Omega_k^{(1)}(W_r H[\omega], \varepsilon)$ and $\Omega_k^{(2)}(W_r H[\omega], \varepsilon)$ as in the Introduction. Further, let $R_k = R_k(\varepsilon)$ be the unique solution of the equation

$$\frac{1}{k^r}\left(\frac{\varepsilon}{R_k}\right)^{r/2k} \omega\left(\frac{1}{k}\left(\frac{\varepsilon}{R_k}\right)^{1/2k}\right) = R_k. \tag{1}$$

Then we have the following

THEOREM 1. *Let $r, k \in \mathbf{Z}_+$, $\varepsilon > 0$ and $\omega(\delta)$ a modulus of continuity. Then there exist absolute constants $1 \leqq c_0, c_1, c_2$ and a positive constant $c_3(r)$ depending only on r, such that*

A) *if* $k \geqq 1$ *and* $0 < \varepsilon \leqq \frac{1}{k^r}\omega\left(\frac{1}{k}\right)$ *then*

$$c_3(r) R_k(\varepsilon) \leqq \Omega_k^{(j)}(W_r H[\omega], \varepsilon) \leqq c_0 R_k(c_1^{2k}\varepsilon) \quad (j = 1, 2); \tag{2}$$

B) *if* $k = 0$ *or* $\varepsilon > \frac{1}{k^r}\omega\left(\frac{1}{k}\right)$ *then*

$$\varepsilon \leqq \Omega_k^{(j)}(W_r H[\omega], \varepsilon) \leqq c_2 \varepsilon \quad (j = 1, 2). \tag{3}$$

REMARK 1. Evidently $R_k(c_1^{2k}\varepsilon) \leqq c_4(r) R_k(\varepsilon)$ and (2) implies that in the case A

$$c_3(r) R_k(\varepsilon) \leqq \Omega_k^{(j)}(W_r H[\omega], \varepsilon) \leqq c_4(r) R_k(\varepsilon) \quad (j = 1, 2). \tag{4}$$

COROLLARY. *Let* $r \in \mathbf{Z}_+$, $k \in \mathbf{N}$, $\omega(\delta) = \delta^\alpha$, $0 < \alpha \leqq 1$, $0 < \varepsilon \leqq k^{-(r+\alpha)}$. *Then*

$$c_3(r)\left(\frac{1}{k}\right)^{2k(r+\alpha)/(2k+r+\alpha)} \varepsilon^{(r+\alpha)/(2k+r+\alpha)} \leqq \Omega_k^{(j)}(W_r H[\omega], \varepsilon) \leqq \tag{5}$$

$$\leqq c_0\left(\frac{c_1}{k}\right)^{2k(r+\alpha)/(2k+r+\alpha)} \varepsilon^{(r+\alpha)/(2k+r+\alpha)}.$$

In what follows $c_i(\dots)$ denote constants depending only on quantities specified in the brackets, while c_i denote absolute constants.

Some lemmas

Let us formulate the first lemma. Its proof can be found in [1].

Let $\{\varphi_i\}_{i=0}^n$ be a Čebysev system on $[a, b]$ and consider the system of points $(a \leqq) x_0 < x_1 < \dots < x_{n+1} (\leqq b)$. The number of these points is $n+2$ so if we omit the rth point $(0 \leqq r \leqq n+1)$ and consider the set of points $x_0 < x_1 < \dots < x_{r-1} < < x_{r+1} < \dots < x_{n+1}$ then for any $0 \leqq i \leqq n+1$, $i \neq r$ the fundamental polynomials $L_{i,r}^n(x)$ of the Lagrange interpolation corresponding to this set can be defined, i.e.

$$\begin{cases} L_{i,r}^n(x_j) = 0 & \text{for} \quad 0 \leqq j \leqq n+1, \quad j \neq r, i; \\ L_{i,r}^n(x_i) = 1 \end{cases}$$

$(0 \leqq i \leqq n+1,\ i \neq r)$. Further let

$$\lambda_n = \max_{0 \leqq r \leqq n+1} \left\| \sum_{\substack{i=0 \\ i \neq r}}^{n+1} |L_{i,r}^n(x)| \right\|. \tag{6}$$

LEMMA 1. *Let* $\mu > 0$ *and* $q_n(x)$ *a polynomial of the system* $\{\varphi_i\}_{i=0}^n$ *satisfying the inequalities*

$$\gamma(-1)^{i+1} q_n(x_i) \leqq \mu \quad (i = 0, 1, \dots, n+1;\ \gamma = \pm 1).$$

Then

$$\|q_n\| \leqq \lambda_n \mu. \tag{7}$$

Now we shall give an estimate for λ_n in the periodic case which is verified in [12]. Consider the set of points $(-\pi \leqq) x_1 < x_2 < \dots < x_{2k+2} (< \pi)$. Then

$$L_{i,r}^k(x) = \frac{\prod\limits_{\substack{j=1 \\ j \neq i,r}}^{2k+2} \sin \frac{x - x_j}{2}}{\prod\limits_{\substack{j=1 \\ j \neq i,r}}^{2k+2} \sin \frac{x_i - x_j}{2}} \quad (1 \leqq i \leqq 2k+2,\ i \neq r,\ 1 \leqq r \leqq 2k+2).$$

Further define λ_k as in (6). Then we have the following

LEMMA 2. *Let the set of points* $(-\pi \leqq) x_1 < x_2 < \ldots < x_{2k+2} (< \pi)$, $k \geqq 1$ *satisfy the inequalities*

$$x_{i+1} - x_i \geqq d \quad (i = 1, 2, \ldots, 2k+1),$$
$$2\pi - (x_{2k+2} - x_1) \geqq d,$$

where $0 < d < \frac{\pi}{k+1}$. *Then*

$$\lambda_k \leqq \max_{1 \leqq r \leqq 2k+2} \sum_{\substack{i=1 \\ i \neq r}}^{2k+2} \frac{1}{\left| \prod_{\substack{j=1 \\ j \neq i, r}}^{2k+2} \sin \frac{x_i - x_j}{2} \right|} \leqq \left(\frac{\pi e}{kd} \right)^{2k}. \tag{8}$$

The next lemma will be proved with the help of the following Kolmogorov-type inequality proved in [6]: for any $f \in W_r H[\omega]$

$$\|f^{(l)}\| \leqq c_6(r, l) \frac{\|f\|}{\lambda^l_{r,\omega}(\|f\|)} \quad (1 \leqq l \leqq r) \tag{9}$$

holds where $\lambda_{r,\omega}(y) = \min \{x: x^r \omega(x) = y\}$. In [6], $c_6(r, l) \leqq \pi 2^l$ is essentially proved, hence for $l = 1$ we have

$$\|f'\| \leqq 2\pi \frac{\|f\|}{\lambda_{r,\omega}(\|f\|)}. \tag{10}$$

LEMMA 3. *For any* $f \in W_r H[\omega]$ *and* $x, y \in [-\pi, \pi]$ *such that* $|f(x) - f(y)| \geqq \|f\|$

$$|x - y| \geqq \frac{1}{2\pi} \lambda_{r,\omega}(\|f\|) \tag{11}$$

holds.

PROOF. For $r = 0$ the statement is trivial. Let $r \geqq 1$. Then using (10) we obtain

$$\|f\| \leqq |f(x) - f(y)| \leqq \|f\|' |x - y| \leqq 2\pi \frac{\|f\|}{\lambda_{r,\omega}(\|f\|)} |x - y|$$

and the lemma is proved.

COROLLARY. *For any* $f \in W_r \operatorname{Lip} \alpha$, $0 < \alpha \leqq 1$ *and* $x, y \in [-\pi, \pi]$ *such that* $|f(x) - f(y)| \geqq \|f\|$,

$$|x - y| \geqq \frac{1}{2\pi} \|f\|^{1/(r+\alpha)}.$$

Now we shall investigate the following question: how can $\omega(p_k^{(r)}(f), \delta)$ be estimated from above if $f \in W_r H[\omega]$. By a theorem of S. B. STEČKIN [10], in this case $\omega\big(p_k^{(r)}(f), \delta\big) \leqq c_7(r)\omega(\delta)$. In the following lemma (which will be proved by the method used in [10]) we obtain an upper estimation for $c_7(r)$.

LEMMA 4. *For any* $f \in W_r H[\omega]$, $r \in \mathbf{Z}_+$, $k \in \mathbf{N}$

$$\omega\big(p_k^{(r)}(f), \delta\big) \leqq c_8 \ln(r+2) \omega(\delta) \tag{12}$$

holds.

PROOF. Assume at first that $\delta \geqq \frac{1}{k}$. Then

$$(13) \qquad \omega\big(p_k^{(r)}(f), \delta\big) \leqq \omega(f^{(r)}, \delta) + 2\|f^{(r)} - p_k^{(r)}(f)\| \leqq \omega(\delta) + 2\|f^{(r)} - p_k^{(r)}(f)\|.$$

By a theorem proved in [13]

$$\|f^{(r)} - p_k^{(r)}(f)\| \leqq c_9 \ln(r+2) E_k(f^{(r)}).$$

Hence and from (13) we obtain for $\delta \geqq \frac{1}{k}$

$$(14) \qquad \omega\big(p_k^{(r)}(f), \delta\big) \leqq \omega(\delta) + 2c_9 \ln(r+2) E_k(f^{(r)}) \leqq$$

$$\leqq \omega(\delta) + c_{10} \ln(r+2)\, \omega\left(\frac{1}{k}\right) \leqq c_{11} \ln(r+2)\, \omega(\delta).$$

Set now $0 < \delta < \frac{1}{k}$. Then using (14) for $\delta = \frac{1}{k}$ and STEČKIN's inequality [10] we obtain

$$\omega\big(p_k^{(r)}(f), \delta\big) \leqq \delta \|p_k^{(r+1)}(f)\| \leqq \frac{\delta k}{2}\, \omega\left(p_k^{(r)}(f), \frac{\pi}{k}\right) \leqq$$

$$\leqq c_{12} \delta k \ln(r+2)\, \omega\left(\frac{1}{k}\right) \leqq c_{13} \ln(r+2)\, \omega(\delta).$$

This inequality together with (14) completes the proof of the Lemma.

Now we shall prove our main lemma.

LEMMA 5. *Let $k \in \mathbf{N}$, $r \in \mathbf{Z}_+$, $f \in W_r H[\omega] \setminus T_k$. Then for any $q_k \in T_k$*

$$(15) \qquad \|p_k(f) - q_k\| \leqq \left(\frac{c_{14}}{k \lambda_{r,\omega}\big(c_0^* E_k(f)\big)}\right)^{2k} \{\|f - q_k\| - E_k(f)\}$$

holds.

PROOF. Set $\bar{f} = f - p_k(f)$. Then by (12) $\bar{f} \in W_r H[\bar{\omega}]$, where $\bar{\omega}(\delta) = c_{15} \ln(r+2)\, \omega(\delta)$. Without loss of generality we may assume that $c_{15} \ln(r+2) > 1$. Further set $\bar{q}_k = q_k - p_k(f)$, then

$$(16) \qquad \|\bar{f} - \bar{q}_k\| = E_k(f) + \mu$$

where $\mu = \|f - q_k\| - E_k(f)$. There exists a system of points $(-\pi \leqq) x_1 < x_2 < \dots$ $\dots < x_{2k+2} (< \pi)$ such that

$$(17) \qquad \bar{f}(x_i) = \gamma(-1)^i \|\bar{f}\| = \gamma(-1)^i E_k(f) \quad (i = 1, 2, \dots, 2k+2,\ |\gamma| = 1).$$

(16) and (17) imply

$$\gamma(-1)^{i+1} \bar{q}_k(x_i) \leqq \mu \quad (i = 1, 2, \dots, 2k+2).$$

Then by Lemma 1

$$(18) \qquad \|\bar{q}_k\| \leqq \lambda_k \mu.$$

Further $\bar{f} \in W_r H[\bar{\omega}]$ and $|\bar{f}(x_{i+1}) - \bar{f}(x_i)| = 2E_k(f) = 2\|\bar{f}\|$. Hence (11) implies that the set of points $\{x_i\}_{i=1}^{2k+2}$ satisfies the conditions of Lemma 2 with $d=\frac{1}{2\pi}\lambda_{r,\bar\omega}(\|\bar f\|)$ and using (18) and (8) we have

$$\|\bar q_k\| \leqq \left(\frac{2\pi^2 e}{k\lambda_{r,\bar\omega}(E_k(f))}\right)^{2k}\mu. \tag{19}$$

Let us compare $\lambda_{r,\bar\omega}(E_k(f))$ with $\lambda_{r,\omega}(E_k(f))$. Evidently $\lambda_{0,\bar\omega}(E_k(f))=\lambda_{0,\omega}(c_0^* E_k(f))$. Let $r\geqq 1$. By $\bar\omega>\omega$, $\lambda_{r,\bar\omega}(E_k(f))<\lambda_{r,\omega}(E_k(f))$ holds. Further, set $\lambda_{r,\bar\omega}(E_k(f))=\bar a$, $\lambda_{r,\omega}(E_k(f))=a$, $E_k(f)=b$. Then

$$\begin{cases} \bar a^r\bar\omega(\bar a)=b \\ a^r\omega(a)=b \\ \bar a<a \end{cases}$$

hence $b=\bar a^r\bar\omega(\bar a)\leqq\bar a^r\bar\omega(a)\leqq\bar a^r c_{15}\ln(r+2)\omega(a)$. On the other hand $b=a^r\omega(a)$ i.e. $a^r\omega(a)\leqq c_{15}\ln(r+2)\bar a^r\omega(a)$ or $a\leqq(c_{15}\ln(r+2))^{1/r}\bar a\leqq c_{16}\bar a$ i.e.

$$\lambda_{r,\bar\omega}(E_k(f))\geqq\frac{1}{c_{16}}\lambda_{r,\omega}(E_k(f))$$

and this together with (19) completes the proof of Lemma 5.

Lemma 5 is the strong unicity theorem in the periodic case (see [2]). It immediately implies the periodic version of FREUD's theorem [1].

COROLLARY 1. *Let* $k\in\mathbf{N}$, $r\in\mathbf{Z}_+$, $f\in W_rH[\omega]\setminus T_k$. *Then for any* $f_1\in C[-\pi,\pi]$

$$\|p_k(f)-p_k(f_1)\|\leqq\left(\frac{c_{17}}{k\lambda_{r,\omega}(c_0^* E_k(f))}\right)^{2k}\|f-f_1\|. \tag{20}$$

PROOF. (20) follows from (15) and the obvious inequality

$$\|f-p_k(f_1)\|\leqq E_k(f)+2\|f-f_1\|. \tag{21}$$

COROLLARY 2. *Let* $k\in\mathbf{N}$, $r\in\mathbf{Z}_+$, $0<\alpha\leqq1$, $M>0$, $f\in W_r\,\mathrm{Lip}_M\,\alpha$. *Then*

a) *for any* $q_k\in T_k$

$$\|p_k(f)-q_k\|\leqq\left(\frac{c_{14}}{k}\right)^{2k}\left(\frac{c_0' M}{E_k(f)}\right)^{2k/(r+\alpha)}\{\|f-q_k\|-E_k(f)\}; \tag{22}$$

b) *for any* $f_1\in C[-\pi,\pi]$

$$\|p_k(f)-p_k(f_1)\|\leqq\left(\frac{c_{17}}{k}\right)^{2k}\left(\frac{c_0' M}{E_k(f)}\right)^{2k/(r+\alpha)}\|f-f_1\| \tag{23}$$

hold.

Proof of the upper estimation of the Theorem

Let us prove the upper estimation for $\Omega_k^{(2)}(W_rH[\omega],\varepsilon)$.

Let $f\in W_rH[\omega]$ and $q_k\in T_k$ such that

$$\|f-q_k\|\leqq E_k(f)+\varepsilon. \tag{24}$$

Set $\bar f=f-p_k(f)$, $\bar q_k=q_k-p_k(f)$.

Case I: $k=0$. Then there exist points $x, y\in[-\pi, \pi]$ such that $\bar{f}(x)=\|\bar{f}\|$, $\bar{f}(y)=-\|\bar{f}\|$. Further (24) implies that $\|\bar{f}-\bar{q}_0\|\leqq\|\bar{f}\|+\varepsilon$; hence $\bar{q}_0(x)\geqq-\varepsilon$, $\bar{q}_0(y)\leqq\varepsilon$ i.e. $\|\bar{q}_0\|\leqq\varepsilon$.

Case II: $k\geqq 1$, $\varepsilon>\frac{1}{k^r}\,\omega\left(\frac{1}{k}\right)$. Then using (24) we obtain

$$\|p_k(f)-q_k\| \leqq E_k(f)+\|f-q_k\| \leqq 2E_k(f)+\varepsilon. \tag{25}$$

But $E_k(f)\leqq\frac{c_{18}}{k^r}\,\omega\left(\frac{1}{k}\right)\leqq c_{18}\varepsilon$; hence and from (25) $\|p_k(f)-q_k\|\leqq c_{19}\varepsilon$ and the upper estimation of the theorem in Case B is proved.

Case III: $k\geqq 1$, $0<\varepsilon\leqq\frac{1}{k^r}\,\omega\left(\frac{1}{k}\right)$, $f\in T_k$. Then $E_k(f)=0$ and hence

$$\|p_k(f)-q_k\| \leqq 2E_k(f)+\varepsilon = \varepsilon \leqq R_k(c_1^{2k}\varepsilon).$$

Case IV: $k\geqq 1$, $0<\varepsilon\leqq\frac{1}{k^r}\,\omega\left(\frac{1}{k}\right)$; $f\in W_rH[\omega]\setminus T_k$. Then (15) and (24) imply

$$\|p_k(f)-q_k\| \leqq \frac{\varepsilon}{\{c_{20}k\lambda_{r,\omega}(c_0^* E_k(f))\}^{2k}} \tag{26}$$

where we may assume that $c_{20}\leqq 1$. Thus by (25)

$$\|p_k(f)-q_k\| \leqq \bar{c}_0 \sup_{E>0} \min\left\{E+\varepsilon, \frac{\varepsilon}{(c_{20}k\lambda_{r,\omega}(E))^{2k}}\right\}. \tag{27}$$

It is known [11] that there exists a strictly increasing modulus of continuity ω_1 such that $\omega\leqq\omega_1\leqq 2\omega$, therefore without loss of generality we may assume that ω is strictly increasing. Then if $r\geqq 1$ the continuous function $\frac{\varepsilon}{(c_{20}k\lambda_{r,\omega}(E))^{2k}}$ is monotonically decreasing from $+\infty$ to 0 as E varies from 0 to $+\infty$. If $r=0$ then it decreases from $+\infty$ to $\frac{\varepsilon}{(c_{20}k\pi)^{2k}}<\varepsilon<\varepsilon+\omega(\pi)$. (Here we assumed that $k\geqq\frac{1}{c_{20}}$. If $1\leqq k<\frac{1}{c_{20}}$ then we can replace (27) by

$$\|p_k(f)-q_k\| \leqq \frac{\bar{c}_0}{c_{20}^{2/c_{20}}} \sup_{E>0} \min\left\{E+\varepsilon, \frac{\varepsilon}{(k\lambda_{r,\omega}(E))^{2k}}\right\}$$

and do the same transformations.)
Hence for any $r\geqq 0$

$$\sup_{E>0} \min\left\{E+\varepsilon, \frac{\varepsilon}{(c_{20}k\lambda_{r,\omega}(E))^{2k}}\right\} = R_k'+\varepsilon \tag{28}$$

where $R_k'=R_k'(\varepsilon)$ satisfies the equation

$$R_k'+\varepsilon = \frac{\varepsilon}{(c_{20}k\lambda_{r,\omega}(R_k'))^{2k}}$$

or

$$R'_k = \frac{1}{c_{20}^r k^r} \left(\frac{\varepsilon}{R'_k+\varepsilon}\right)^{r/2k} \omega\left(\frac{1}{c_{20}k}\left(\frac{\varepsilon}{R'_k+\varepsilon}\right)^{1/2k}\right).$$

Consider the equation

$$R^*_k = \frac{1}{(c_{20}k)^r} \left(\frac{\varepsilon}{R^*_k}\right)^{r/2k} \omega\left(\frac{1}{c_{20}k}\left(\frac{\varepsilon}{R^*_k}\right)^{1/2k}\right).$$

Obviously $R'_k \leqq R^*_k$, hence (27) and (28) imply

$$\|p_k(f)-q_k\| \leqq \bar{c}_0(R'_k+\varepsilon) \leqq \bar{c}_0(R^*_k+\varepsilon)$$

and using that $c_{20} \leqq 1$ we evidently have that $R^*_k > \varepsilon$ if $0<\varepsilon<\frac{1}{k^r}\omega\left(\frac{1}{k}\right)$, thus

$$\|p_k(f)-q_k\| \leqq c_0 R^*_k(\varepsilon) = c_0 R_k\left(\left(\frac{1}{c_{20}}\right)^{2k}\varepsilon\right).$$

This completes the proof of the upper estimation of the Theorem for $\Omega_k^{(2)}(W_r H[\omega], \varepsilon)$. Using (21) we can analogously prove the upper estimation for $\Omega_k^{(1)}(W_r H[\omega], \varepsilon)$.

The proof of the lower estimation of the Theorem

Evidently, it is enough to prove the lower estimation for $\Omega_k^{(1)}(W_r H[\omega], \varepsilon)$. If $k=0$ or $\varepsilon>\frac{1}{k^r}\omega\left(\frac{1}{k}\right)$ we may set $g=f+\varepsilon$; then $\|p_k(f)-p_k(g)\|=\varepsilon$.

Let $k\geqq 1$ and $0<\varepsilon<\frac{1}{k^r}\omega\left(\frac{1}{k}\right)$. We shall use the counter-example constructed in [5]. Consider the set of points

$$x_1 = -\frac{\pi}{4}, \quad x_i = x_1+(i-1)2^r h \quad (h>0,\ i=2,3,\dots,4k+4).$$

Then, as in [5], we can construct a function satisfying the following properties:

a) $f\in W^r H[\omega]$;

b) $f(x_{2i}) = (-1)^{i+1}\|f\|$, $f(x_{2i-1}) = 0$, $(i=1,2,\dots,2k+2)$;

c) $\|f\| = c_{21}(r)h^r\omega(h)$;

d) $f=0$ if $x\in[-\pi, x_1]$; $f<0$ if $x\in(x_{4k+3},\pi)$; $\operatorname{sign} f = (-1)^i$, $x\in[x_{2i+1}, x_{2i+3}]$ $(i=0,1,\dots,2k)$.

Now we shall define h. Let $R_k = R_k(\varepsilon)$ satisfy equation (1). It is identical with

$$R_k = \frac{\varepsilon}{k^{2k}\lambda_{r,\omega}^{2k}(R_k)}.$$

Further set

$$h = \frac{1}{2^r 100} \lambda_{r,\omega}(R_k). \tag{29}$$

Then

$$(2^r 100h)^r \omega(2^r 100h) = R_k = \frac{\varepsilon}{k^{2k}(2^r 100h)^{2k}}$$

i.e.

$$(2^r 100h)^{r+2k} \omega(2^r 100h) = \frac{\varepsilon}{k^{2k}} < \left(\frac{1}{k}\right)^{2k+r} \omega\left(\frac{1}{k}\right),$$

$$h < \frac{1}{2^r 100k}. \tag{30}$$

Then $\{x_i\}_{i=1}^{4k+4} \in \left[-\frac{\pi}{4}, \frac{\pi}{4}\right]$. Further property c) of the function f implies

$$c_{22}(r) R_k(\varepsilon) \leqq \|f\| \leqq c_{23}(r) R_k(\varepsilon). \tag{31}$$

Define

$$g(x) = \begin{cases} f(x) & \text{if} \quad -\pi \leqq x \leqq x_{4k+3}; \\ \dfrac{\|f\|+\varepsilon}{\|f\|} f(x) & \text{if} \quad x_{4k+3} \leqq x \leqq \pi. \end{cases}$$

Then $g \in C[-\pi, \pi]$ and $\|f-g\| = \varepsilon$. It is easy to show that $E_k(g) = \|f\| + \mu$, where $\mu > 0$, hence

$$(-1)^i p_k(g, x_{2i}) \leqq \mu \quad (i = 1, 2, \dots, 2k+2).$$

Then setting $a = x_2$, $b = x_{4k+4}$, $\gamma = 1$ in Lemma 1 and using (8) with $d = 2^{r+1}h$ we have

$$\|p_k(g)\|_{[x_2, x_{4k+4}]} \leqq \mu \max_{1 \leqq m \leqq 2k+2} \left\| \sum_{\substack{i=1 \\ i \neq m}}^{2k+2} \left| \frac{\prod\limits_{\substack{j=1 \\ j \neq i, m}}^{2k+2} \sin \dfrac{x - x_{2j}}{2}}{\prod\limits_{\substack{j=1 \\ j \neq i, m}}^{2k+2} \sin \dfrac{x_{2i} - x_{2j}}{2}} \right| \right\|_{[x_2, x_{4k+4}]} \leqq$$

$$\leqq \mu \sin^{2k} \frac{x_{4k+4} - x_2}{2} \max_{1 \leqq m \leqq 2k+2} \sum_{\substack{i=1 \\ i \neq m}}^{2k+2} \frac{1}{\left| \prod\limits_{\substack{j=1 \\ j \neq i, m}}^{2k+2} \sin \dfrac{x_{2i} - x_{2j}}{2} \right|} \leqq$$

$$\leqq \mu [(2k+1) 2^r h]^{2k} \left(\frac{e\pi}{2^{r+1}hk}\right)^{2k} \leqq \mu (5\pi)^{2k}.$$

Using this inequality we obtain

$$\|f\| + \varepsilon - \mu(5\pi)^{2k} \leqq |g(x_{4k+4})| - |p_k(g, x_{4k+4})| \leqq |g(x_{4k+4}) - p_k(g, x_{4k+4})| \leqq \|f\| + \mu$$

i.e.

$$\varepsilon \leqq \left(1 + (5\pi)^{2k}\right) \mu \leqq (6\pi)^{2k} \mu. \tag{32}$$

Assume that $\|p_k(g)\|<\|f\|$. There exists a system of points $(-\pi\leqq)\bar{x}_1<\dots$ $\dots<\bar{x}_{2k+2}(<\pi)$ such that

$$(33)\qquad g(\bar{x}_i)-p_k(g,\bar{x}_i)=\gamma(-1)^i E_k(g)\quad (i=1,2,\dots,2k+2;\ |\gamma|=1).$$

Further, let us prove that there is no point $\bar{x}\in[x_{4k+3},\pi]$ such that

$$g(\bar{x})-p_k(g,\bar{x})=E_k(g).$$

Indeed, assume the contrary. Then

$$g(\bar{x})=p_k(g,\bar{x})+E_k(g)>\|f\|+\mu-\|f\|=\mu>0$$

but $g\leqq 0$ if $x\in[x_{4k+3},\pi]$. Evidently $\bar{x}_i\notin[-\pi,x_1]$ $(i=1,2,\dots,2k+2)$, hence $\{\bar{x}_i\}_{i=1}^{2k+1}\in[x_1,x_{4k+3}]$ and using that $|g|\leqq\|f\|$ on $[x_1,x_{4k+3}]$ we obtain from (33)

$$(34)\qquad \gamma(-1)^{i+1}p_k(g,\bar{x}_i)=E_k(g)+\gamma(-1)^{i+1}g(\bar{x}_i)\geqq\|f\|+\mu-\|f\|=\mu$$

$$(i=1,2,\dots,2k+1).$$

By Lagrange interpolatory formula and (34)

$$\gamma p_k(g,\pi)=\gamma\sum_{i=1}^{2k+1}p_k(g,\bar{x}_i)\frac{\prod\limits_{\substack{j=1\\ j\neq i}}^{2k+1}\sin\frac{\pi-\bar{x}_j}{2}}{\prod\limits_{\substack{j=1\\ j\neq i}}^{2k+1}\sin\frac{\bar{x}_i-\bar{x}_j}{2}}=$$

$$=\sum_{i=1}^{2k+1}\gamma(-1)^{i+1}p_k(g,\bar{x}_i)\frac{\prod\limits_{\substack{j=1\\ j\neq i}}^{2k+1}\sin\frac{\pi-\bar{x}_j}{2}}{\left|\prod\limits_{\substack{j=1\\ j\neq i}}^{2k+1}\sin\frac{\bar{x}_i-\bar{x}_j}{2}\right|}\geqq$$

$$\geqq\mu\left(\frac{\sqrt{2}}{2}\right)^{2k}\frac{2k+1}{\left(\frac{x_{4k+3}-x_1}{2}\right)^{2k}}\geqq\frac{2^k\mu}{(4k+2)^{2k}2^{2kr}h^{2k}},$$

hence and from (32) and (29) we get

$$\|p_k(g)\|\geqq\frac{2^k\varepsilon}{(6\pi)^{2k}(4k+2)^{2k}2^{2kr}h^{2k}}=\left(\frac{\sqrt{2}\,100}{36\pi}\right)^{2k}\frac{\varepsilon}{k^{2k}\lambda_{r,\omega}^{2k}(R_k)}>\frac{\varepsilon}{k^{2k}\lambda_{r,\omega}^{2k}(R_k)}=R_k(\varepsilon).$$

We obtained this inequality assuming that $\|p_k(g)\|\leqq\|f\|$. Hence and from (31) we have

$$\|p_k(g)\|\geqq\min\{\|f\|,\ R_k(\varepsilon)\}=c_{24}(r)\,R_k(\varepsilon).$$

The Theorem is proved.

§ 2

As it was mentioned already, the operator of best approximation satisfies a point-wise Lip 1 condition and evidently this is the best possible result. By (5) we can see that the uniform Lipschitz exponent for the class $W_r \operatorname{Lip} \alpha$ is $\frac{r+\alpha}{2k+r+\alpha}$ i.e. it remains below 1. But this exponent tends to 1 as $r \to \infty$ hence one may expect that for classes of analytic or even of infinitely differentiable functions, the uniform modulus of continuity of the operator of best approximation will be close to Lip 1. In this section we shall obtain some results of this type, thus answering some problems raised by S. B. Stečkin. More precisely his question was the following:

Let F_ϱ be a bounded open set of the complex plane including the interval $[-\pi, \pi]$ and let γ_ϱ be its boundary satisfying

$$\varrho = \inf_{\substack{x \in [-\pi, \pi] \\ z \in \gamma_\varrho}} |x-z|, \quad \varrho > 0.$$

Further, let $A(M, \varrho)$ be the set of 2π-periodic, real-valued functions analytic on $[-\pi, \pi]$ which can be analytically extended into F_ϱ, are continuous on γ_ϱ and satisfy $\sup_{z \in \gamma_\varrho} |f(z)| \leqq M$ with a given $M > 0$. Then the question is whether there exists a positive constant $\beta > 0$ such that

$$\Omega_k^{(j)}\big(A(M, \varrho), \varepsilon\big) \leqq c_{25}(k, M, \varrho)\varepsilon |\ln \varepsilon|^\beta \quad (j = 1, 2) \tag{35}$$

for ε small enough.

Here the conjecture is that the uniform moduli will satisfy an "almost" Lip 1 condition. First of all let us prove that $|\ln \varepsilon|^\beta$ cannot be omitted on the right side of (35), i.e. the uniform moduli do not satisfy the Lip 1 condition.

Indeed, assume the contrary. Then for example $\Omega_k^{(1)}\big(A(M, \varrho), \varepsilon\big) \leqq c_{25}(k, M, \varrho)\varepsilon$. Then for any $\alpha > 0$

$$\Omega_k^{(1)}\big(\alpha A(M, \varrho), \varepsilon\big) \leqq c_{25}(k, M, \varrho)\varepsilon \tag{36}$$

holds (see e.g. [9]). In the previous section we constructed continuous functions f_ε and g_ε such that $\|f_\varepsilon - g_\varepsilon\| = \varepsilon$ and $\|p_k(f_\varepsilon) - p_k(g_\varepsilon)\| \geqq c_{26}(k)\varepsilon^\delta$ for any $0 < \varepsilon < \varepsilon_0$ where $0 < \delta < 1$. Further $p_m(f_\varepsilon) \rightrightarrows f_\varepsilon$, $p_m(g_\varepsilon) \rightrightarrows g_\varepsilon$; $p_k\big(p_m(f_\varepsilon)\big) \rightrightarrows p_k(f_\varepsilon)$; $p_k\big(p_m(g_\varepsilon)\big) \rightrightarrows p_k(g_\varepsilon)$ when $m \to \infty$. Hence for m large enough

$$\|p_m(f_\varepsilon) - p_m(g_\varepsilon)\| < 2\varepsilon,$$

$$\|p_k\big(p_m(f_\varepsilon)\big) - p_k\big(p_m(g_\varepsilon)\big)\| \geqq \frac{c_{26}(k)}{2}\varepsilon^\delta$$

hold. But $p_m(f_\varepsilon) \in \alpha A(M, \varrho)$ for any $0 < \varepsilon < \varepsilon_0$ and α large enough, and this contradicts (36).

We shall prove a more general theorem which, in particular, will imply (35).

Let $a \geqq 0$, $N > 1$ and consider the class of functions $C_{a,N}^\infty$ consisting of 2π-periodic real valued infinitely differentiable functions f such that $\|f^{(r)}\| \leqq N^r r^{ar}$ for any $r \geqq 1$. Then for $a = 0$ obviously $C_{0,N}^\infty \subseteqq T_N$; if $0 < a < 1$ then $C_{a,N}^\infty$ contains only entire functions and for $a=1$, $C^\infty_{1,N}$ consists only of analytic functions (in this case the opposite statement is also true: if f is analytic then $f\in C^\infty_{1,N}$ with some N).

EXAMPLE. It is easy to show that with appropriate $N=N(\varrho, M)$, $A(M,\varrho)\subseteqq C^\infty_{1,N}$.

THEOREM 2. *Let* $k\in\mathbf{N}$, $\varepsilon>0$, $N>1$, $a\geqq 0$.

A$_1$) *If* $a>0$, $k\geqq k_0=1+\left[N\left\{\frac{e}{a}\left(\ln\frac{\pi}{2}+2a\right)\right\}^a\right]$, *then*

$$\Omega_k^{(j)}(C^\infty_{a,N},\varepsilon)\leqq\begin{cases}\left(\dfrac{c_{27}}{k}\right)^{2k}N^{2k}|\ln\varepsilon|^{2ka}\varepsilon & \text{if}\quad 0<\varepsilon<\dfrac{\pi}{2}e^a e^{-(k/N)^{1/a}a/e};\\ 4\varepsilon, & \text{otherwise},\quad (j=1,2).\end{cases}\tag{37}$$

A$_2$) *If* $a>0$, $1\leqq k\leqq k_0$ *then for any* $0<\varepsilon<c_{28}(a)$

$$\Omega_k^{(j)}(C^\infty_{a,N},\varepsilon)\leqq\left(\frac{c_{29}(a)}{k}\right)^{2k}N^{2k}|\ln\varepsilon|^{2ka}\varepsilon;\quad (j=1,2)\tag{38}$$

holds.

B) *If* $a=0$ *then*

$$\Omega_k^{(j)}(C^\infty_{0,N},\varepsilon)\leqq\Omega_k^{(j)}(T_N,\varepsilon)\equiv c_{30,j}(k,N)\varepsilon\quad (j=1,2),\tag{39}$$

where the constants $c_{30,j}(k,N)$ $(j=1,2)$ *satisfy*

$$c_{30,j}(k,N)\leqq\begin{cases}2, & \text{if}\quad k\geqq N\\ \left(\dfrac{c_{27}}{k}\right)^{2k}N^{2k}, & \text{if}\quad 1\leqq k<N.\end{cases}\tag{40}$$

We shall prove some lemmas.

The first lemma presents a result which is essentially the strong unicity theorem (Case I) and Freud's continuity theorem (Case II) for the functions from $C^\infty_{a,N}$.

LEMMA 6. *Let* $k\in\mathbf{N}$, $a\geqq 0$, $N>1$, $f\in C^\infty_{a,N}\setminus T_k$. *Then*

I) *For any* $q_k\in T_k$

$$\|p_k(f)-q_k\|\leqq\left(\frac{c_{14}}{k}\right)^{2k}N^{2k}\{\|f-q_k\|-E_k(f)\}\tag{41}$$

if $a=0$ *and*

$$\|p_k(f)-q_k\|\leqq\left(\frac{c_{31}}{k}\right)^{2k}N^{2k}|\ln E_k(f)|^{2ka}\{\|f-q_k\|-E_k(f)\}\tag{42}$$

if $a>0$ *and* $0<E_k(f)\leqq e^{-a}$.

II) *For any* $f_1\in C[-\pi,\pi]$

$$\|p_k(f)-p_k(f_1)\|\leqq\left(\frac{c_{17}}{k}\right)^{2k}N^{2k}\|f-f_1\|\tag{43}$$

if $a=0$ *and*

$$\|p_k(f)-p_k(f_1)\|\leqq\left(\frac{c_{32}}{k}\right)^{2k}N^{2k}|\ln E_k(f)|^{2ka}\|f-f_1\|\tag{44}$$

if $a>0$ *and* $0<E_k(f)\leqq e^{-a}$.

REMARK. Inequalities (41) and (43) were stated under the conditions that $f\in C^{\infty}_{a,N}\setminus T_k$ and $a=0$, and this is equivalent with $f\in T_N\setminus T_k$ $(N>k)$. Hence we assume $N>k$ if $a=0$ in the statement of Lemma 6.

PROOF. Let us prove (41). If $f\in C^{\infty}_{a,N}\setminus T_k$ then $f\in W_{r-1}\,\mathrm{Lip}_{\|f^r\|}\,1\setminus T_k$ for any $r\geqq 1$. Thus by (22) for any $q_k\in T_k$ and $r\geqq 1$

$$\|p_k(f)-q_k\|\leqq\left(\frac{c_{14}}{k}\right)^{2k}\left(\frac{\|f^{(r)}\|}{E_k(f)}\right)^{2k/r}\{\|f-q_k\|-E_k(f)\}. \tag{45}$$

If $a=0$ then $f\in T_N\setminus T_k$, hence $\|f^{(r)}\|\leqq N^r\|f\|$. Then (45) implies

$$\|p_k(f)-q_k\|\leqq\left(\frac{c_{14}}{k}\right)^{2k}N^{2k}\left(\frac{\|f\|}{E_k(f)}\right)^{2k/r}\{\|f-q_k\|-E_k(f)\}.$$

But this inequality is valid for any $r\geqq 1$ therefore it implies (41).

Assume now that $a>0$ and $0<E_k(f)\leqq e^{-a}$. Using that $f\in C^{\infty}_{a,N}\setminus T_k$ and (45) we obtain

$$\|p_k(f)-q_k\|\leqq\left(\frac{c_{14}}{k}\right)^{2k}N^{2k}\frac{r^{2ka}}{E_k^{2k/r}(f)}\{\|f-q_k\|-E_k(f)\}. \tag{46}$$

(46) holds for any $r\geqq 1$, $r^{2ka}/E_k^{2k/r}(f)$ attains its minimum if $r=-\dfrac{\ln E_k(f)}{a}=$ $=\dfrac{|\ln E_k(f)|}{a}$ hence it is advisable to set $r=\left[\dfrac{|\ln E_k(f)|}{a}\right]+1$. Then we get by easy calculation

$$\frac{r^{2ka}}{E_k^{2k/r}(f)}\leqq c_{33}^{2k}|\ln E_k(f)|^{2ka}$$

and combining this inequality with (46) we obtain (42).

Using (23) we can analogously obtain (43) and (44). Lemma 6 is proved.

In the proof of Theorem 2 we shall need an estimation on the order of approximation of functions from $C^{\infty}_{a,N}$.

LEMMA 7. *Let* $f\in C^{\infty}_{a,N}$, $a>0$, $N>1$. *Then*

$$E_k(f)\leqq\begin{cases}\dfrac{\pi}{2}e^a e^{-(k/N)^{1/a}a/e}, & \textit{if}\quad k>Ne^a;\\[2mm] \dfrac{\pi}{2}\dfrac{N}{k}, & \textit{if}\quad 1\leqq k\leqq Ne^a.\end{cases} \tag{47}$$

PROOF. By Favard's inequality

$$E_k(f)\leqq M_r\frac{\|f^{(r)}\|}{k^r}\leqq\frac{\pi}{2}\frac{N^r r^{ar}}{k^r}. \tag{48}$$

Let us minimize the right hand side here. If $k>Ne^a$ then set $r=\left[1/e\left(\frac{N}{k}\right)^{1/a}\right]+1$. Then by (48) we get

$$E_k(f) \leqq \frac{\pi}{2}\left(\frac{N}{k} r^a\right)^r \leqq \frac{\pi}{2} e^a e^{-(k/N)^{1/a} a/e}$$

Assume now that $1 \leqq k \leqq Ne^a$. Then $\frac{N^r r^{ar}}{k^r}$ is monotonically increasing with r, hence setting $r=1$ in (48) we obtain the desired inequality. Q. e. d.

The proof of Theorem 2

We shall prove Theorem 2 for $j=2$. The estimations for $\Omega_k^{(1)}(C_{a,N}^\infty, \varepsilon)$ can be obtained analogously.

Take $f \in C_{a,N}^\infty$, $\varepsilon>0$.

Case 1. $a=0$, $k \geqq N$. Then $f \in T_k$, hence for any $q_k \in T_k$ satisfying $\|f-q_k\| \leqq E_k(f)+\varepsilon=\varepsilon$ we have

$$\|p_k(f)-q_k\| = \|f-q_k\| \leqq \varepsilon. \tag{49}$$

Case 2. $a=0$, $1 \leqq k < N$. Then $f \in T_N$. Further, if $f \in T_k$ we have (49) and if $f \in T_N \setminus T_k$ then for any $q_k \in T_k$ satisfying $\|f-q_k\| \leqq E_k(f)+\varepsilon$ we obtain by (41)

$$\|p_k(f)-q_k\| \leqq \left(\frac{c_{14}}{k}\right)^{2k} N^{2k} \varepsilon. \tag{50}$$

Of course we may assume that $1<c_{14}$, then $\left(\frac{c_{14}}{k}\right)^{2k} N^{2k} > c_{14}^{2k} > 1$ if $k<N$. Hence (50) is more general than (49) thus (50) fulfils for any $f \in T_N$.

By a theorem proved in [9], if K is a cone then $\Omega_k^{(j)}(K, \varepsilon) \equiv C\varepsilon$ or ∞. Using this theorem and inequalities (49) and (50) we obtain the statement B of the Theorem.

Case 3. $a>0$, $k \geqq k_0 = 1+\left[N\left\{\frac{e}{a}\left(\ln\frac{\pi}{2}+2a\right)\right\}^a\right]$, $0<\varepsilon<\frac{\pi}{2} e^a e^{-(k/N)^{1/a} a/e}$. Then $k>Ne^a$ and by (47)

$$E_k(f) \leqq \frac{\pi}{2} e^a e^{-(k/N)^{1/a} a/e} \leqq \frac{\pi}{2} e^a e^{-(k_0/N)^{1/a} a/e} \leqq e^{-a}. \tag{51}$$

Let $f \in C_{a,N}^\infty \setminus T_k$ and $q_k \in T_k$ satisfy $\|f-q_k\| \leqq E_k(f)+\varepsilon$. Then by (51) and (42) we obtain

$$\|p_k(f)-q_k\| \leqq \left(\frac{c_{31}}{k}\right)^{2k} N^{2k} |\ln E_k(f)|^{2ka} \varepsilon. \tag{52}$$

On the other hand we obviously have

$$\|p_k(f)-q_k\| \leqq 2E_k(f)+\varepsilon \tag{53}$$

and combining (52) and (53) we obtain

$$\|p_k(f)-q_k\| \leqq 2 \sup_{E>0} \min \left\{E+\varepsilon, \left(\frac{c_{31}}{k}\right)^{2k} N^{2k} |\ln E|^{2ka} \varepsilon\right\} = 2(E^*+\varepsilon) \tag{54}$$

where $E^*=E^*(\varepsilon)$ is the unique solution of the equation

$$E^*+\varepsilon = \left(\frac{c_{31}}{k}\right)^{2k} N^{2k} |\ln E^*|^{2ka} \varepsilon.$$

Let E be the unique solution of the equation

$$E = \left(\frac{c_{31}}{k}\right)^{2k} N^{2k} |\ln E|^{2ka} \varepsilon, \tag{55}$$

then evidently

$$E^* < E. \tag{56}$$

Further, ε tends to zero faster than $E(\varepsilon)$ $(\varepsilon \to +0)$, and the equation $E(\varepsilon)=\varepsilon$ has the unique solution

$$\varepsilon_0 = e^{-(k/c_{31}N)^{1/a}}. \tag{57}$$

Therefore $\varepsilon < E(\varepsilon)$ if $0<\varepsilon<\varepsilon_0$. It can be shown by easy calculations that if we choose the constant c_{31} larger than an absolute constant c_{32} then ε_0 defined by (57) is larger than $\frac{\pi}{2} e^a e^{-(k/N)^{1/a} a/e}$. But inequality (52) (where c_{31} appears) remains valid for larger constants also, i.e. we may assume that the solution $E(\varepsilon)$ of (55) is larger than ε for any $0<\varepsilon<\frac{\pi}{2} e^a e^{-(k/N)^{1/a} a/e}$. This together with (54), (55) and (56) implies

$$\begin{aligned} \|p_k(f)-q_k\| &\leqq 2(E^*+\varepsilon) \leqq 2(E+\varepsilon) \leqq 4E(\varepsilon) = \\ &= 4\left(\frac{c_{31}}{k}\right)^{2k} N^{2k} |\ln E|^{2ka} \varepsilon \leqq \left(\frac{c_{33}}{k}\right)^{2k} N^{2k} |\ln \varepsilon|^{2ka} \varepsilon. \end{aligned} \tag{58}$$

If $f \in T_k$, then $\|p_k(f)-q_k\| \leqq \varepsilon$ and this inequality is contained in (58). Hence we verified (58) in *Case 3*.

Case 4. $a>0$, $k \geqq k_0$, $\varepsilon \geqq \frac{\pi}{2} e^a e^{-(k/N)^{1/a} a/e}$. If $k \geqq k_0$ then $k > Ne^a$ and by (47)

$$E_k(f) \leqq \frac{\pi}{2} e^a e^{-(k/N)^{1/a} a/e} \leqq \varepsilon.$$

Thus using this inequality and (53) we obtain in this case that for any $q_k \in T_k$ satisfying $\|f-q_k\| \leqq E_k(f)+\varepsilon$,

$$\|p_k(f)-q_k\| \leqq 3\varepsilon.$$

This together with (58) completes the proof of statement A_1 of the Theorem.

Case 5. $a>0$, $1\leqq k<k_0$. Let $q_k\in T_k$ satisfy $\|f-q_k\|\leqq E_k(f)+\varepsilon$. If $f\in T_k$, then

(59) $$\|p_k(f)-q_k\| = \|f-q_k\| \leqq \varepsilon.$$

Let $f\in C^\infty_{a,N}\setminus T_k$. Then if $E_k(f)<e^{-a}$ and $0<\varepsilon<\frac{\pi}{2}e^a e^{-(k/N)^{1/a}a/e}$ we obtain (58) similarly to Case 3. But for $k<k_0$, $\frac{\pi}{2}e^a e^{-(k/N)^{1/a}a/e}\geqq c_{34}(a)$ i.e. if $0<E_k(f)<e^{-a}$ then for any $0<\varepsilon<c_{34}(a)$ (58) holds. On the other hand, if $E_k(f)\geqq e^{-a}$, then setting $r=1$ in (45) we obtain

(60) $$\|p_k(f)-q_k\| \leqq \left(\frac{c_{14}}{k}\right)^{2k}\left(\frac{\|f'\|}{E_k(f)}\right)^{2k}\varepsilon \leqq \left(\frac{c_{14}}{k}\right)^{2k} N^{2k}e^{2ka}\varepsilon \leqq \left(\frac{c_{35}(a)}{k}\right)^{2k} N^{2k}\varepsilon.$$

Now we have to compare this inequality with (58) and (59). It can be easily shown that in this case by choosing an appropriate constant $c_{36}(a)$ and combining (58), (59) and (60) we have for any $0<\varepsilon\leqq c_{36}(a)$

$$\|p_k(f)-q_k\| \leqq \left(\frac{c_{37}(a)}{k}\right)^{2k} N^{2k}|\ln\varepsilon|^{2ka}\varepsilon.$$

I.e. we obtained statement A_2 of the Theorem hence its proof is complete.

During our attempts to verify the sharpness of the estimation of Theorem 2 we met some technical difficulties connected with calculation of parameters a and N for $f\in C^\infty[-\pi,\pi]$. Now we shall give lower estimations for $\Omega_k^{(j)}(C^\infty_{a,N},\varepsilon)$ $(j=1,2)$ when $a>1$.

Take $f(x)=e^{-[x(h-x)]^{-b}}$, where $b>0$, $x\in[0,h]$, $0<h<1$. Then for some a and N $f\in C^\infty_{a,N}$; in fact $f\in C^\infty_{1+1/b,N}$. In order to prove this, first of all we must calculate the derivatives of f. The Cauchy's theorem gives a representation

(61) $$f^{(r)}(x) = \frac{r!}{2\pi i}\oint_C \frac{f(z)dz}{(z-x)^{r+1}}$$

where C is a closed curve on the complex plane containing x in its interior. We shall use this representation together with the following

LEMMA 8. *For any* $r\geqq 1$, $x\in[0,h]$

(62) $$f^{(r)}(x) = f(x)\sum_{l=1}^{r}\frac{P_{r,l}(x)}{[x(h-x)]^{lb+r}}$$

where $P_{r,l}(x)$ $(1\leqq l\leqq r)$ *are algebraic polynomials of degree at most* r, $P_{1,1}=$ $=b(h-2x)$ *and*

(63) $$P_{r+1,l} = P'_{r,l}x(h-x)-(lb+r)(h-2x)P_{r,l}+b(h-2x)P_{r,l-1}$$

$(1\leqq l\leqq r+1)$ *where* $P_{r,0}\equiv P_{r,r+1}\equiv 0$.

Moreover, for any $1\leqq l\leqq r$

(64) $$r^{-(r-l)}\|P_{r,l}(x)\|_{[0,h]} \leqq (2b+2)^r \quad (r=1,2,\dots).$$

PROOF. (62)—(63) is easily seen using induction on r.

(64) whill be proved also by induction. $\|P_{1,1}\|_{[0,h]} \leqq b \leqq 2b+2$. Assume that (64) holds for $r=m$ and verify it for $r=m+1$. Using (63) and Bernstein's inequality we have for any $x \in [0, h]$, $1 \leqq l \leqq m+1$

$$|P_{m+1,l}| \leqq |P'_{m,l}|\,|x(h-x)| + [(m+1)b+m]\,|P_{m,l}| + b\,|P_{m,l-1}| \leqq$$

$$\leqq m\,\|P_{m,l}\|_{[0,h]} \sqrt{x(h-x)} + [(m+1)b+m]\,\|P_{m,l}\|_{[0,h]} + b\,\|P_{m,l-1}\|_{[0,h]} \leqq$$

$$\leqq [(m+1)b+2m]\,\|P_{m,l}\|_{[0,h]} + b\,\|P_{m,l-1}\|_{[0,h]}.$$

Hence and from (64) for $r=m$ we obtain for any $1 \leqq l \leqq m+1$

$$\frac{1}{(m+1)^{m+1-l}}\,\|P_{m+1,l}(x)\|_{[0,h]} \leqq \frac{(m+1)b+2m}{m+1}\,\frac{\|P_{m,l}\|_{[0,h]}}{(m+1)^{m-l}} +$$

$$+\frac{b\,\|P_{m,l-1}\|_{[0,h]}}{(m+1)^{m+1-l}} \leqq (b+2)\,\frac{\|P_{m,l}\|_{[0,h]}}{m^{m-l}} + b\,\frac{\|P_{m,l-1}\|_{[0,h]}}{m^{m-(l-1)}} \leqq$$

$$\leqq (b+2)(2b+2)^m + b(2b+2)^m = (2b+2)^{m+1}$$

and this implies (64). Q. e. d.

LEMMA 9. *For any* $r \geqq 1$, $x \in [0, h]$

$$|f^{(r)}(x)| \leqq N^r r^{(1+1/b)r} \tag{65}$$

with some $N = N(b)$.

PROOF. Set $I_r = [0, r^{-1/b}] \cup [h - r^{-1/b}, h]$;

$$C_r = \left\{\frac{r^{-1/b}}{2} + iy,\ |y| \leqq \frac{r^{-1/b}}{4b}\right\} \cup \left\{x + i\,\frac{r^{-1/b}}{4b},\ \frac{r^{-1/b}}{2} \leqq x \leqq h - \frac{r^{-1/b}}{2}\right\} \cup$$

$$\cup \left\{h - \frac{r^{-1/b}}{2} + iy,\ |y| \leqq \frac{r^{-1/b}}{4b}\right\} \cup \left\{x - i\,\frac{r^{-1/b}}{4b},\ \frac{r^{-1/b}}{2} \leqq x \leqq h - \frac{r^{-1/b}}{2}\right\}.$$

CASE A. $x \in [0, h] \setminus I_r$. Then by (61)

$$f^{(r)}(x) = \frac{r!}{2\pi i} \oint_{C_r} \frac{f(z)\,dz}{(z-x)^{r+1}}. \tag{66}$$

Evidently for $z \in C_r$ $|\arg z| \leqq \operatorname{arctg}\frac{1}{2b}$, $|\arg(h-z)| \leqq \operatorname{arctg}\frac{1}{2b}$, hence

$$|\arg[z(h-z)]^{-b}| = b\,|\arg z + \arg(h-z)| \leqq 2b \operatorname{arctg}\frac{1}{2b} \leqq 2b\,\frac{\pi}{2}\,\frac{1}{2b} = \frac{\pi}{2}.$$

It means that $\operatorname{Re}[z(h-z)]^{-b}\geqq 0$ and $|f(z)|\leqq 1$ for $z\in C_r$. Therefore for any $x\in[0,h]\diagdown I_r$

$$|f^{(r)}(x)|\leqq\frac{r!}{2\pi}\left(2+\frac{1}{b}\right)\frac{1}{\left[r^{-1/b}\min\left\{\frac{1}{2},\frac{1}{4b}\right\}\right]^{r+1}}\leqq c_{38}^r(b)\,r^{r+(r+1)/b}\leqq c_{39}^r\,r^{r(1+1/b)}$$

i.e. (65) holds in this case.

CASE B. $x\in I_r$. Then by (62) and (64) we have

$$|f^{(r)}(x)|=\frac{|f(x)|}{|x(h-x)|^{rb+r}}\left|\sum_{l=1}^{r}[x(h-x)]^{(r-l)b}P_{r,l}(x)\right|\leqq$$

$$\leqq\left\|\frac{f}{x(h-x)^{rb+r}}\right\|_{[0,h]}\sum_{l=1}^{r}\frac{1}{r^{r-l}}\|P_{r,l}\|_{[0,h]}\leqq r(2b+2)^r\|u^{rb+r}e^{-u^b}\|_{[0,\infty]}\leqq c_{40}^r(b)\,r^{r(1+1/b)}.$$

Q. e. d.

Now extend $f(x)$ to the interval $[-(k+1)h,(k+1)h]$ as odd $2h$-periodic function and assume that h is so small that $[-(k+1)h,(k+1)h]\subset\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$. Then consider the function

$$\bar f(x)=\begin{cases}f(x),\ x\in[-(k+1)h,(k+1)h];\\ 0,\quad\text{otherwise.}\end{cases}$$

Then evidently $p_k(\bar f)\equiv 0$, $\bar f\in C^{\infty}_{1+1/b,N}$, $\|\bar f\|=e^{-(2/h)^{2b}}$. Further let $\bar x$ be the greatest isolated zero of $\bar f(x)$ and set

$$g(x)=\begin{cases}\bar f(x), & -\pi\leqq x\leqq\bar x;\\ \dfrac{\|\bar f\|+\varepsilon}{\|\bar f\|}\bar f(x), & \bar x\leqq x\leqq\pi.\end{cases}$$

Then $\|\bar f-g\|=\varepsilon$ and similarly to the proof of the lower estimation of Theorem 1 we can verify that

$$\|p_k(\bar f)-p_k(g)\|\geqq c_{41}^k\min\left\{\frac{\varepsilon}{k^{2k}h^{2k}},\|\bar f\|\right\}\tag{67}$$

holds for h small enough. Let E be the unique solution of the equation $E=\frac{\varepsilon\,|\ln E|^{k/b}}{k^{2k}}$, $0<E<1$ for ε small enough, and set $h=\frac{2}{|\ln E|^{1/2b}}$. Then $\|\bar f\|=$ $=e^{-(2/h)^{2b}}=E$; hence and from (67)

$$\|p_k(\bar f)-p_k(g)\|\geqq c_{41}^k\min\left\{\frac{E}{2^{2k}},E\right\}=c_{42}^{2k}E=\tag{68}$$

$$=\left(\frac{c_{42}}{k}\right)^{2k}\varepsilon\,|\ln E|^{k/b}\geqq\left(\frac{c_{43}(b)}{k}\right)^{2k}\varepsilon\,|\ln\varepsilon|^{k/b}$$

for ε small enough. Set $a=1+\frac{1}{b}$, then $\bar{f}\in C^{\infty}_{a,N}$ and by (68)

$$\Omega_k^{(1)}(C^{\infty}_{a,N},\varepsilon)\geqq\left(\frac{c_{43}(a)}{k}\right)^{2k}\varepsilon\,|\ln\varepsilon|^{k(a-1)}$$

i.e. we obtained a lower estimation for $\Omega_k^{(1)}(C^{\infty}_{a,N},\varepsilon)$ if $a>1$. $\Omega_k^{(2)}(C^{\infty}_{a,N},\varepsilon)$ can be estimated from below similarly. These lower estimations for $a>1$ are rather close to the upper estimations given by Theorem 2 but nevertheless they are not the sharpest.

The main idea of getting lower estimations for $\Omega_k^{(j)}$ was based on constructing functions which had all their points of Čebysev alternation condensed on a very small interval. We achieved this by setting these functions zero on the whole period except a small interval. But functions of this type cannot be analytic, hence this method is not appropriate for getting lower estimations for $\Omega_k^{(j)}$ in the case $0<a\leqq 1$.

The author is indebted to J. Szabados for his useful remarks.

References

[1] G. Freud, Eine Ungleichung für Tschebyscheffsche Approximationspolynome, *Acta Sci. Math. Szeged,* **19** (1958), 162—164.

[2] D. J. Newman, H. S. Shapiro, Some theorems on Čebysev approximation, *Duke Math. J.,* **30** (1963), 673—681.

[3] R. Holmes, B. Kripke, Smoothness of approximation, *Mich. Math. J.,* **15** (1968), 225—248.

[4] B. O. Björnestål, *The preprint series of Department of Mathematics,* Royal Institute of Technology, Stockholm, TRITA—MAT, **17** (1974), **20** (1974), **12** (1975).

[5] П. В. Галкин, О модуле непрерывности оператора наилучшего приближения в пространстве непрерывных функций, *Матем. заметки,* **10** (1971), 601—613.

[6] А. Кроо, Задача корректности наилучших приближений тригонометрическими полиномами класса $W_0^r H[\omega]_c$, *Матем. заметки,* **22** (1977), 85—101.

[7] А. В. Колушов, Задача корректности наилучшего приближения в пространстве непрерывных функций, *Матем. заметки,* **23** (1978), 351—361.

[8] A. Kroó, On the continuity of best approximations in the space of integrable functions, *Acta Math. Ac. Sci. Hungar.,* **32** (1978), 331—348.

[9] A. Kroó, The continuity of best approximations, *Acta Math. Acad. Sci. Hungar.,* **30** (1977), 175—188.

[10] С. Б. Стечкин, О порядке наилучших приближений непрерывных функций, *Изв. АН СССР,* **15** (1951), 219—242.

[11] А. В. Ефимов, Линейные методы приближения непрерывных периодических функций, *Матем. сб.,* **54** (1960), 51—90.

[12] A. Kroó, A comparison of uniform and discrete polynomial approximation, *Analysis Math.,* **5** (1979), 35—49.

[13] В. Н. Малоземов, *Совместное приближение функции и её производных,* ЛГУ (1973).

(Received November 11, 1977)

MATHEMATICAL INSTITUTE OF THE
HUNGARIAN ACADEMY OF SCIENCES
1053 BUDAPEST, REÁLTANODA U. 13—15.

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (1—2), (1979), 205—221.

ON THE ϱ-SYSTEMS OF CIRCLES

By

A. FLORIAN (Salzburg), L. HÁRS (Budapest) and J. MOLNÁR (Budapest)

Let $\{C_i\}$ be a packing of cirles in the Euclidean plane. A circle C is said to be a *supporting circle* of the circle system $\{C_i\}$ if it has no common interior point with $\{C_i\}$ and touches at least three circles of $\{C_i\}$. If ϱ is the greatest lower bound of the radii ϱ^* of all supporting circles of $\{C_i\}$ and if $\varrho=\inf \varrho^*>0$, then $\{C_i\}$ is called a *ϱ-system of circles.*

The *density of a circle system* $\{C_i\}$ *with respect to the Euclidean plane* is defined by

$$\delta = \overline{\lim_{R\to\infty}} \frac{\sum\limits_i (C_i \cap C(R))}{C(R)}\ ^1$$

where $C(R)$ is a circle of radius R centred at a fixed point O of the plane.[2]

Subsequent to the investigations of MOLNÁR ([10], [11]), concerning ϱ-systems of circles, we prove the following

THEOREM.[3] *If d denotes the density of a packing in the Euclidean plane by a ϱ-system of circles of radii contained in the interval* $[\varepsilon, 1]$, *where* $\varepsilon>0$, *then*

$$d \leqq \frac{\arccos \dfrac{1}{1+\varrho}}{\sqrt{2\varrho+\varrho^2}}.$$

Equality holds if $\varrho=\dfrac{2\sqrt{3}}{3}-1$, $\sqrt{2}-1$ *and* 1 *and the ϱ-system consists only of unit circles.*[4]

Consider three circles of radii 1, 1, ϱ and centres A, B, C mutually touching one another (Fig. 4). Then $d(\varrho)=\dfrac{\arccos \dfrac{1}{1+\varrho}}{\sqrt{2\varrho+\varrho^2}}$ is the density of the unit circles in the triangle ABC, namely the ratio of the area of the part of the triangle ABC covered by the unit circles to the area of the whole triangle.

[1] We denote a domain and its area by the same symbol.

[2] It is easy to see that δ does not depend on the choice of O; see FEJES TÓTH [1].

[3] Attention should also be drawn to the quadrilateral tessellation and the lemmas employed in the proof of this theorem which may be useful for future density investigations. Lemmas 5 and 6 are due to Hárs and Florian respectively, the remaining part of the article is the work of Molnár.

[4] See Fig. 1, 2 and 3.

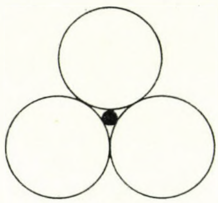


Fig. 1

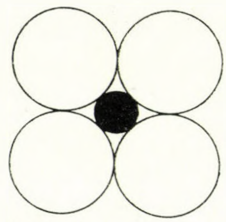


Fig. 2

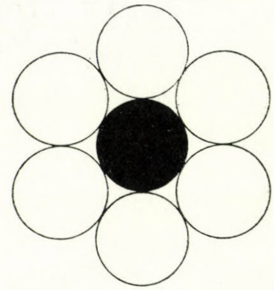


Fig. 3

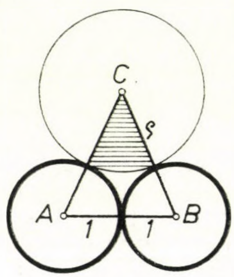


Fig. 4

Without loss of generality we may suppose that the packing of the ϱ-system of circles is saturated. We shall construct a tessellation with quadrilateral faces, the vertices of which are alternatively centres $O_1, O_2, \ldots$ of the circles $C_1, C_2, \ldots$ and centres $V_1, V_2, \ldots$ of the supporting circles of $\{C_i\}$. In order to prove our assertion we shall show that in each quadrangle of the tessellation the density of $\{C_i\}$ does not exceed $d(\varrho)$.

We introduce the notion of the (algebraic) *distance* $d(P, C)=\overline{OP}-r$ of a point P from a circle C of radius r centred at O. Let us associate with any circle C_i the set S_i of all points P lying "nearer" to C_i than to any other circle C_j, i.e. $d(P, C_i)<d(P, C_j)$ $(j\neq i)$.[5] It is not difficult to show that S_i is a star region with respect to the pole O_i (Fig. 5). The star regions $\{S_i\}$ are bounded by arcs of hyperbolae and segments of straight lines.

Obviously the star regions $S_1, S_2, \ldots$ constitute a tessellation S. Joining the centre O_i $(i=1, 2, \ldots)$ with the vertices $V_1, V_2, \ldots$ of the corresponding star region S_i, we obtain a new tessellation T with quadrilateral faces (Fig. 6).

We proceed to show that in each quadrilateral face (quadrangle) of T the density of $\{C_i\}$ is $\leqq d(\varrho)$.

To prove this statement we need a certain number of lemmas.

[5] See Fejes Tóth—Molnár [2].

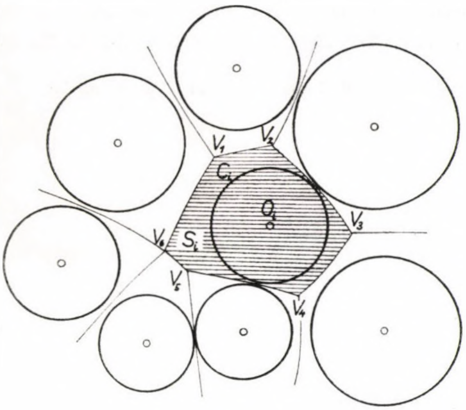


Fig. 5

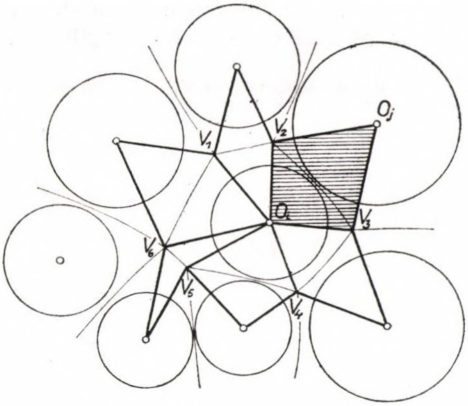


Fig. 6

LEMMA 1. *Let AOB be a triangle of $\sphericalangle OAB \geqq \frac{\pi}{2}$. If M is the midpoint of the side AB then $\sphericalangle AOM > \sphericalangle BOM$.*

PROOF. The condition $\sphericalangle OAB \geqq \frac{\pi}{2}$ implies $\overline{OB} > \overline{OA}$ (Fig. 7). Let O^* be the mirror point of O with respect to M. Considering the triangle OO^*B, we have $\sphericalangle BOM < \sphericalangle BO^*M = \sphericalangle AOM$ in consequence of $\overline{O^*B} = \overline{OA} < \overline{OB}$.

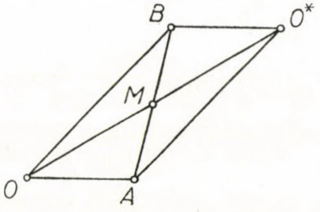


Fig. 7

Let O_1AO_2B be a simple quadrangle and let C_1, C_2 be two circles of centres O_1, O_2. Denote by $C_1(AO_1B)$ and $C_2(AO_2B)$ the sectors of the circles C_1 and C_2 corresponding to the angles $\sphericalangle AO_1B$ and $\sphericalangle AO_2B$ of O_1AO_2B. We define the *density of the circles C_1, C_2 with respect to the quadrangle O_1AO_2B* by

$$d_{12}(AB) = \frac{C_1(AO_1B) + C_2(AO_2B)}{O_1AO_2B}. \tag{1}$$

LEMMA 2. *In the Euclidean plane, consider two circles C_1, C_2 of centres O_1, O_2 and of radii r_1, r_2 ($r_1 < r_2$), resp. Let A, B be two different points, both at the same distance from C_1 and C_2 and on the same side of the straight line O_1O_2. If $\sphericalangle BAO_1 \geqq \frac{\pi}{2}$ (hence $\overline{O_1A} < \overline{O_1B}$), then for any point P of the segment AB we have $d_{12}(AP) \geqq d_{12}(PB)$.*

PROOF.[6] We first remark that A, B are points on the same branch of the hyperbola H of foci O_1, O_2, the length of the transverse axis is $r_2 - r_1$. Since the line AB is a secant of H and $\sphericalangle BAO_1 \geqq \frac{\pi}{2}$ we have also $\sphericalangle BAO_2 \geqq \frac{\pi}{2}$ (Fig. 8).

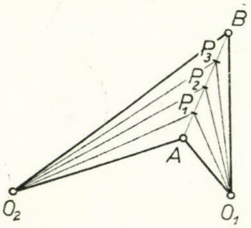


Fig. 8

Let $P_1, P_2, \dots, P_n$ be equidistant points on AB, i.e. $\overline{AP_1} = \overline{P_1P_2} = \dots = \overline{P_{n-1}P_n} = \overline{P_nB}$. In view of Lemma 1, the angles $\sphericalangle AO_1P_1 = \alpha_1$, $\sphericalangle P_1O_1P_2 = \alpha_2, \dots$, $\sphericalangle P_nO_1B = \alpha_{n+1}$ and $\sphericalangle AO_2P_1 = \beta_1$, $\sphericalangle P_1O_2P_2 = \beta_2, \dots$, $\sphericalangle P_nO_2B = \beta_{n+1}$ form two decreasing sequences. On the other hand, the quadrangles $O_1AO_2P_1$, $O_1P_1P_2P_2$, $\dots$, $O_1P_nO_2B$ have all the same area. Therefore, employing the notation introduced in (1), we see that the sequence $d_{12}(AP_1), d_{12}(P_1P_2), \dots, d_{12}(P_nB)$ decreases monotonically. Consequently we get

$$d_{12}(AP_i) \geqq d_{12}(P_{i-1}P_i) > d_{12}(P_iP_{i+1}) \geqq d_{12}(P_iB) \quad (i = 1, \dots, n).$$

But the inequality $d_{12}(AP_i) > d_{12}(P_iB)$ is true for any n and $i = 1, \dots, n$. This concludes the proof of Lemma 2.

Obviously, $d_{12}(AP) \geqq d_{12}(PB)$ implies $d_{12}(AP) \geqq d_{12}(AB)$.

LEMMA 3. *Let H be a hyperbola branch and F the focus lying in the convex domain bounded by H. Let us denote by H^* one of the half branches of H determined by the transverse axis of H. The circle of diameter FP, where P is a point of H^*, has at most one further common point with H^*.*

PROOF. Let $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ be the equation of the hyperbola H and let H^* be the half branch of H lying in the first quadrant of the coordinate system (Fig. 9). Let $F(c, 0)$ be the corresponding focus of H and $P(\lambda, \mu)$ a point of H^*. The equation of the circle C with diameter FP is

$$x^2 + y^2 - (\lambda + c)x - \mu y + \lambda c = 0.$$

The abscissae of the common points of C and H satisfy the equation

$$f(x) \equiv \frac{c^2}{a^2}x^2 - (\lambda + c)x - \frac{a}{b}\mu\sqrt{x^2 - a^2} + \lambda c - b^2 = 0.$$

[6] A different proof which does not make use of Lemma 1, was given later by A. Florian.

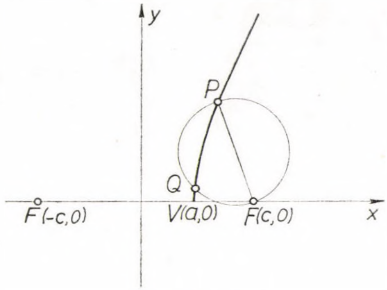


Fig. 9

Obviously, $f(x)$ is a strictly convex function for $x \geqq a$ which vanishes at $x=\lambda$. Now we distinguish two cases:

(i) $\lambda > a$ implies

$$f'(\lambda) = \lambda \frac{c^2}{a^2} - c > \frac{c}{a}(c-a) > 0$$

and $f(a)=(\lambda-a)(c-a)>0$. Therefore, the function $f(x)$ has precisely two zeros in the interval $x>a$, the greater of which is λ. Let $V=(a, 0)$ be the vertex of H^*. If $\lambda>a$ and consequently $P \neq V$ then C intersects H^* in exactly two points, namely P and $Q \neq P$, where Q lies between P and V. From this we deduce immediately that any point P^* on the open arc PQ of H^* has the property $\sphericalangle PP^*F > \frac{\pi}{2}$.

(ii) For $\lambda=a$ the function $f(x)$ vanishes only at $x=a$ and $x=\frac{a(ac-b^2)}{c^2}$; but $ac-b^2<c^2$. It follows that, if $P=V$, the circle C and the hyperbola half branch H^* touch each other at V and do not have any other point in common.

LEMMA 4. *In the Euclidean plane, consider two non-overlapping circles C_1, C_2 of centres O_1, O_2 and radii r_1, r_2 $(r_1<r_2)$. Let A, B be two different points, both equidistant from C_1 and C_2 and on the same side of the straight line O_1O_2. Let H be that branch of the hyperbola of foci O_1, O_2 having the length r_2-r_1 of the transverse axis which contains A, B. If $\overline{O_1A}<\overline{O_1B}$, then for any interior point P of the arc AB of H we have, using the notation* (1),

$$d_{12}(AP) \geqq d_{12}(PB). \tag{2}$$

PROOF. It suffices to prove the lemma under the assumption that A is not the vertex V of H, carrying out the limiting process $A \to V$ in the other case.

We can find, on the basis of Lemma 3, on the open arc AB of H a sequence of points $P_1, P_2, \dots, P_n$, so that the angles $\sphericalangle P_1AO_1$, $\sphericalangle P_2P_1O_1, \dots, \sphericalangle BP_nO_1$ are obtuse. A sequence of this property we call admissible. Since the tangent at any point P of H is the bisecting line of $\sphericalangle O_1PO_2$, the angles $\sphericalangle P_1AO_2$, $\sphericalangle P_2P_1O_2$, $\dots, BP_nO_2$ are obtuse too (Fig. 10).

We shall first prove the inequality

$$\lim_{M\to P_i} d_{12}(MP_i) > \lim_{N\to P_i} d_{12}(P_iN) \quad (i=1, \dots, n)$$

where M and N are pointson the segments $P_{i-1}P_i$ and P_iP_{i+1} respectively (Fig. 11).

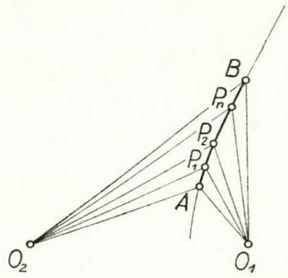


Fig. 10

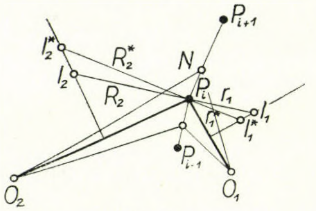


Fig. 11

Write $\sphericalangle MO_1P_i=\varepsilon_1$, $\sphericalangle MO_2P_i=\varepsilon_2$, $\sphericalangle NO_1P_i=\varepsilon_1^*$, $\sphericalangle NO_2P_i=\varepsilon_2^*$ and denote by I_1, I_1^* and I_2, I_2^* the intersections of the perpendicular bisector of the segment O_1P_i and O_2P_i respectively, with the straight lines perpendicular to $P_{i-1}P_i$ and P_iP_{i+1} at P_i.

Taking into account that $\overline{P_iI_1}=R_1>\overline{P_iI_1^*}=R_1^*$, $\overline{P_iI_2}=R_2<\overline{P_iI_2^*}=R_2^*$ and that

$$\frac{r_2^2}{r_1^2} > \frac{\overline{O_2P_i}^2}{\overline{O_1P_i}^2} = \frac{(r_2+\varrho)^2}{(r_1+\varrho)^2}$$

it is easy to see that

$$\lim_{M\to P_i} d_{12}(MP_i) \geqq \lim_{\varepsilon_1\to 0} \frac{\dfrac{\varepsilon_1}{\varepsilon_2}r_1^2+r_2^2}{\dfrac{\varepsilon_1}{\varepsilon_2}\overline{O_1P_i}^2+\overline{O_2P_i}^2} = \frac{\dfrac{R_2}{R_1}r_1^2+r_2^2}{\dfrac{R_2}{R_1}\overline{O_1P_i}^2+\overline{O_2P_i}^2} >$$

$$> \frac{\dfrac{R_2^*}{R_1^*}r_1^2+r_2^2}{\dfrac{R_2^*}{R_1^*}\overline{O_1P_i}^2+\overline{O_2P_i}^2} = \lim_{\varepsilon_1^*\to 0} \frac{\dfrac{\varepsilon_1^*}{\varepsilon_2^*}r_1^2+r_2^2}{\dfrac{\varepsilon_1^*}{\varepsilon_2^*}\overline{O_1P_i}^2+\overline{O_2P_i}^2} \geqq \lim_{N\to P_i} d_{12}(P_iN).$$

Therefore, and in view of Lemma 2, we obtain

$$d_{12}(AP_1) > d_{12}(P_1P_2) > \dots > d_{12}(P_nB),$$

whence

$$d_{12}(AP_i) \geqq d_{12}(P_{i-1}P_i) > d_{12}(P_iP_{i+1}) \geqq d_{12}(P_iB) \quad (i=1, \dots, n).$$

Let P be an arbitrary interpolating point on the open arc $P_{i-1}P_i$ of H. Recalling the property $\sphericalangle P_iP_{i-1}O_1>\frac{\pi}{2}$, we note that $\sphericalangle PP_{i-1}Q_1>\frac{\pi}{2}$ and, according to Lemma 3, $\sphericalangle P_iPO_1>\frac{\pi}{2}$. Thus the sequence $P_1, \dots, P_{i-1}, P, P_i, \dots, P_n$ is also admissible and the inequality (2) is shown.

Let ABC be a triangle where the lengths of the sides AC and BC are supposed to be fixed. The notation is chosen so that $\overline{AC} \leqq \overline{BC}$. We draw attention to the density

$$\delta(x) = \frac{\lambda\alpha + \mu\beta + \nu x}{\frac{1}{2}\overline{AC} \cdot \overline{BC} \cdot \sin x}$$

where x indicates the angle enclosed by AC and BC (Fig. 12). Herein λ, μ, ν denote non-negative constants, not all of which are zero. $\delta(x)$ represents the ratio of a weighted sum of the angles to the area of the triangle ABC.

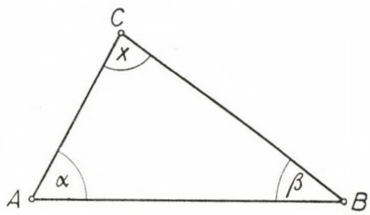


Fig. 12

LEMMA 5.[7] *Let us vary the angle x of the triangle ABC, so that $0 \leqq x \leqq \pi$. Then, in any subinterval of $(0, \pi)$, $\delta(x)$ attains its maximum at one of the endpoints.*

Functions having this property we shall call in the following *quasiconvex*. For the sake of simplicity let us consider, instead of $\delta(x)$, its constant multiple

$$S(x) = \frac{\lambda\alpha + \mu\beta + \nu x}{\sin x}. \tag{3}$$

We remark that $S(x)$ is continuous in $(0, \pi)$.

Making use of the cosine theorem and introducing the notation $\overline{AC}/\overline{BC} = p$ $(p \leqq 1)$, we have

$$\cos\alpha = \frac{p - \cos x}{\sqrt{1 + p^2 - 2p\cos x}}. \tag{4}$$

Differentiation yields

$$\alpha' = \frac{d\alpha}{dx} = \frac{p\cos x - 1}{1 + p^2 - 2p\cos x}, \tag{5}$$

and, in view of $\alpha + \beta + x = \pi$,

$$\beta' = \frac{d\beta}{dx} = -1 - \alpha' = \frac{p\cos x - p^2}{1 + p^2 - 2p\cos x}. \tag{6}$$

[7] See L. HÁRS [4]. In a previous paper [3] A. FLORIAN proved a more special result in a similar way.

Putting, for brevity's sake, $\lambda\alpha+\mu\beta+\nu x=y$, we obtain

$$y'=\frac{-[(\mu-\nu)p^2+\lambda-\nu]+p\cos x[\lambda+\mu-2\nu]}{1+p^2-2p\cos x} \tag{7}$$

and, with the notation

$$A=(\mu-\nu)p^2+\lambda-\nu; \quad B=p(\lambda+\mu-2\nu), \tag{8}$$

$$y'=\frac{-A+B\cos x}{1+p^2-2p\cos x}. \tag{9}$$

Differentiating once again, we have

$$y''=\frac{p(p^2-1)(\mu-\lambda)\sin x}{(1+p^2-2p\cos x)^2} \tag{10}$$

and, owing to (3),

$$S'=\frac{y'\sin x-y\cos x}{\sin^2 x}=\frac{P}{\sin^2 x}. \tag{11}$$

We observe that S' has the same sign as P. The function

$$Q=\frac{y'\sin x-y\cos x}{\cos x}=\frac{P}{\cos x}=y'\operatorname{tg} x-y \tag{12}$$

is continuous on the set $[0, \pi/2)\cup(\pi/2, \pi]$ and has the values

$$Q(0)=-y(0)\leqq 0, \quad Q(\pi)=-y(\pi)\leqq 0. \tag{13}$$

Since $\cos x>0$ in $[0, \pi/2)$ and $\cos x<0$ in $(\pi/2, \pi]$ we can state that

a1) for $x<\pi/2$ S is increasing if $Q>0$,
a2) S is decreasing if $Q<0$,
b1) for $\pi/2<x\leqq\pi$ S is decreasing if $Q>0$,
b2) S is increasing if $Q<0$.

To examine the sign of Q it will be useful to see whether Q is increasing or decreasing in a given interval. For this purpose we shall need its derivative

$$Q'=\frac{p(p^2-1)(\mu-\lambda)\sin^2 x}{(1+p^2-2p\cos x)^2\cos x}+\frac{(-A+B\cos x)\sin^2 x}{(1+p^2-2p\cos x)\cos^2 x}.$$

Since

$$\operatorname{sgn} Q'=\operatorname{sgn}\left[\frac{\cos^2 x}{\sin^2 x}(1+p^2-2p\cos x)^2 Q'\right]$$

in $(0, \pi/2)$ and $(\pi/2, \pi)$ we have to consider the function

$$R(x)=\frac{\cos^2 x}{\sin^2 x}(1+p^2-2p\cos x)^2 Q'=$$

$$=-2pB\cos^2 x+[(p^2-1)(\mu-\lambda)p+2Ap+B(1+p^2)]\cos x-A(1+p^2).$$

But, by (8), the coefficient of $\cos x$ is $4pA$, so that we finally have

$$(14) \qquad R(x) = -2pB\cos^2 x + 4pA\cos x - A(1+p^2).$$

Obviously, $R(x)$ is a polynomial in $\cos x$ of degree $\leqq 2$. Denoting it by F

$$F(z) = -2pBz^2 + 4pAz - A(1+p^2)$$

then $R(x) = F(\cos x)$. The discriminant of F is

$$(15) \qquad D = 8Ap^2(p^2-1)(\mu-\lambda).$$

In proving S to be quasiconvex, we have to distinguish several cases and subcases.

I. $B=0$. If also $A=0$, then from (9) it follows that $y'=0$ and from (11) that $\operatorname{sgn} S' = -\operatorname{sgn}(\cos x)$. Therefore, S is decreasing for $x<\pi/2$ and increasing for $x>\pi/2$, which means that S is quasiconvex.

If, however, $A\neq 0$, then $F(z) = A[4pz-(p^2+1)]$ is a linear polynomial in z having the root $\dfrac{p^2+1}{4p}>0$.

I.1. $A>0$. For $\cos x = z \leqq 0$ we have $R(x)<0$ by (14). More generally, if $R(x)\leqq 0$ for $x>\pi/2$ (Q is decreasing) or $R(x)<0$ for $\pi/2<x<x_1$ and $R(x)>0$ in $x_1<x<\pi$ with any $x_1\in(\pi/2, \pi)$ (Q is decreasing in $(\pi/2, x_1)$ and increasing in (x_1, π)), we shall refer to it as *case* c). Since $Q(\pi)\leqq 0$ by (13), in this case Q is either negative or positive in the whole interval $(\pi/2, \pi)$, or positive in a certain subinterval $(\pi/2, x_0)$ and negative in (x_0, π). Then we can state that:

in the first case (case b2)) S is increasing,

in the second case (case b1)) S is decreasing and

in the third case (case b1) in $(\pi/2, x_0)$ and case b2) in (x_0, π)) S is decreasing in $(\pi/2, x_0)$ and increasing in (x_0, π).

If we can show, moreover, that for $x<\pi/2$ we have $S'(x)\leqq 0$ (this will be supposed to hold in case c)) then S follows to be quasiconvex.

In fact, for $0<x<\pi/2$ $\cos x$ and $\sin x$ are positive, so that, in view of (11), it will be sufficient to verify the inequality $y'<0$. But this is trivial by (9) and $B=0$, $A>0$.

I.2. $A<0$. Then we have $R(\pi/2)>0$. More generally, if $R(x)\geqq 0$ for $0<x<$ $<\pi/2$ (Q is increasing) or $R(x)<0$ in $(0, x_2)$ and $R(x)>0$ in $(x_2, \pi/2)$ with any $x_2\in(0, \pi/2)$ (Q is decreasing in $(0, x_2)$ and increasing in $(x_2, \pi/2)$), we shall refer to it s *case* d). Since $Q(0)\leqq 0$ by (13) Q is, in this case, either negative or positive in the whole interval $(0, \pi/2)$, or negative in a certain subinterval $(0, x_0)$ and positive in $(x_0, \pi/2)$. Therefore, we again have to distinguish three cases here:

In the first case (case a2)) S is decreasing

in the second case (case a1)) S is increasing and

in the third case S is decreasing in $(0, x_0)$ and increasing in $(x_0, \pi/2)$.

It is easy to see that if for $x>\pi/2$ we have $S'(x)\geqq 0$ (this is supposed to be valid in case d)), then S is proved to be quasiconvex.

But now $\cos x<0$, $\sin x>0$ and, by (11), we have only to show that $y'>0$. This inequality follows from (9) in view of $A<0$, $B=0$.

From now on we can suppose that $B \neq 0$.

II. $D \leqq 0$ (see (15)). In this case the polynomial F does not change its sign.

II.1. $R \leqq 0$ in $(0, \pi)$. Combining $Q'(x) \leqq 0$ in $(0, \pi/2)$ with $Q(0) \leqq 0$ by (13), we find that $Q(x) \leqq 0$ and, by (12) and (11), that also $S'(x) \leqq 0$ for $x \leqq \pi/2$. Since case c) is realized here the function $S(x)$ turns out to be quasiconvex.

II.2. $R \geqq 0$ in $(0, \pi)$. Combining $Q'(x) \geqq 0$ in $(\pi/2, \pi)$ with $Q(\pi) \leqq 0$ by (13) we get $Q(x) \leqq 0$ and, owing to (12) and (11), $S'(x) \geqq 0$ for $x \geqq \pi/2$. Since the conditions of case d) are fulfilled, the function $S(x)$ is quasiconvex.

Consequently, in the following we shall confine ourselves to the more complicated case $D > 0$.

III. $D > 0$. This assumption ensures that $p < 1$ and

$$A(\mu - \lambda) < 0, \tag{16}$$

as can be seen from (15). The quadratic equation $F(z) = 0$ has exactly two different real roots

$$z_{+,-} = \frac{-2A \pm \sqrt{2A(p^2-1)(\mu-\lambda)}}{-2B}. \tag{17}$$

Obviously, they have the same sign if and only if

$$AB > 0. \tag{18}$$

Consequently, we have to study four subcases corresponding to the signs of A and B.

III.1. $A > 0$, $B > 0$. The graph of $F(z)$ is exhibited in Fig. 13a. Since $z_+ + z_- = 2\frac{A}{B}$ we obtain $z_+, z_- > 0$.

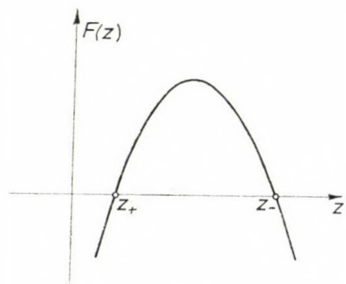


Fig. 13a

We proceed to show that the conditions of case c) are satisfied. For $x \geqq \pi/2$ is $z = \cos x \leqq 0$, hence $F(z) < 0$ and $R(x) < 0$. Now let $x < \pi/2$, then by (11) $S'(x) \leqq 0$, provided $y' < 0$ or $-A + B \cos x < 0$. But

$$-A + B \cos x < -A + B = (p-1)[-(\mu-\nu)p + \lambda - \nu] \tag{19}$$

where the first factor is negative. As $A>0$ we deduce from (16) $\lambda>\mu$, whence $\lambda-\nu>\mu-\nu$. Since $B=p(\lambda+\mu-2\nu)>0$, we have $(\lambda-\nu)+(\mu-\nu)>0$ and therefore $\lambda-\nu>0$. Consequently, we obtain $\lambda-\nu>p(\mu-\nu)$, so that the second factor in (19) is positive and $-A+B\cos x<0$, according to our assertion.

III.2. $A>0$, $B<0$ (see Fig. 13b). Then $z_-<0<z_+$. Observing that trivially $-A+B\cos x<0$ or $S'(x)<0$ for $x\leqq\pi/2$, we state that there is case c) again.

III.3. $A<0$, $B>0$ (see Fig. 13c). Then $z_+<0<z_-$.

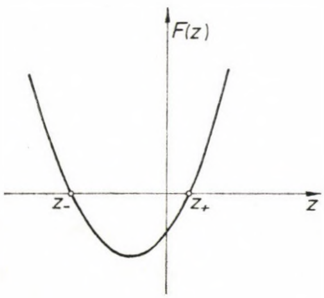


Fig. 13b

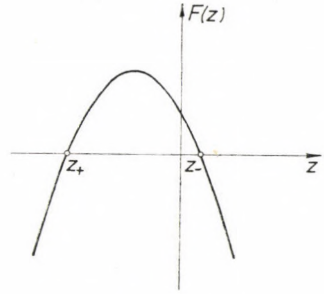


Fig. 13c

We shall show that the conditions of case d) are now fulfilled. To do this, we have yet to verify that $S'(x)>0$ or, owing to (11), that $y'\sin x-y\cos x>0$ for $x>\pi/2$. We have

$$(20)\quad y'\sin x-y\cos x=\frac{-A+p(\lambda+\mu-2\nu)\cos x}{1+p^2-2p\cos x}\sin x-(\lambda\alpha+\mu\beta+\nu x)\cos x>$$

$$>\left[\frac{\lambda+\mu-2\nu}{1+p^2-2p\cos x}\,p\sin x-\mu\beta\right]\cos x\geqq\left[\frac{\lambda+\mu-2\nu}{1+p^2-2p\cos x}\,p\sin x-\mu\sin\beta\right]\cos x,$$

the second factor being negative. On the other hand, we obtain, employing the sine theorem on the triangle ABC,

$$\sin\beta=\frac{p\sin x}{\sqrt{1+p^2-2p\cos x}},$$

hence the first factor on the right hand of (20) is

$$\frac{\lambda+\mu-2\nu}{1+p^2-2p\cos x}\,p\sin x-\mu\sin\beta=$$

$$=\frac{p\sin x}{1+p^2-2p\cos x}\left[(\lambda+\mu-2\nu)-\mu\sqrt{1+p^2-2p\cos x}\right].$$

It follows from (16) that $\mu>\lambda$ or $\mu-\nu>\lambda-\nu$. Since $0<B/p=(\mu-\nu)+(\lambda-\nu)$, we have $\mu-\nu>0$. On the other hand, $0>A=(\mu-\nu)p^2+(\lambda-\nu)$, whence $\lambda-\nu<0$ and $\lambda-2\nu<0$. Taking into account, further, that $1+p^2-2p\cos x>1$, we see that the expression in brackets is negative, and consequently the statement $y'\sin x-y\cos x>0$ for $x>\pi/2$ is true.

III.4. $A<0$, $B<0$ (see Fig. 13d). Then $0<z_-<z_+$.

We proceed to show that the assumptions of case d) are fulfilled again. For $x\geqq\pi/2$ is $y'>0$ by (9), hence $S'(x)>0$. Further, the vertex of the parabola $F(z)$ has the abszissa $z_0=\frac{A}{B}$. We claim that $z_0>1$. This inequality is equivalent to $-A+B>0$ or, by (19), to $-(\mu-\nu)p+(\lambda-\nu)<0$. It follows from (16) that $\lambda<\mu$ or $\lambda-\nu<\mu-\nu$. But $0>B/p=(\lambda-\nu)+(\mu-\nu)$, hence $\lambda-\nu<0$. Consequently, $-(\mu-\nu)p+(\lambda-\nu)<-(\mu-\nu)p+(\lambda-\nu)p=(\lambda-\mu)p<0$. Therefore, for the greater root of $F(z)$, $z_+>1$ holds, confirming our assertion.
Now, the proof of Lemma 5 is complete.

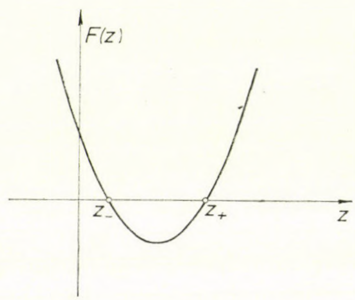


Fig. 13d

LEMMA 6. *Let C_a, C_1 and C_p be three circles of radii a ($0<a\leqq1$), 1 and p (>0, fixed), respectively, and mutually touching one another (Fig. 14). Then the density δ of C_a and C_1 with respect to the triangle Δ, determined by the centres of the three circles, attains its maximum only for $a=1$.*

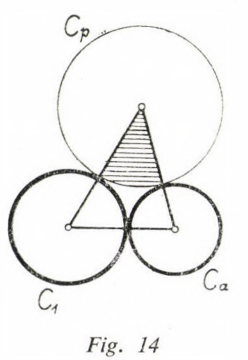


Fig. 14

PROOF. Obviously,

$$\delta=\frac{a^2\varphi_a+\varphi_1}{2\Delta}$$

where φ_a and φ_1 denote the central angles belonging to C_a and C_1. By elementary calculation we find

$$\sqrt{p}\delta = f(a, p) \equiv \frac{1}{\sqrt{a(a+p+1)}}\left[a^2 \operatorname{arctg}\sqrt{\frac{p}{a(a+p+1)}} + \operatorname{arctg}\sqrt{\frac{ap}{a+p+1}}\right].$$

To examine this function, we differentiate it partially and obtain

$$\frac{\partial f}{\partial a} = \frac{2a+p+1}{2[a(a+p+1)]^{3/2}} f_1(a, p)$$

with

$$f_1(a, p) = \frac{a^2(2a+3p+3)}{2a+p+1} \operatorname{arctg}\sqrt{\frac{p}{a(a+p+1)}} - \operatorname{arctg}\sqrt{\frac{ap}{a+p+1}} -$$

$$-\frac{a^2+(a+p)(a-1)}{(a+p)(2a+p+1)}\sqrt{ap(a+p+1)}.$$

Further differentiation yields

$$\frac{\partial f_1}{\partial a} = 2a\frac{4a^2+6ap+6a+3p^2+6p+3}{(2a+p+1)^2} f_2(a, p)$$

where

$$f_2(a, p) = \operatorname{arctg}\sqrt{\frac{p}{a(a+p+1)}} - \frac{1}{2}\frac{\sqrt{ap(a+p+1)}}{(a+1)^2(4a^2+6ap+6a+3p^2+6p+3)}\cdot$$

$$\cdot\left[\left(\frac{2a+p+1}{a+p}\right)^2(2a^2+3ap+3a+4p) + \frac{4a+p+3}{a}\right].$$

After some laborious calculations we obtain, putting $a^m p^n = (m, n)$,

$$4a(a+p)^3(a+1)^2(4a^2+6ap+6a+3p^2+6p+3)^2\sqrt{\frac{a(a+p+1)}{p}}\frac{\partial f_2}{\partial a} =$$

$$= f_3(a, p) \equiv -96\ (7,4) - 128\ (7,1) - 336\ (6,3) - 880\ (6,2) -$$

$$-320\ (6,1) - 456\ (5,4) - 1920\ (5,3) - 1688\ (5,2) - 288\ (5,1) -$$

$$-294\ (4,5) - 1918\ (4,4) - 2850\ (4,3) - 1298\ (4,2) - 104\ (4,1) -$$

$$-87\ (3,6) - 944\ (3,5) - 2086\ (3,4) - 1620\ (3,3) - 435\ (3,2) -$$

$$-12\ (3,1) - 9\ (2,7) - 216\ (2,6) - 658\ (2,5) - 700\ (2,4) -$$

$$-309\ (2,3) - 60\ (2,2) - 18\ (1,7) - 63\ (1,6) - 48\ (1,5) + 18\ (1,4) +$$

$$+18\ (1,3) - 3\ (1,2) + 3\ (0,7) + 18\ (0,6) + 36\ (0,5) + 30\ (0,4) + 9\ (0,3).$$

Since $\frac{\partial^2 f_3}{\partial a^2} < 0$, f_2 is a concave function of a. Note that $f_3(0, p) > 0$ and

$$f_3(1, p) = -24p^7 - 348p^6 - 1908p^5 - 5112p^4 - 7008p^3 - 4460p^2 - 852p < 0$$

for any positive p. Therefore, $f_3(a, p)$ and also $\frac{\partial f_2}{\partial a}$ passes from positive to negative values when a varies, increasing from 0 to 1.

We observe that

$$f_2(1,p) = \operatorname{arctg}\sqrt{\frac{p}{p+2}} - \frac{2p^3+14p^2+27p+13}{2(p+1)^2(3p^2+12p+13)}\sqrt{p(p+2)},$$

hence

$$\frac{d}{dp}f_2(1,p) = \frac{1}{2(p+1)^3(3p^2+12p+13)^2}\sqrt{\frac{p}{p+2}}\cdot$$

$$\cdot[21p^5+170p^4+527p^3+767p^2+496p+91] > 0.$$

Since $f_2(1,0)=0$, we have $f_2(1,p)>0$. Combining this with $\lim_{a\to 0} f_2(a,p) = -\infty$, we deduce that $f_2(a,p)$ and also $\frac{\partial f_1}{\partial a}$ passes from negative to positive values when a increases. In view of $f_1(0,p)=0$ it follows that $f(a,p)$ as a function of a assumes its maximum only in a boundary point of the interval $0 \leqq a \leqq 1$. But it is easily proved, in a similar way as above, that

$$\frac{\sqrt{p+2}}{2}[f(1,p)-f(0,p)] = \operatorname{arctg}\sqrt{\frac{p}{p+2}} - \frac{\sqrt{p(p+2)}}{2(p+1)} > 0$$

for $p>0$. This completes the proof of Lemma 6.

Finally, it is very easy to prove the following two lemmas:[8]

LEMMA 7. *Let $\Delta_k = OTP_k$ $(k=1,2)$ be right triangles $(\sphericalangle OTP_k = \pi/2)$, where the sides TP_k do not have common interior points with the circle C of centre O (Fig. 15). If $\overline{OP_1} < \overline{OP_2}$ then*

$$\frac{C\cap\Delta_1}{\Delta_1} > \frac{C\cap\Delta_2}{\Delta_2}.$$

LEMMA 8. *Let $\Delta_k = OT_kP$ $(k=1,2)$ be right triangles $(\sphericalangle OT_kP = \pi/2)$, where the sides T_kP do not have common interior points with the circle C of centre O (Fig. 16). If $\sphericalangle T_1OP < \sphericalangle T_2OP$ then*

$$\frac{C\cap\Delta_1}{\Delta_1} < \frac{C\cap\Delta_2}{\Delta_2}.$$

Let us now return to the proof of our theorem.

For simplicity's sake, let us denote by $O_1V_1O_2V_2$ an arbitrary quadrangle of the tessellation T, where O_1, O_2 are the centres of the circles C_1, C_2 of $\{C_i\}$ and V_1, V_2 are the corresponding vertices of the tessellation S.

[8] See MOLNÁR [5] [6], [7], [8], [9].

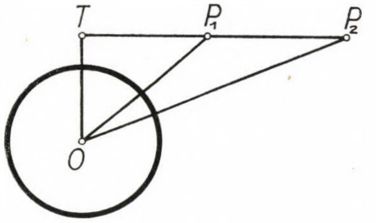


Fig. 15

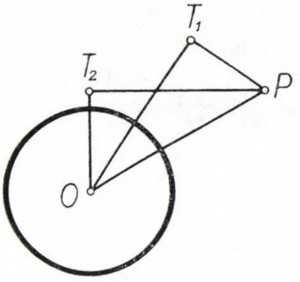


Fig. 16

We now proceed to show that in $O_1V_1O_2V_2$ the density of $\{C_i\}$ does not exceed $d(\varrho)$, i.e.

$$d_{12}(V_1V_2) \leqq d(\varrho) = \frac{\arccos\dfrac{1}{1+\varrho}}{\sqrt{2\varrho+\varrho^2}},$$

and distinguish the following two cases:

a) *$O_1V_1O_2V_2$ is convex.* In this case we decompose $O_1V_1O_2V_2$ into two triangles $\Delta_1=O_1O_2V_1$ and $\Delta_2=O_1O_2V_2$. Obviously, the inequality $d_{12}(V_1V_2)\leqq \leqq d(\varrho)$ is valid if we can show that

$$d_{12}(V_i) = \frac{C_1\cap\Delta_i + C_2\cap\Delta_i}{\Delta_i} \leqq d(\varrho) \quad (i = 1, 2),$$

and it suffices for $i=1$.

Consider three circles of radii r, r, ϱ^* and of centres A, B, C, respectively, mutually touching one another. We denote with $d(r, r, \varrho^*)$ the density of the circles of centres A, B with respect to the triangle ABC; obviously $d(r, r, \varrho^*)= =d\left(1, 1, \frac{\varrho^*}{r}\right)=d\left(\frac{\varrho^*}{r}\right)$.

Let $\varrho^*\geqq\varrho$ be the radius of the supporting circle C centred at V_1 which touches C_1, C_2.

If the segment O_1O_2 has no common interior points with C (Fig. 17) then, in view of Lemma 5, $d_{12}(V_1)$ attains its maximum for one of the following configurations:

(i) C_1 and C_2 (radii r_1 and r_2, $r_1\leqq r_2$) touch one another (Fig. 18). Making use of Lemma 6 and, if necessary, of Lemma 7, we obtain $d_{12}(V_1)\leqq d(r_2, r_2, \varrho^*)= =d\left(1, 1, \frac{\varrho^*}{r_2}\right)\leqq d(\varrho)$.

(ii) The segment O_1O_2 touches C (Fig. 19). We draw the tangents from V_1 to C_1 and C_2 and denote the points of tangency with T_1 and T_2. In view of Lemma 8 we get

$$\frac{C_i\cap O_iPV_1}{O_iPV_1} < \frac{C_i\cap O_iT_iV_1}{O_iT_iV_1} = d(r_i, r_i, \varrho^*) \quad (i = 1, 2),$$

where P is the foot of the perpendicular from V_1 to O_1O_2. Thus $d_{12}(V_1) \leqq \leqq d(r_2, r_2, \varrho^*) \leqq d(\varrho)$.

If the segment O_1O_2 has common interior points with C, the inequality $d_{12}(V_1) \leqq \leqq d(\varrho)$ can be proved in the same way as in case (ii) of a).

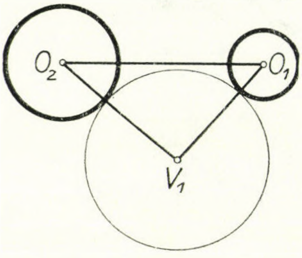


Fig. 17

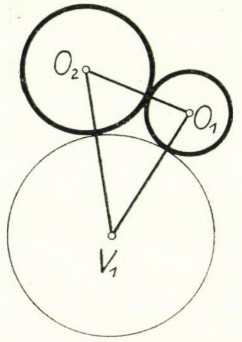


Fig. 18

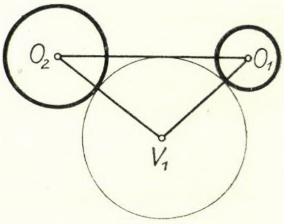


Fig. 19

b) $O_1V_1O_2V_2$ *is concave.* Let $\overline{O_1V_1} < \overline{O_1V_2}$ (Fig. 20), then by Lemma 4 we have $d_{12}(V_1V_2) \leqq d_{12}(V_1)$. But we have already seen that $d_{12}(V_1) \leqq d(\varrho)$.

This completes the proof of our statement that in each quadrangle of the tessellation T the density of $\{C_i\}$ is not greater than $d(\varrho)$. In order to deduce, finally, the inequality $d \leqq d(\varrho)$, we remark that, in view of $\sup r_i \leqq 1$, the circumradii of

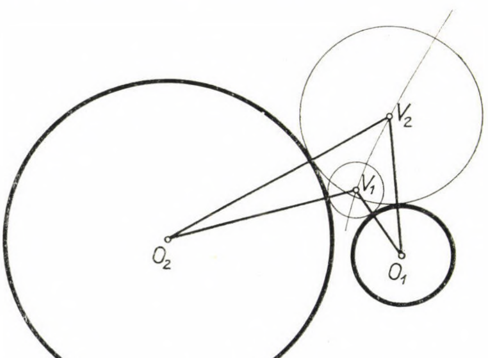


Fig. 20

the quadrangles of the tessellation T have also a finite upper bound b. Denoting by $Q_{ij} = O_iAO_jB$ the quadrangle of the tessellation T corresponding to the circles C_i, C_j and taking into account that

$$C_i(AO_iB) + C_j(AO_jB) \leqq d(\varrho)Q_{ij},$$

we obtain

$$\frac{1}{\pi R^2}\sum_i (C_i \cap C(R)) \leqq \frac{1}{\pi R^2} \sum_{C \cap C(R) \neq \emptyset} C_i \leqq \frac{d(\varrho)}{\pi R^2} \sum Q_{ij} \leqq$$

$$\leqq \frac{\pi (R+2+2b)^2}{\pi R^2} d(\varrho) = \left(1+\frac{2+2b}{R}\right)^2 d(\varrho).$$

From this the desired inequality $d \leqq d(\varrho)$ follows immediately.

REMARK. The Lemmas 2, 4, 5, 7, 8 continue to be valid whenever the "measure" of C_i is an arbitrary positive value $\varphi(r_i)$. The system of values $\{\varphi(r_i)\}$ associated to $\{C_i\}$ is called a functional system of $\{C_i\}$ and the corresponding density a functional density.

Lemma 6, however, is no longer valid for an arbitrary functional system. The case of a decreasing function $\varphi(r)$ yields a trivial counterexample. It is easy to give counterexamples also for certain increasing functions $\varphi(r)$. But it seems likely that Lemma 6 continues to hold for some particular functional systems of $\{C_i\}$.

References

[1] L. FEJES TÓTH, *Lagerungen in der Ebene, auf der Kugel und im Raum.* Zweite Auflage (Berlin—Heidelberg—New York, 1972).

[2] L. FEJES TÓTH and J. MOLNÁR, Unterdeckung und Überdeckung der Ebene durch Kreise, *Math. Nachrichten,* **18** (1958), 235—243.

[3] A. FLORIAN, Ausfüllung der Ebene durch Kreise, *Rend. Circ. Mat. Palermo,* **(2) 9** (1960), 300—312.

[4] L. HÁRS, Weighted circle systems (Thesis, in Hungarian), L. Eötvös University (Budapest, 1977).

[5] J. MOLNÁR, Collocazioni di cerchi sulla superficie di curvatura constante, *Celebrazioni archimedee del secolo XX, Siracusa, 11—16 aprile 1961,* pp. 61—72.

[6] J. MOLNÁR, Körelhelyezések állandó görbületű felületeken, *Magyar Tud. Akad. III. Oszt. Közl.,* **12** (1962), 223—263 (in Hungarian).

[7] J. MOLNÁR, Kreislagerungen auf Flächen konstanter Krümmung, *Math. Annalen,* **158** (1965), 365—376.

[8] J. MOLNÁR, Collocazioni di cerchi di Minkowski, *Rend. Sem. Mat. di Messina,* **10** (1966), 100—104.

[9] J. MOLNÁR, Aggregati di cerchi di Minkowski, *Annali di Matematica pura ed applicata,* **71** (1966), 101—108.

[10] J. MOLNÁR, On the ϱ-system of unit circles, *Annales Univ. Sci. Budapest, Sectio Math.,* **20** (1977), 195—203.

[11] J. MOLNÁR, Sur les empilements de ϱ-système de cercles dans un domaine convexe, *Coll. on Intuitive Geometry* (Tihany, 1975).

(Received November 16, 1977)

DEPARTMENT OF MATHEMATICS,
UNIVERSITY OF SALZBURG,
PETERSBRUNNSTRAßE 19,
A—5020 SALZBURG, AUSTRIA.

DEPARTMENT OF GEOMETRY,
L. EÖTVÖS UNIVERSITY,
MÚZEUM KRT. 6—8,
H—1088 BUDAPEST.

INDEX

Printed in Hungary

A kiadásért felel az Akadémiai kiadó igazgatója Műszaki szerkesztö: Zacsik Annamária
A kézirat nyomdába érkezett: 1979. I. 16. — Terjedelem: 19,60 (A/5) ív 29 ábra

79-244 — Szegedi Nyomda — F. v.: Dobó József

The Acta Mathematica publishes papers on mathematics in English, German, Frencn and Russian. It appears in parts of various size, making up volumes. Manuscripts should be sent to:

Acta Mathematica, H-1053 Budapest, Reáltanoda u. 13—15.

Correspondence with the publishers should be sent to:

Acta Mathematica, H-1363 Budapest, Pf. 24.

The rate of subscription is $36.00 a volume. Orders may be placed with "*Kultura*" Foreign Trade Company (1011 Budapest, Fő u. 32. Account No. 218-10990) or with representatives abroad.

Instructions for authors. Manuscripts should be typed on standard size papers (25 rows; 50 characters in each row). Mathematical symbols (except sin, log, etc.) will be set in italics, so *use handwritten letters, or indicate them by simple underlining.* The text of theorems, lemmas, corollaries are to be printed also in italics. The words THEOREM, LEMMA, COROLLARY, PROOF, REMARK, PROBLEM, etc. (or their German and French equivalents) should be typed in capital letters. Sub-headings will be set in boldface lower case; indicate them by red underlining. Other special types of symbols (such as German, Greek, script, etc.) should also be clearly identified. Mark footnotes by consecutive superscript numbers. When listing references, please follow the following pattern:

[1] G. Szegő, *Orthogonal polynomials*, AMS Coll. Publ. Vol. XXIII (Providence, 1939).
[2] A. Zygmund, Smooth functions, *Duke Math. J.*, **12** (1945), 47—76.

For abbreviation of names of journals follow the Mathematical Reviews. After the references give the author's affiliation.

Do not send abstract and second copy of the manuscript. Authors will receive only galley-proofs (one copy). Manuscripts will not be sent back to authors (neither for the purpose of proof-reading nor when rejecting a paper).

Authors obtain 100 reprints free of charge. Additional copies may be ordered from the publishers.

All the reviews of the Hungarian Academy of Sciences may be obtained among others from the following bookshops:

1979. XI. 30. HU ISSN 0001—5954 **Index: 26014**

7634

ACTA MATHEMATICA

ACADEMIAE SCIENTIARUM HUNGARICAE

ADIUVANTIBUS

REDIGIT

K. TANDORI

CORREDACTOR

J. SZABADOS

TOMUS XXXIV FASCICULI 3—4

AKADÉMIAI KIADÓ, BUDAPEST
1979

ACTA MATH. HUNG.

ACTA MATHEMATICA
ACADEMIAE SCIENTIARUM HUNGARICAE

A MAGYAR TUDOMÁNYOS AKADÉMIA III. OSZTÁLYÁNAK
MATEMATIKAI KÖZLEMÉNYEI

SZERKESZTŐSÉG: 1053 BUDAPEST, REÁLTANODA U. 13—15.

KIADÓHIVATAL: 1363 BUDAPEST, PF. 24.

Az Acta Mathematica angol, német, francia és orosz nyelven közöl értekezéseket a matematika köréből. Váltakozó terjedelmű füzetekben jelenik meg, több füzet alkot egy kötetet. A közlésre szánt kéziratok a szerkesztőség, minden más levelezés a kiadóhivatal címére küldendő.

Megrendelhető a belföld számára az Akadémiai Kiadónál (1054 Budapest, Alkotmány utca 21. Bankszámla 215-11488), a külföld számára pedig a Kultúra Külkereskedelmi Vállalatnál (1011 Budapest, Fő utca 32. Bankszámla 218-10990), vagy annak külföldi képviseleteinél és bizományosainál.

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (3—4), (1979), 223—231.

THE HOM AND TENSOR FUNCTORS WITH MIXED LIMITS

By

I. A. ASSEM, C. G. CHEHATA and M. EL GENDY* (Alexandria)

Introduction

In a previous paper [2], we have introduced the concepts of mixed systems and limits. In that paper, the existence and uniqueness of the limits was proved and an explicit construction of these limits and the canonical morphism was given. We also defined a category $\mathcal{M}(I\times L, \mathcal{C})$ of mixed systems in the category $\mathcal{C}$ over $I\times L$. It is the purpose of the present work to study the relation between the mixed limit functors and the functors Hom and $\otimes$. We prove that we have, in each case, a commutative diagram of functors and functorial morphisms. We also prove that the mixed limit functors on the category of modules are representable (if L is directed) and the representing objects are finitely generated projective modules. We use the same notations as in [2].

1. The Hom functor and mixed systems

Given an arbitrary object G in a category $\mathcal{C}$ with inductive and projective limits, we shall consider the action of the covariant functor $h_G(-)=\mathrm{Hom}_{\mathcal{C}}(G, -)$ and the contravariant functor $h^G(-)=\mathrm{Hom}_{\mathcal{C}}(-, G)$ on some mixed system $(E_\alpha^\lambda, f_{\alpha\beta}^{\mu\lambda})_{I\times L}$ of $\mathcal{C}$. We shall need the following lemmas;

LEMMA 1.1. *Let $(E_\alpha, g_{\alpha\beta})_I$ be a projective system in $\mathcal{C}$, then $\big(h_G(E_\alpha), h_G(g_{\alpha\beta})\big)_I$ is a projective system of sets. Moreover there exists a unique bijection:*

$$s\colon \varprojlim_{\alpha\in I} h_G(E_\alpha) \to h_G\big(\varprojlim_{\alpha\in I} E_\alpha\big)$$

such that, for all $\alpha\in I$, $h_G(g_\alpha)s=k_\alpha$ (where k_α and g_α represent the respective canonical projections).

The proof can be found in [3].

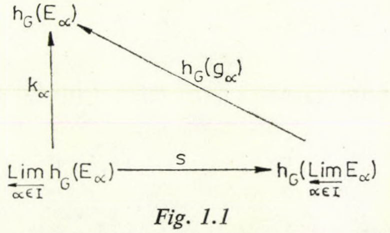


Fig. 1.1

* Research was carried out at the Science Centre, University of Alexandria. The authors wish to express their thanks to Professor R. Wiegandt for the interest he took in this work.

COROLLARY 1.1. *The family* $(h_G(g_\alpha))_{\alpha\in I}$ *is monomorphic.*

Let indeed (u, v) be a pair of maps such that, for all $\alpha\in I$, $h_G(g_\alpha)u=h_G(g_\alpha)v$. Then $k_\alpha s^{-1}u=k_\alpha s^{-1}v$. Since $(k_\alpha)_{\alpha\in I}$ is a monomorphic family, and s^{-1} is bijective, this implies $u=v$.

LEMMA 1.2. *Let* $(E^\lambda, h^{\mu\lambda})_L$ *be an inductive system in* $\mathscr{C}$, *then* $(h^G(E^\lambda), h^G(h^{\mu\lambda}))_L$ *is a projective system of sets. Moreover there exists a unique bijection*

$$s\colon \varprojlim_{\lambda\in L} h^G(E^\lambda) \to h^G(\varinjlim_{\lambda\in L} E^\lambda)$$

such that, for all $\lambda\in L$, $h^G(h^\lambda)s=k_\lambda$ *(where* k_λ *and* h^λ *represent the respective canonical morphisms).*

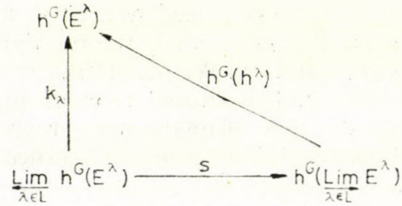


Fig. 1.2

The proof can be found in [3].

COROLLARY 1.2. *The family* $(h^G(h^\lambda))_{\lambda\in L}$ *is monomorphic.*

This can be proved exactly as Corollary 1.1.

LEMMA 1.3. *Let* $\mathscr{E}=(E_\alpha^\lambda, f_{\alpha\beta}^{\mu\lambda})_{I\times L}$ *be a mixed system in* $\mathscr{C}$ *over* $I\times L$, *then* $h_G(\mathscr{E})=$ $=(h_G(E_\alpha^\lambda), h_G(f_{\alpha\beta}^{\mu\lambda}))_{I\times L}$ *is a mixed system of sets over* $I\times L$.

This follows from the covariance of h_G.
From now on, we shall write

$$K_\alpha^\lambda = h_G(E_\alpha^\lambda),\quad k_{\alpha\beta}^\lambda = h_G(g_{\alpha\beta}^\lambda)\ (\alpha\leqq\beta),\quad l_\alpha^{\gamma\lambda} = h_G(h_\alpha^{\mu\lambda})\ (\lambda\leqq\mu),$$

$$(K_\alpha, l_\alpha^\lambda) = \varinjlim_{\lambda\in L} (K_\alpha^\lambda, l_\alpha^{\mu\lambda}),\quad (K^\lambda, k_\alpha^\lambda) = \varprojlim_{\alpha\in I} (K_\alpha^\lambda, k_{\alpha\beta}^\lambda),$$

$$k_{\alpha\beta} = \varinjlim_{\lambda\in L} k_{\alpha\beta}^\lambda\ (\alpha\leqq\beta),\quad l^{\mu\lambda} = \varprojlim_{\alpha\in I} l_\alpha^{\mu\lambda}\ (\lambda\leqq\mu),$$

$$(J, l^\lambda) = \varinjlim_{\lambda\in l} (K^\lambda, l^{\mu\lambda}),\quad (K, k_\alpha) = \varprojlim_{\alpha\in I} (K_\alpha, k_{\alpha\beta})$$

and $g\colon J\to K$ will denote the canonical morphism of the system $h_G(\mathscr{E})$. Consider now Figure 1.3. Our aim is to prove the existence of unique maps $(s^\lambda)_{\lambda\in L}$, $(r_\alpha)_{\alpha\in I}$, p and q such that the following diagram is commutative for any $(\alpha, \lambda)\in I\times L$:

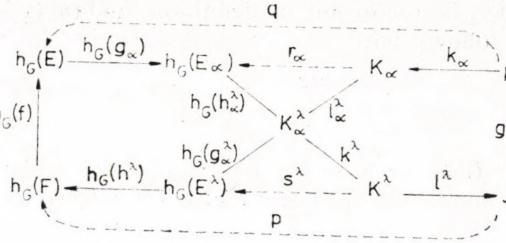


Fig. 1.3

By Lemma 1.1, for any $\lambda \in L$, there exists a unique bijection $s^\lambda \colon K^\lambda \to h_G(E^\lambda)$ such that, for all $\alpha \in I$,

$$h_G(g^\lambda_\alpha) s^\lambda = k^\lambda_\alpha. \tag{1.1}$$

It is easy to see that, if $\lambda \leqq \mu$,

$$h_G(h^{\mu\lambda}) s^\lambda = s^\mu l^{\mu\lambda} \tag{1.2}$$

and hence $(s^\lambda)_{\lambda \in L}$ is an inductive system of bijections.

Since by the covariance of h_G, $h_G(h^\lambda) = h_G(h^\mu) h_G(h^{\mu\lambda})$, this implies $h_G(h^\lambda) s^\lambda = h_G(h^\mu) s^\mu l^{\mu\lambda}$. Therefore there exists a unique $p \colon J \to h_G(F)$ such that, for all $\lambda \in L$,

$$p l^\lambda = h_G(h^\lambda) s^\lambda. \tag{1.3}$$

This completes the construction of the maps in the lower half of Figure 1.3 and the proof of the commutativity of this lower half.

Let now $\lambda \leqq \mu$ in L and $\alpha \in I$ be arbitrary. The covariance of h_G implies $h_G(h^\lambda_\alpha) = h_G(h^\mu_\alpha) l^{\mu\lambda}_\alpha$ hence there exists a unique $r_\alpha \colon K_\alpha \to h_G(E_\alpha)$ such that, for all $\lambda \in L$,

$$r_\alpha l^\lambda_\alpha = h_G(h^\lambda_\alpha). \tag{1.4}$$

It is easy to prove that $(r_\alpha)_{\alpha \in I}$ is a projective system of maps. Let $r = \varprojlim_{\alpha \in I} r_\alpha$; it is by definition the only map such that, for all $\alpha \in I$,

$$k^*_\alpha r = r_\alpha k_\alpha \tag{1.5}$$

where $k^*_\alpha \colon \varprojlim_{\alpha \in I} h_G(E_\alpha) \to h_G(E_\alpha)$ is the projection map.

By Lemma 1.1, there exists a unique bijection $r^* \colon \varprojlim_{\alpha \in I} h_G(E_\alpha) \to h_G(E)$ such that, for all $\alpha \in I$,

$$h_G(g_\alpha) r^* = k^*_\alpha. \tag{1.6}$$

Let us put $q = r^* r$, then, for any $\alpha \in I$,

$$h_G(g_\alpha) q = r_\alpha k_\alpha. \tag{1.7}$$

There only remains to establish the commutativity of the outer square of Figure 1.3. For any $(\alpha, \lambda) \in I \times L$,

$$h_G(g_\alpha) q g l^\lambda = r_\alpha k_\alpha g l^\lambda = r_\alpha l^\lambda_\alpha k^\lambda_\alpha = h_G(h^\lambda_\alpha) k^\lambda_\alpha = h_G(h^\lambda_\alpha) h_G(g^\lambda_\alpha) s^\lambda =$$
$$= h_G(g_\alpha) h_G(f) h_G(h^\lambda) s^\lambda = h_G(g_\alpha) h_G(f) p l^\lambda.$$

Since the family $(l^\lambda)_{\lambda\in L}$ is epimorphic by definition, and $(h_G(g_\alpha))_{\alpha\in I}$ is monomorphic by Corollary 1.1, it follows that

$$qg = h_G(f)p. \tag{1.8}$$

Thus we have:

LEMMA 1.4. *Let G be an arbitrary object of $\mathscr{C}$, then, for every mixed system $\mathscr{E}$ in $\mathscr{C}$, there exist maps p and q such that Figure 1.4 below is a commutative diagram.*

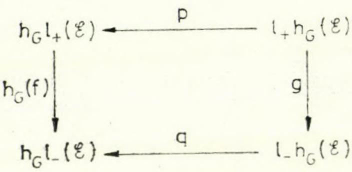


Fig. 1.4

Now it is clear (cf. [2], Corollary 3.3) that $h_G(f)$ and g are functorial maps. On the other hand, following step by step their construction, it is easy to show that the maps $(s^\lambda)_{\lambda\in L}$, p, $(r_\alpha)_{\alpha\in I}$, r, r^* and q are functorial. We have proved:

THEOREM 1.1. *Let G be an arbitrary object of $\mathscr{C}$. Then we have the following commutative diagram of covariant functors and of functorial morphisms of $\mathscr{M}(I\times L, \mathscr{C})$ into the category of sets:*

Fig. 1.5

In general, p and q are not isomorphisms, so that the functors l_+ and l_- are not representable. Now let $\mathscr{C}$ be the category ${}_R$Mod of left modules over an arbitrary ring R. Then we have

LEMMA 1.5. *Let $(E^\lambda, h^{\mu\lambda})_L$ be an inductive system of (left) R-modules over a directed set L, then:*

(a) *For every (left) R-module G, there exists a unique abelian group homomorphism*

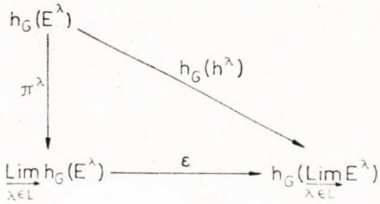
$$\varepsilon\colon \varinjlim_{\lambda\in L} h_G(E^\lambda) \to h_G(\varinjlim_{\lambda\in L} E^\lambda)$$

such that, for all $\lambda\in L$, $\varepsilon\pi^\lambda = h_G(h^\lambda)$ where h^λ and π^λ are the respective canonical morphisms.

(b) *If G is finitely generated, ε is a monomorphism.*

(c) *If G is finitely generated projective, ε is an isomorphism.*

(Cf. [1], Problems of Chapter II.)



Let G be a left R-module and $\mathscr{E}=(E_\alpha^\lambda, f_{\alpha\beta}^{\mu\lambda})_{I\times L}$ a mixed system of left R-modules with L directed. All the sets and maps whose existence was proved above are abelian groups and homomorphisms.

For any $\lambda\in L$, s^λ is an isomorphism, hence so is $s=\underset{\lambda\in L}{\underrightarrow{\mathrm{Lim}}}\, s^\lambda$: $J\to\underset{\lambda\in L}{\underrightarrow{\mathrm{Lim}}}\, h_G(E^\lambda)$, and the relation $sl^\lambda=\pi^\lambda s^\lambda$ implies that

$$\varepsilon s l^\lambda = \varepsilon\pi^\lambda s^\lambda = h_G(h^\lambda)s^\lambda = pl^\lambda.$$

Since the $(l^\lambda)_{\lambda\in L}$ form an epimorphic family of homomorphisms, we deduce $\varepsilon s=p$. Now s is an isomorphism, thus

(i) if G is finitely generated, then ε is a monomorphism and so is p,

(ii) if G is finitely generated projective, then ε is an isomorphism and so is p.

Consider the maps $(r_\alpha)_{\alpha\in I}$ each of these is constructed as ε, and q was defined as r^*r, where $r=\underset{\alpha\in I}{\underleftarrow{\mathrm{Lim}}}\, r_\alpha$ and r^*: $\underset{\alpha\in I}{\underleftarrow{\mathrm{Lim}}}\, h_G(E_\alpha)\to h_G(E)$ is an abelian group isomorphism, hence

(i) if G is finitely generated, each of the r_α is a monomorphism, and so are r and q,

(ii) if G is finitely generated projective, each of the r_α is an isomorphism, and so are r and q.

Thus we have:

THEOREM 1.2. *Let $\mathscr{E}$ be a mixed system of left R-modules over $I\times L$, where L is directed, and let G be a finitely generated R-module. Then:*

(i) *the canonical maps of Lemma* 1.4, p: $l_+h_G(\mathscr{E})\to h_G l_+(\mathscr{E})$ *and* q: $l_-h_G(\mathscr{E})\to h_G l_-(\mathscr{E})$ *are monomorphisms,*

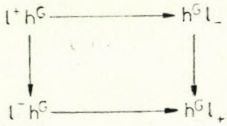
(ii) *if, moreover, G is projective, p and q are isomorphisms, and hence the functors l_+ and l_- are representable.*

COROLLARY 1.3. *Let $\mathscr{E}$ be a mixed system of left R-modules over $I\times L$, where L is directed, and let G be a finitely generated projective R-module. If the canonical map f of $\mathscr{E}$ is an isomorphism, so is the canonical map g of $h_G(\mathscr{E})$.*

This follows at once from equation (1.8) and Theorem 1.2.

We now let $h^G(-)$ be the contravariant functor $\mathrm{Hom}_{\mathscr{C}}(-, G)$. Let also $l^+(-)$ and $l^-(-)$ denote the functors $\underset{\alpha\in I}{\underrightarrow{\mathrm{Lim}}}\,\underset{\alpha\in L}{\underleftarrow{\mathrm{Lim}}}\,(-)$ and $\underset{\lambda\in L}{\underleftarrow{\mathrm{Lim}}}\,\underset{\alpha\in I}{\underrightarrow{\mathrm{Lim}}}\,(-)$, respectively. We can prove exactly as above:

THEOREM 1.3. *Let G be an arbitrary object of $\mathscr{C}$. Then we have the following commutative diagram of contravariant functors and of functorial maps of $\mathscr{M}(I\times L, \mathscr{C})$ into the category of sets:*



2. The bifunctor Hom (—, —) and two inductive systems

Let $(A_\alpha, \varphi_{\beta\alpha})_I$ and $(B^\lambda, \psi^{\mu\lambda})_L$ be two inductive systems in some category $\mathscr{C}$ having inductive limits. Then:

LEMMA 2.1. *The system* $(\mathrm{Hom}\,(A_\alpha, B^\lambda),\ \mathrm{Hom}\,(\varphi_{\beta\alpha}, \psi^{\mu\lambda}))_{I\times L}$ *is a mixed system of sets.*

This follows from the fact that the bifunctor Hom $(-, -)$ is contravariant in the first variable, and covariant in the second.

Let
$$E^\lambda_\alpha = \mathrm{Hom}\,(A_\alpha, B^\lambda),\ f^{\mu\lambda}_{\alpha\beta} = \mathrm{Hom}\,(\varphi_{\beta\alpha}, \psi^{\mu\lambda}),$$
$$(A, \varphi_\alpha) = \varinjlim_{\alpha\in I}(A_\alpha, \varphi_{\beta\alpha}),\ (B, \psi^\lambda) = \varinjlim_{\lambda\in L}(B^\lambda, \psi^{\mu\lambda}).$$

The canonical pentagon of the mixed system $(E^\lambda_\alpha, f^{\mu\lambda}_{\alpha\beta})_{I\times L}$ will be as in Figure 2.1.

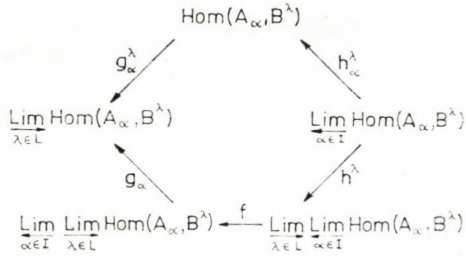


Fig. 2.1

The covariance of Hom in the second variable implies that $(\mathrm{Hom}\,(A, B^\lambda), \mathrm{Hom}\,(A, \psi^{\mu\lambda}))_L$ is an inductive system. We shall write $k^{\mu\lambda} = \mathrm{Hom}\,(A, \psi^{\mu\lambda})$ and denote the canonical morphisms by $k^\lambda\colon \mathrm{Hom}\,(A, B^\lambda)\to\varinjlim_{\lambda\in L}\mathrm{Hom}\,(A, B^\lambda)$. On the other hand, the contravariance of Hom in the second variable implies that $(\mathrm{Hom}\,(A_\alpha, B), \mathrm{Hom}\,(\varphi_{\beta\alpha}, B))_I$ is projective. We shall write $l_{\alpha\beta} = \mathrm{Hom}\,(\varphi_{\beta\alpha}, B)$ and denote the canonical projection by $l_\alpha\colon \varprojlim_{\alpha\in I}\mathrm{Hom}\,(A_\alpha, B)\to\mathrm{Hom}\,(A_\alpha, B)$.

By Lemma 1.2, for each $\lambda\in L$, there exists, a unique bijection
$$r^\lambda\colon \varprojlim_{\alpha\in I}\mathrm{Hom}\,(A_\alpha, B^\lambda)\to\mathrm{Hom}\,(A, B^\lambda)$$

such that, for all $\lambda\in L$,

$$\text{Hom}(\varphi_\alpha, B^\lambda)r^\lambda = g_\alpha^\lambda. \tag{2.1}$$

It is easy to prove that, if $\lambda\leqq\mu$,

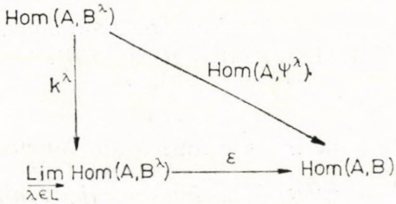
$$r^\mu h^{\mu\lambda} = \text{Hom}(A, \psi^{\mu\lambda})r^\lambda, \tag{2.2}$$

that is, the $(r^\lambda)_{\lambda\in L}$ form an inductive system of bijections. Therefore $r=\varinjlim_{\lambda\in L} r^\lambda$ is also a bijection. It is actually the only map of $\varinjlim_{\lambda\in L}\varprojlim_{\alpha\in I}\text{Hom}(A_\alpha, B)$ into $\varinjlim_{\lambda\in L}\text{Hom}(A, B^\lambda)$ such that, for all $\lambda\in L$,

$$rh^\lambda = k^\lambda r^\lambda. \tag{2.3}$$

On the other hand, since Hom $(A, -)$ is a covariant functor, $\lambda\leqq\mu$ implies $\text{Hom}(A, \psi^\lambda)=\text{Hom}(A, \psi^\mu)k^{\mu\lambda}$ and hence there exists a unique map ε such that, for all $\lambda\in L$,

$$\varepsilon k^\lambda = \text{Hom}(A, \psi^\lambda). \tag{2.4}$$



If $u=\varepsilon r$, then, for all $\lambda\in L$, $uh^\lambda=\varepsilon k^\lambda r^\lambda$ or

$$uh^\lambda = \text{Hom}(A, \psi^\lambda)r^\lambda. \tag{2.5}$$

Let $(\lambda, \mu)\in L^2$, $\lambda\leqq\mu$ and $\alpha\in I$ be arbitrary. Then the relation $\text{Hom}(A_\alpha, \psi^\lambda)=\text{Hom}(A_\alpha, \psi^\mu)h_\alpha^{\mu\lambda}$ implies that for each α, there exists, a unique map ε_α of $\varinjlim_{\lambda\in L}\text{Hom}(A_\alpha, B^\lambda)$ into $\text{Hom}(A_\alpha, B)$ such that, for all $\lambda\in L$,

$$\varepsilon_\alpha h_\alpha^\lambda = \text{Hom}(A_\alpha, \psi^\lambda). \tag{2.6}$$

It is readily seen that $(\varepsilon_\alpha)_{\alpha\in I}$ is a projective system of maps. Its limit $\varepsilon^*=\varprojlim_{\alpha\in I}\varepsilon_\alpha$ is by definition the only map of $\varprojlim_{\alpha\in I}\varinjlim_{\lambda\in L}\text{Hom}(A_\alpha, B^\lambda)$ into $\varprojlim_{\alpha\in I}\text{Hom}(A_\alpha, B)$ such that, for all $\alpha\in I$,

$$l_\alpha\varepsilon^* = \varepsilon_\alpha g_\alpha. \tag{2.7}$$

Lemma 1.2 asserts that there exists a unique bijection s: $\varprojlim_{\alpha\in I}\text{Hom}(A_\alpha, B)\to\text{Hom}(A, B)$, such that for all $\alpha\in I$,

$$\text{Hom}(\varphi_\alpha, B)s = l_\alpha. \tag{2.8}$$

If $v=s\varepsilon^*$, we have, for all $\alpha\in I$,

$$\text{Hom}(\varphi_\alpha, B)v = \varepsilon_\alpha g_\alpha. \tag{2.9}$$

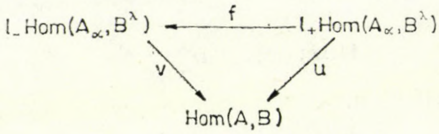


Fig. 2.2

We have thus obtained a triangle (see Fig. 2.2). We shall prove that it is commutative. For every $(\alpha, \lambda) \in I \times L$, we have;

$$\mathrm{Hom}\,(\varphi_\alpha, B) v f h^\lambda = \varepsilon_\alpha g_\alpha f h^\lambda = \varepsilon_\alpha h^\lambda_\alpha g^\lambda_\alpha = \mathrm{Hom}\,(A_\alpha, \psi^\lambda) g^\lambda_\alpha =$$

$$= \mathrm{Hom}\,(A_\alpha, \psi^\lambda)\,\mathrm{Hom}(\varphi_\alpha, B^\lambda) r^\lambda = \mathrm{Hom}\,(\varphi_\alpha \psi^\lambda) r^\lambda = \mathrm{Hom}\,(\varphi_\alpha, B)\,\mathrm{Hom}\,(A, \psi^\lambda) r^\lambda =$$

$$= \mathrm{Hom}\,(\varphi_\alpha, B) u h^\lambda.$$

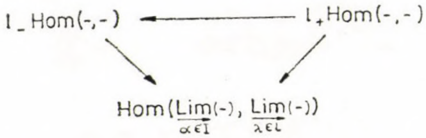
By Corollary 1.2 $\big(\mathrm{Hom}\,(\varphi_\alpha, B)\big)_{\alpha \in I}$ is monomorphic, while $(h^\lambda)_{\lambda \in L}$ is epimorphic hence

(2.10) $$vf = u.$$

LEMMA 2.2. *Let $(A_\alpha, \varphi_{\beta\alpha})_I$ and $(B^\lambda, \psi^{\mu\lambda})_L$ be two inductive systems in a category $\mathscr{C}$ having inductive limits. If $A = \varinjlim_{\alpha \in I} A_\alpha$ and $B = \varinjlim_{\lambda \in L} B^\lambda$, then Fig. 2.2 is a commutative diagram.*

It is easy to prove that the maps v and u are functorial, and thus;

THEOREM 2.1. *Let $\mathscr{C}$ be a category having inductive limits. Then we have the following commutative diagram of bifunctors and of functorial morphisms of $\mathscr{C} \times \mathscr{C}$ into the category of sets:*



If $\mathscr{C}$ is the category ${}_R\mathrm{Mod}$, all the sets considered are abelian groups and maps are homomorphisms.

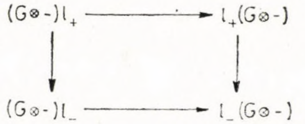
If now the system $(A_\alpha, \varphi_{\beta\alpha})_I$ is composed of finitely generated projective modules and L is directed, then each of the ε_α is an isomorphism. Hence so is ε^* and finally so is v. If each of the A_α is finitely generated projective, we cannot assert that $A = \varinjlim_{\alpha \in I} A_\alpha$ is also finitely generated projective. But of A is also finitely generated projective, ε is an isomorphism, hence so is u. We get:

COROLLARY 2.1. *If $(A_\alpha, \varphi_{\beta\alpha})_I$ is an inductive system of finitely generated projective R-modules, whose limit is also finitely generated projective, and $(B^\lambda, \psi^{\mu\lambda})_L$ is an inductive system of R-modules over a directed set L, the canonical map f of the system* $\big(\mathrm{Hom}\,(A_\alpha, B^\lambda),\ \mathrm{Hom}\,(\varphi_{\beta\alpha}, \psi^{\mu\lambda})\big)_{I \times L}$ *is an isomorphism.*

Such inductive system always exists, since we may take the ring R to be semisimple artinian, then every R-module is projective.

3. The tensor functor and mixed limits

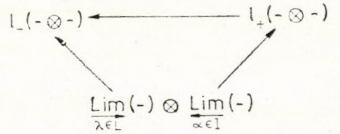
THEOREM 3.1. *Given a right R-module G, we have the following commutative diagram of covariant functors and of functorial morphisms of* $\mathcal{M}(I\times L, {}_R\mathrm{Mod})$ *into the category Ab of Abelian groups:*



The proof is similar to that of Theorem 1.1.

Let now $(A_\alpha, \varphi_{\alpha\beta})_I$ be a projective system in the category ${}_R\mathrm{Mod}$ of left R-modules, and $(B^\lambda, \psi^{\mu\lambda})_L$ be an inductive system in the category Mod_R of right R-modules. Then $(B^\lambda\otimes A_\alpha, \psi^{\mu\lambda}\otimes\varphi_{\alpha\beta})_{I\times L}$ is clearly a mixed system of abelian groups. We can prove as in Theorem 2.1.

THEOREM 3.2. *We have the following commutative diagram of bifunctors and of functorial morphisms of* $\mathrm{Mod}_R\times{}_R\mathrm{Mod}$ *into Ab:*



References

[1] N. BOURBAKI, *Elements de Mathématique, Algèbre*, Hermann (Paris, 1970).
[2] C. G. CHEHATA, M. EL GENDY and I. A. ASSEM, Limits of bifunctors into a category (to appear).
[3] J. P. LAFON, *Les formalismes fondamentaux de l'algèbre commutative*, Hermann (Paris, 1974).

(Received August 29, 1977)

FACULTY OF SCIENCE,
UNIVERSITY OF ALEXANDRIA,
ALEXANDRIA,
EGYPT.

AND

FACULTY OF ENGINEERING,
UNIVERSITY OF ALEXANDRIA,
ALEXANDRIA,
EGYPT.

Acta Mathematica Academiae Scietiarum Hungaricae
Tomus 34 (3—4), (1979), 233—237.

A REMARK ON STRASSEN'S THEOREM

By
I. BERKES (Budapest)

Let $X_1, X_2, \ldots$ be i.i.d. random variables such that $EX_1=0$, $EX_1^2=1$ and put $S_n=\sum_{i=1}^{n} X_i$ $(S_0=0)$. Let $f(n)>0$ be any function (defined on positive integers) and let $\psi_n^{(f)}(t)$ denote the random function in $[0, 1]$ which is linear in the subintervals $[(k-1)/n, k/n]$ $(1\leqq k\leqq n)$ and $\psi_n^{(f)}(k/n)=f(n)^{-1}S_k$ $(0\leqq k\leqq n)$. STRASSEN'S celebrated theorem (see [5]) states that if $f(n)=(2n\log\log n)^{1/2}$ then the sequence $\psi_n^{(f)}(t)$ is relatively compact in $C[0, 1]$ with probability one and the set of its norm-limit points coincides with the set

$$K=\left\{x(t)\colon x(t) \text{ is absolutely continuous in } [0,1], x(0)=0 \text{ and } \int_0^1 \dot{x}(t)^2\,dt\leqq 1\right\}.$$

The purpose of the present paper is to prove the following

THEOREM. *Let us suppose $f(n)/n^{1/2}\to\infty$, $f(n)/(2n\log\log n)^{1/2}\downarrow 0$.*[1] *Then, with probability one, the derived set of $\psi_n^{(f)}(t)$ coincides with the set of all continuous functions $x(t)$ $(0\leqq t\leqq 1)$ such that $x(0)=0$. The same conclusion holds if $f(n)=\sqrt{n}$.*[2]

COROLLARY. *Let us assume $f(n)/n^{1/2}\uparrow\infty$, $f(n)/(2n\log\log n)^{1/2}\downarrow 0$ (or $f(n)=\sqrt{n}$) and suppose in addition that $f(n)$ is regularly varying with exponent $1/2$.*[3] *Let (a, b) be any interval and let ν_n denote the number of integers k, $1\leqq k\leqq n$ such that $a<S_k/f(k)<b$. Then we have $\liminf_{n\to\infty}\nu_n/n=0$, $\limsup_{n\to\infty}\nu_n/n=1$ with probability one.*

PROOF OF THE THEOREM. Let $f(n)/n^{1/2}\to\infty$, $f(n)/(2n\log\log n)^{1/2}\downarrow 0$ and put $\xi_n^{(f)}(t)=f(n)^{-1}\xi(nt)$ $(0\leqq t\leqq 1)$ where ξ is a standard Wiener-process with continuous paths. Following Strassen, we first prove that the statement of the theorem is valid for the sequence $\xi_n^{(f)}(t)$ instead of $\psi_n^{(f)}(t)$. Let H denote the set of all continuous functions $x(t)$ $(0\leqq t\leqq 1)$ such that $x(0)=0$. Since there exists a countable subset $\{x_n(t)\}$ of H which consists of continuously differentiable functions and which is dense in H, it is sufficient to prove that

$$P\left(\sup_{0\leqq t\leqq 1}|\xi_n^{(f)}(t)-x(t)|<5\varepsilon \text{ i.o.}\right)=1 \tag{1}$$

[1] The symbols $\downarrow$ and $\uparrow$ mean convergence monotonically decreasing and monotonically increasing, respectively.

[2] For a result related to the case $f(n)=\sqrt{n}$ of our theorem, see [4].

[3] For a definition see [2] p. 269.

for every $\varepsilon>0$ and for every function $x(t)$ which is continuously differentiable in $[0, 1]$ and $x(0)=0$. Let M denote a constant such that $|\dot{x}(t)|\leqq M$, let us choose an integer $m\geqq 2$ for which $\varepsilon^2 m\geqq 64M^2$, $M/m<\varepsilon$ and put

$$A_n = \left\{\left|\xi_n^{(f)}(i/m)-\xi_n^{(f)}((i-1)/m)-\Delta x_i\right| < \varepsilon/m \quad \text{for} \quad i = 2, 3, \dots, m\right\},$$

$$B_n = \Big\{ \sup_{\substack{0\leqq t<t'\leqq 1 \\ |t-t'|\leqq 1/m}} |\xi_n^{(f)}(t)-\xi_n^{(f)}(t')| > \varepsilon\Big\}$$

where $\Delta x_i=x(i/m)-x((i-1)/m)$. In the same way as in [5] we get (using the inequality

$$\int_a^b e^{-x^2/2}\,dx \geqq \frac{1}{b}\, e^{-a^2/2}(1-e^{-(b^2-a^2)/2})$$

and the fact that $|\Delta x_i|\leqq M/m$)

$$P(A_n) = \prod_{i=2}^{m} \int_{(\Delta x_i-\varepsilon/m)\sqrt{mg(n)}}^{(\Delta x_i+\varepsilon/m)\sqrt{mg(n)}} \frac{1}{\sqrt{2\pi}}\, e^{-s^2/2}\,ds \geqq \tag{2}$$

$$\geqq \prod_{i=2}^{m} \int_{|\Delta x_i|\sqrt{mg(n)}}^{(|\Delta x_i|+\varepsilon/m)\sqrt{mg(n)}} \frac{1}{\sqrt{2\pi}}\, e^{-s^2/2}\,ds \geqq C\frac{1}{g(n)^{\frac{m-1}{2}}}\, e^{-\frac{1}{2}g(n)\sum\limits_{i=2}^{m} m(\Delta x_i)^2} \geqq$$

$$\geqq C\frac{1}{g(n)^{\frac{m-1}{2}}}\, e^{-\frac{1}{2}g(n)M^2} \geqq Ce^{-M^2 g(n)}$$

for sufficiently large n; here $g(n)=f^2(n)/n$ and C is a positive constant independent of n. On the other hand, by the choice of m and well-known properties of the Wiener-Process we have

$$P(B_n) \leqq \sum_{i=1}^{m} P\Big(\sup_{(i-1)/m\leqq t\leqq i/m} \left|\xi_n^{(f)}(t)-\xi_n^{(f)}((i-1)/m)\right| > \varepsilon/4\Big) \leqq \tag{3}$$

$$\leqq 2\sum_{i=1}^{m} P\left(\left|\xi_n^{(f)}(i/m)-\xi_n^{(f)}((i-1)/m)\right| > \varepsilon/4\right) = 4m\left(1-\Phi\left(\frac{\varepsilon}{4}\sqrt{mg(n)}\right)\right) \leqq$$

$$\leqq 4m(2\pi)^{-1/2}\left(\frac{\varepsilon}{4}\sqrt{mg(n)}\right)^{-1} e^{-\frac{\varepsilon^2 m}{32}g(n)} \leqq e^{-2M^2 g(n)}$$

for sufficiently large n. Let us now put $a_k=e^{M^2 g(m^k)}$. The assumption made on $f(n)$ implies that $g(n)=\log\log n\cdot h(n)$ where $h(n)\downarrow 0$, thus $a_k=\exp(M^2\log\log m^k\cdot \cdot h(m^k))=(k\log m)^{M^2 h(m^k)}$. Hence we get

$$a_{k+1} = ((k+1)\log m)^{M^2 h(m^{k+1})} \leqq ((k+1)\log m)^{M^2 h(m^k)} =$$

$$= a_k\left(1+\frac{1}{k}\right)^{M^2 h(m^k)} < a_k\left(1+\frac{1}{k}\right) < a_k+\frac{1}{k}(k\log m)^{1/2} < a_k+1/2$$

for sufficiently large k. We can therefore choose a sequence $l_1<l_2<\dots$ of positive integers such that $a_{l_k}\sim k$ if $k\to\infty$[4] and if we put $n_k=m^{l_k}$ then we have $e^{M^2g(n_k)}\sim k$. By virtue of (2), (3) this implies that $\sum_{k=1}^{\infty}P(A_{n_k})=\infty$, $\sum_{k=1}^{\infty}P(B_{n_k})<\infty$ and since the events A_{n_k} (actually also the events A_{m^k}) are independent, we have $P(A_{n_k}\text{ i.o.})=1$, $P(B_{n_k}\text{ i.o.})=0$. To prove (1) it remains to observe that the events A_n and $\bar{B}_n$[5] together imply the event $\sup_{0\le t\le 1}|\xi_n^{(f)}(t)-x(t)|<5\varepsilon$ (note that $\sup_{|t-t'|\le 1/m}|x(t)-x(t')|\le M/m<\varepsilon$) and thus we have by the above relations

$$(4)\qquad P\big(\sup_{0\le t\le 1}|\xi_{n_k}^{(f)}(t)-x(t)|<5\varepsilon\ \text{i.o.}\big)=1.$$

Let now $X_1, X_2, \dots$ be i.i.d. random variables such that $EX_1=0$, $EX_1^2=1$. We shall show that

$$(5)\qquad P\big(\sup_{0\le t\le 1}|\psi_n^{(f)}(t)-x(t)|<7\varepsilon\ \text{i.o.}\big)=1$$

for any $\varepsilon>0$ and for every continuously differentiable $x(t)$ such that $x(0)=0$. In view of the Skorokhod representation theorem we can assume, without loss of generality, that $X_k=\xi(\tau_1+\dots+\tau_k)-\xi(\tau_1+\dots+\tau_{k-1})$ for every $k\ge 1$ where $\xi(t)$ is a standard Wiener process with continuous paths and $\tau_1, \tau_2, \dots$ are i.i.d. nonnegative random variables such that $E\tau_1=1$. Let us fix $x(t)$, ε and let $M, m, \{n_k\}$ denote the same as above. By virtue of relation (4), for the proof of (5) it suffices to show that almost surely

$$(6)\qquad \sup_{0\le t\le 1}|\psi_{n_k}^{(f)}(t)-\xi_{n_k}^{(f)}(t)|<2\varepsilon$$

for sufficiently large k. Let us now observe that the above form of the X_k's and the piecewise linearity of $\psi_n^{(f)}(t)$ imply that $\psi_n^{(f)}(t)$ lies between $f(n)^{-1}\xi(\sum_{i=1}^{[nt]}\tau_i)$ and $f(n)^{-1}\xi(\sum_{i=1}^{[nt]+1}\tau_i)$ for every $0\le t\le 1$ and thus

$$(7)\qquad |\psi_n^{(f)}(t)-\xi_n^{(f)}(t)|\le f(n)^{-1}\max\left\{\left|\xi\left(\sum_{i=1}^{[nt]}\tau_i\right)-\xi(nt)\right|,\ \left|\xi\left(\sum_{i=1}^{[nt]+1}\tau_i\right)-\xi(nt)\right|\right\}=$$

$$=\max\left\{\left|\xi_n^{(f)}\left(\frac{1}{n}\sum_{i=1}^{[nt]}\tau_i\right)-\xi_n^{(f)}(t)\right|,\ \left|\xi_n^{(f)}\left(\frac{1}{n}\sum_{i=1}^{[nt]+1}\tau_i\right)-\xi_n^{(f)}(t)\right|\right\}.$$

(Here the quantities $n^{-1}\sum_{i=1}^{[nt]}\tau_i$ and $n^{-1}\sum_{i=1}^{[nt]+1}\tau_i$ can lie outside of the interval [0, 1] but this does not cause trouble because $\xi_n^{(f)}(t)$ is defined for $0\le t<+\infty$.) The strong law of large numbers implies that for almost all ω there exists an $L=L(\omega)$ such that for $nt\ge L$ we have

$$0\le nt(1-1/2m)<\sum_{i=1}^{[nt]}\tau_i\le\sum_{i=1}^{[nt]+1}\tau_i<nt(1+1/2m)\le 2nt$$

[4] The symbol $a_n\sim b_n$ means $\lim_{n\to\infty}a_n/b_n=1$.

[5] $\bar{B}_n$ denotes the complement of B_n.

and hence

$$(8)\qquad |\psi_n^{(f)}(t)-\xi_n^{(f)}(t)| \leqq \sup_{\substack{0\leqq s<s'\leqq 2\\ |s-s'|\leqq 1/m}} |\xi_n^{(f)}(s)-\xi_n^{(f)}(s')| \quad 0\leqq t\leqq 1,\ nt\geqq L.$$

On the other hand, for a fixed ω, $\xi\left(\sum_{i=1}^{[nt]}\tau_i\right)$, $\xi\left(\sum_{i=1}^{[nt]+1}\tau_i\right)$ take only finitely many values and $\xi(nt)$ is bounded for $0\leqq nt\leqq L$, thus the first relation of (7) gives

$$(9)\qquad |\psi_n^{(f)}(t)-\xi_n^{(f)}(t)|\leqq f(n)^{-1}R \quad (0\leqq nt\leqq L)$$

where R depends only on ω. By (8) and (9) we have almost surely for sufficiently large n

$$|\psi_n^{(f)}(t)-\xi_n^{(f)}(t)|\leqq \sup_{\substack{0\leqq s<s'\leqq 2\\ |s-s'|\leqq 1/m}} |\xi_n^{(f)}(s)-\xi_n^{(f)}(s')|+\varepsilon = T_n+\varepsilon \quad (0\leqq t\leqq 1).$$

Put $C_n=\{T_n\geqq\varepsilon\}$. Exactly in the same way as in (3), we get $P(C_n)\leqq 2e^{-2M^2g(n)}$. Hence $e^{M^2g(n_k)}\sim k$ and the Borel—Cantelli lemma yields $T_{n_k}<\varepsilon$ almost surely for sufficiently large k. The latter relation, together with (10), evidently implies (6) and the proof is complete.

In the case $f(n)=\sqrt{n}$ the statement of the theorem is almost evident. Let $x(t)$ be a function which is continuous in [0, 1] and $x(0)=0$. Put

$$D_n=\Big\{\sup_{0\leqq t\leqq 1}|\psi_n^{(f)}(t)-x(t)|<\varepsilon\Big\},\quad D=\Big\{\sup_{0\leqq t\leqq 1}|\xi(t)-x(t)|<\varepsilon\Big\}$$

where $\xi(t)$ is a standard Wiener process with continuous paths and $\varepsilon>0$ is a continuity point of the d.f. of the random variable $\sup_{0\leqq t\leqq 1}|\xi(t)-x(t)|$. Since there exist arbitrary small ε's having this property, it suffices to show that $P(D_n \text{ i.o.})=1$ for any such value of ε. By Donsker's invariance principle (see [1] p. 68) we have $\lim_{n\to\infty}P(D_n)=P(D)$ and here the limit $P(D)$ is known to be positive (see [3], p. 30). Therefore we have $P(D_n \text{ i.o.})>0$ and the proof is completed by applying the zero-one law.

To obtain the corollary let us write $f(n)=n^{1/2}\varphi(n)$ where $\varphi(n)$ is a slowly varying function. We show that, for any given $0<\delta<1$, the event

$$E_n=\{af(k)<S_k<bf(k) \text{ for } \delta n\leqq k\leqq n\}$$

ovcurs for infinitely many n with probability one; this statement evidently implies $\limsup_{n\to\infty} v_n/n=1$. Now E_n can be written as follows:

$$E_n=\left\{a\sqrt{\frac{k}{n}}\frac{\varphi(k)}{\varphi(n)}<\frac{S_k}{f(n)}<b\sqrt{\frac{k}{n}}\frac{\varphi(k)}{\varphi(n)} \quad \text{for} \quad \delta n\leqq k\leqq n\right\}.$$

Since $\varphi(n)$ is increasing and slowly varying, for any given $\varepsilon>0$ there exists an $n_0=n_0(\varepsilon)$ such that for $\delta n\leqq k\leqq n$, $n\geqq n_0$ we have $1-\varepsilon<\varphi(k)/\varphi(n)\leqq 1$. Hence, applying the theorem for $x(t)=c\sqrt{t}$ where $a<c<b$ we get that E_n occurs infinitely often. The relation $\liminf_{n\to\infty} v_n/n=0$ is trivial from the fact that $\limsup_{n\to\infty} v_n/n=1$ is valid also for intervals $(a'b')$ disjoint from (a, b).

References

[1] P. BILLINGSLEY, *Convergence of probability measures*. Wiley (New York, 1968).
[2] W. FELLER, *An introduction to probability theory and its applications*, Vol. II. Wiley (New York, 1966).
[3] D. FREEDMAN, *Brownian motion and diffusion*. Holden—Day (San Francisco, 1971).
[4] H. ROOTZÉN, Rényi — mixing and functional limit theorems for martingales, *Proc. Seventh Prague Conference and European Meet. Statist*. 1974 (to appear).
[5] V. STRASSEN, An invariance principle for the law of the iterated logarithm, *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, **3** (1964), 211—226.

(Received September 20, 1977)

MATHEMATICAL INSTITUTE
OF THE HUNGARIAN ACADEMY OF SCIENCES,
1053 BUDAPEST, REÁLTANODA U. 13—15.

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (3—4), (1979), 239—243.

MATRICES WHICH MAXIMISE ANY ANALYTIC FUNCTION

By

N. J. YOUNG (Glasgow)

Rapid convergence of an iterative process in numerical linear algebra often depends on whether the powers A^m of some $n \times n$ matrix A become small (in some sense) quickly enough. Such considerations led VL. PTÁK [2] to raise the question as to how large $\|A^m\|$ could be for $n \times n$ matrices A satisfying $\|A\| \leqq 1$ and $p(A)=0$, where p is a polynomial of degree n having roots of absolute value less than one and $\|\cdot\|$ denotes the operator norm on n-dimensional Hilbert space. He proved that the maximum value is attained when A is the restriction of the unilateral backward shift S on l_2 to a certain n-dimensional subspace K_p of l_2: in fact $K_p = \{x \in l_2;\ p(S)x=0\}$.

This result has been greatly generalised by B. SZ.-NAGY [3]. He has shown that if φ, ψ are bounded analytic functions in the open unit disc then the maximum of $\|\psi(A)\|$, over all completely non-unitary Hilbert space operators A satisfying $\|A\| \leqq 1$ and $\varphi(A)=0$, is attained when A is the restriction of S to $K_\varphi = \{x \in l_2:\ \varphi(S)x=0\}$, where S again denotes the backward shift on l_2.

Sz.-Nagy's theorem has a very striking qualitative consequence; there is an extremal operator for the problem in question which depends only on φ and is independent of ψ. In the finite-dimensional case it follows that to each φ there corresponds a matrix with a very remarkable extremal property, and it is natural to ask how to find such a matrix. Let us consider the problem of determining

$$\sup \{\|f(A)\|:\ A \in M_n(\mathbf{C}),\ \|A\| \leqq 1,\ p(A) = 0\} \tag{1}$$

where p is a monic polynomial of degree n whose roots lie in the open unit disc and f is a function analytic in a neighbourhood of the roots of p. Sz.-Nagy's result shows[1] that there is a matrix A_p, satisfying the constraints of (1), for which $\|f(A)\|$ is maximised for every f. The purpose of the present note is to give a formula allowing such an extremal matrix A_p to be calculated relatively easily, at least when p has real coefficients.

The author [4] has already given two ways of finding such a matrix A_p, but both are rather unsatisfactory from a practical point of view, in that they require elaborate computation. They are as follows. Let T be the companion matrix of p (see (8) below) and let Y be the solution of the equation

$$Y - T^*YT = E \tag{2}$$

[1] Although the theorem is only stated for functions f analytic in the whole of the unit disc Sz.-Nagy's proof applies equally well to functions f of the type discussed here.

where E is the $n\times n$ matrix having (1, 1) entry 1 and all other entries zero. It is clear that (2) has the unique solution

$$Y=\sum_{s=0}^{\infty}T^{*s}ET^{s}=\sum_{s=0}^{\infty}T^{*ns}T^{ns}, \tag{3}$$

the convergence of the series being ensured by the hypothesis that the roots of p, which are the eigenvalues of T, be of absolute value less than one. (3) shows that $Y\geqq I$, so that $Y^{1/2}$ and $Y^{-1/2}$ can be formed. By [4], Theorem 2,

$$K=Y^{1/2}TY^{-1/2} \tag{4}$$

is an extremal matrix for (1).

The difficulty with calculating K lies in obtaining Y and extracting $Y^{1/2}$ and $Y^{-1/2}$. In the second method we let S denote the $n\times n$ shift matrix ($S=[\delta_{i(k-1)}]$) and let D be an $n\times n$ matrix of norm less than one. Then, by the corollary to Theorem 3 of [4], if

$$\Gamma=(I-DD^*)^{-1/2}(S+D)(I+D^*S)^{-1}(I-D^*D)^{1/2} \tag{5}$$

is such that $p(\Gamma)=0$, then Γ is an extremal for (1). All that is needed, therefore, is to find D such that $\|D\|<1$ and $p(\Gamma)=0$. The method suggested in [4] was to pick D to be the diagonal matrix having the roots of p on its main diagonal: Γ is then an upper triangular matrix with diagonal D. If the roots of p are known (say, $\varrho_1, \dots, \varrho_n$) then an explicit formula for Γ can easily be given: in fact

$$\Gamma=\begin{bmatrix}\varrho_1 & s_1s_2 & -s_1\bar\varrho_2s_3 & s_1\bar\varrho_2\bar\varrho_3s_4 & \dots\\ 0 & \varrho_2 & s_2s_3 & -s_2\bar\varrho_3s_4 & \dots\\ \cdot & \cdot & \cdot & \cdot & \dots\\ 0 & 0 & 0 & 0 & \dots\end{bmatrix} \tag{6}$$

where $s_i=(1-\varrho_i\bar\varrho_i)^{1/2}$. This constitutes a satisfactory solution to the extremal problem provided the ϱ_i are known; when they are not then Γ can only be found by first determining the roots of p, and this is well known to be a lenthy and unstable process.

In the method to be described we produce a matrix D whose entries are given explicitly in terms of the coefficients of p and which is such that Γ in (5) satisfies $p(\Gamma)=0$. In this way we obtain the desired extremal without having to solve a polynomial equation. It is true that it still remains to extract the square root of a positive definite matrix, but this is a considerably speedier and more stable procedure.

Let us write

$$p(\lambda)=\lambda^n-\alpha_n\lambda^{n-1}-\alpha_{n-1}\lambda^{n-2}-\dots-\alpha_1 \tag{7}$$

so that

$$T=\begin{bmatrix}0 & 1 & 0 & \dots & 0\\ 0 & 0 & 1 & \dots & 0\\ \cdot & \cdot & \cdot & \dots & \cdot\\ \alpha_1 & \alpha_2 & \alpha_3 & \dots & \alpha_n\end{bmatrix} \tag{8}$$

and let

(9)
$$B = \begin{bmatrix} 0 & 0 & \dots 0 & \beta_1 \\ 0 & 0 & \dots \beta_1 & \beta_2 \\ . & . & \dots . & . \\ 0 & \beta_1 & \dots \beta_{n-2} & \beta_{n-1} \\ \beta_1 & \beta_2 & \dots \beta_{n-1} & \beta_n \end{bmatrix}$$

where the β_i are given by the equations

(10)
$$\begin{cases} \beta_1 = \alpha_1 \\ \beta_2 = \alpha_2 + \beta_1 \alpha_n \\ \beta_3 = \alpha_3 + \beta_1 \alpha_{n-1} + \beta_2 \alpha_n \\ . \quad . \quad . \quad . \\ \beta_n = \alpha_n + \beta_1 \alpha_2 + \dots + \beta_{n-1} \alpha_n . \end{cases}$$

B is a lower cross-triangular Hankel matrix.

THEOREM. *If p has real coefficients and the roots of p have absolute value less than one, the supremum in* (1) *is attained when*

(11)
$$A = (I - B^2)^{-1/2} T (I - B^2)^{1/2}.$$

PROOF. Let S denote the shift matrix $[\delta_{i(k-1)}]$ as previously. It will be left to the reader to verify that

(12)
$$S + B = T(I + BS)$$

(indeed (12) is equivalent to (10)). Let us suppose for the moment that $I+BS$ and $I-B^2$ are non-singular; we prove that $Y=(I-B^2)^{-1}$ is a solution of (2). In view of (12), (2) can be written

(13)
$$Y - (I + S^* B)^{-1} (S^* + B) Y (S + B)(I + BS)^{-1} = E$$

(the fact that the α_i, and hence the β_i, are real implies that $B=B^*$). Since

$$(I + S^* B) E (I + BS) = E,$$

(13) is equivalent to

(14)
$$(I + S^* B) Y (I + BS) - (S^* + B) Y (S + B) = E.$$

If Y commutes with B then, on expanding the two products in (14), we find that four terms cancel, and the equation simplifies to

(15)
$$Y - BYB - S^* (Y - BYB) S = E.$$

Since $I - S^* S = E$, (15) is satisfied by $Y = (I - B^2)^{-1}$.

It remains to be proved that $I+BS$ and $I-B^2$ are non-singular. It obviously suffices to prove that $\|B\| < 1$. For this purpose we appeal to the following result.

LEMMA. *Let $A_1, A_2, \dots, A_n$ be commuting $n \times n$ matrices, each of norm one and spectral radius less than one. Then $\|A_1 A_2 \dots A_n\| < 1$.*

This was stated and proved by PTÁK [1] for the case in which the A_i are all equal.

PROOF. Let $V_i=\{x\in\mathbf{C}^n:\ \|A_1A_2\dots A_ix\|=\|x\|\}$, $1\leqq i\leqq n$, and let $V_0=\mathbf{C}^n$. Each V_i is a subspace of $\mathbf{C}^n$ and, by reason of the commutativity assumption, $V_{i+1}\subseteqq V_i$. Note that if $x\in V_{i+1}$ then necessarily

$$\|A_{i+1}x\|=\|x\| \tag{16}$$

and so

$$\|A_1A_2\dots A_i(A_{i+1}x)\|=\|x\|=\|A_{i+1}x\|;$$

thus $A_{i+1}V_{i+1}\subseteqq V_i$. If, for any i, $V_i=V_{i+1}\neq\{0\}$, then V_{i+1} is invariant under A_{i+1} and hence contains an eigenvector, and it follows from (16) that A_{i+1} has an eigenvalue of unit modulus, contrary to supposition. Thus the chain $V_0\supseteqq V_1\supseteqq\dots$ $\dots\supseteqq V_n$ descends strictly till it reaches $\{0\}$, and we have $V_n=\{0\}$, which is the desired condusion.

To apply the Lemma introduce the polynomial

$$q(\lambda)=\lambda^np(1/\lambda)=1-\alpha_n\lambda-\alpha_{n-1}\lambda^2-\dots-\alpha_1\lambda^n, \tag{17}$$

and let J denote the matrix with ones on the priucipal cross-diagonal $(i+k=n+1)$ and zeros everywhere else, so that the effect of postmultiplying by J is to reverse the order of the columns. We have

$$q(S)J=\begin{bmatrix}-\alpha_2 & -\alpha_3 & \dots & -\alpha_n & 1\\ -\alpha_3 & -\alpha_4 & \dots & 1 & 0\\ \cdot & \cdot & \dots & \cdot & \cdot\\ 1 & 0 & \dots & 0 & 0\end{bmatrix},$$

and now, using (9) and (10), we find that $q(S)JB=-p(S)$. Since $q(S)$ is nonsingular,

$$B=-Jq(S)^{-1}p(S). \tag{18}$$

If $p(\lambda)=(\lambda-\varrho_1)(\lambda-\varrho_2)\dots(\lambda-\varrho_n)$ then

$$q(\lambda)=(1-\varrho_1\lambda)(1-\varrho_2\lambda)\dots(1-\varrho_n\lambda)=(1-\bar\varrho_1\lambda)(1-\bar\varrho_2\lambda)\dots(1-\bar\varrho_n\lambda),$$

the last step depending on the assumption that the coefficients in p be real. Thus $q(S)^{-1}p(S)=A_1A_2\dots A_n$ where $A_i=(I-\bar\varrho_iS)^{-1}(S-\varrho_iI)$.
Since $\|S\|=1$, any matrix obtained from S by the application of a Möbius transformation also has norm 1, and so $\|A_i\|=1$. The A_i commute, and A_i has the unique eigenvalue $-\varrho_i$, which has modulus less than one, by hypothesis. Hence, by the Lemma, $\|A_1A_2\dots A_n\|<1$, and so $\|B\|<1$ as desired.

We have now shown that the solution of (2) is $Y=(I-B^2)^{-1}$, and the theorem follows from (4).

I leave open the problem of extending the theorem to the complex case. The proof above suggests what the right extension might be: namely, that an extremal for (1) is given by Γ in (5) where $D=-Jq^*(S)^{-1}p(S)$ and q^* is given by $q^*(\lambda)=\lambda^np(1/\bar\lambda)^-$.

One of the main questions with which PTÁK was concerned in [2] was to find the supremum in (1) in the case where $f(A)=A^n$, and I conclude with an observation about this. To find the supremum we could calculate K using the theorem above, form K^n and calculate its norm (for instance, by finding the spectral radius of $K^{*n}K^n$). There is, however, a short cut which saves a great deal of computation: in particular, it avoids the need to extract any square roots of matrices. From the definition of K we have

$$I-K^*K=I-Y^{-1/2}T^*YTY^{-1/2}=Y^{-1/2}(Y-T^*YT)Y^{-1/2}=Y^{-1/2}EY^{-1/2},$$

the last step following from the definition (2) of Y. Multiplying fore and aft by K^{*j} and K^j, and summing from $j=0$ to $n-1$, we find

$$I-K^{*n}K^n=Y^{-1/2}\left(\sum_{j=0}^{n-1}T^{*j}ET^j\right)Y^{-1/2}=Y^{-1/2}IY^{-1/2}=Y^{-1}=I-B^2.$$

Thus $K^{*n}K^n=B^2$ and so $\|K^n\|=\|B\|$. Since B is real and symmetric its norm equals its spectral radius; we can therefore find the supreme in (1) simply by forming B and applying one of the standard routines which calculates the eigenvalue of largest modulus for a real symmetric matrix. This makes the problem very easy to solve with the aid of a computer.

References

[1] V. PTAK, Norms and the spectral radius of matrices, *Czech. Math. J.*, **87** (1962), 555—557.
[2] V. PTAK, Spectral radius, norms of iterates, and the critical exponent, *Linear Algebra and it's Applications*, **1** (1968), 245—260.
[3] B. SZ.-NAGY, Sur la norme des fonctions de certains opérateurs, *Acta Math. Acad. Sci. Hungar.*, **20** (1969), 331—334.
[4] N. J. YOUNG, Analytic programmes in matrix algebras, *Proc. London Math. Soc.*, **36** (1978), 226—242.

(Received October 14, 1977)

UNIVERSITY OF GLASGOW
DEPARTMENT OF MATHEMATICS
GLASGOW, G12 8QW
SCOTLAND

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (3—4), (1979), 245—251.

ÜBER DIE STRUKTUR LINEAR KOMPAKTER RINGE. II

Von

PHAM NGOC ANH (Budapest)

§ 1. Einleitung

Dies ist die Fortsetzung und Ergänzung meiner Arbeit [1]. Wie auch in [1], werde ich die Struktur linear kompakter Ringe untersuchen. Die von uns betrachteten Ringe sind sämtlich assoziativ. Das Radikal eines Ringes wird immer das jacobsonsche Radikal bedeuten. Wir übernehmen in dieser Arbeit die Bezeichnungen, Begriffe und Abkürzungen aus [1], ein l. k. Ring ist also z. B. ein linear kompakter Ring und ein i. e. S. l. k. Modul ein im engeren Sinne linear kompakter Modul (d. h. ein vollständiger linear topologischer Modul, in dessen diskreten Restklassenmoduln die Minimalbedingung für Untermoduln gilt.). Da der Querstrich zur Bezeichnung von Restklassenringen dient, wird zur Bezeichnung des topologischen Abschlusses einer Menge A in einem topologischen Raum das Symbol cl (A) (closure) benutzt.

In der Strukturtheorie artinscher Ringe spielt der Satz von KERTÉSZ—WIDIGER (Satz 3 in [3] oder (K—W) in [1]), der eine vollständige Beschreibung der Klasse artinscher Ringe mit artinschem Radikal liefert, eine wichtige Rolle. Für linear kompakte Ringe mit Umgebungsbasis des Nullelementes aus Idealen gab Widiger eine ähnliche Kennzeichnung dieser Ringe, deren Radikal selbst ein i. e. S. l. k. Ring ist, wie es in [3] oder in [1] zu finden ist. Diese Ergebnisse von Widiger lassen sich in gewisser Weise mit unseren Methoden zum Abschluß bringen; Jeder l. k. Ring mit s. l. k. Radikal ist eine direkte Summe voller Endomorphismenringe von Vektorräumen über unendlichen Schiefkörpern und eines Ringes, dessen Restklassenring nach dem Radikal direkte Summe voller Endomorphismenringe von Vektorräumen über endlichen Körpern ist. Diese Zerlegung ist eindeutig bis auf topologische Isomorphie. Dieses Resultat ist auch eine Verallgemeinerung des Satzes von Leptin über halbeinfache l. k. Ringe vgl. (XVII) [1].

Ein weiteres wichtiges Resultat des ersten Paragraphen ist der folgende Satz; Ein l. k. Ring mit s. l. k. Radikal hat genau dann die wedderburnsche Zerlegung, wenn die Summe aller Einselemente von endlichen einfachen direkten Summanden des Restklassenringes R/I nach dem Radikal I einen (nicht notwendig idempotenten) Vertreter e mit folgender Eigenschaft hat: Für jede Umgebung U des Nullelementes gibt es endlich viele verschiedene Primzahlen $p_1, p_2, \ldots, p_n$ mit $(p_1, \ldots, p_n)e \in U$.

Hier verstehen wir unter der Wedderburnschen Zerlegung eines topologischen Ringes R die direkte Darstellung $R = S \oplus I$ als Gruppen wobei I das Radikal von R und S ein abgeschlossener Unterring von R sind.

Im zweiten Paragraphen beschreiben wir die Struktur von l. k. Ringen mit MHL-Radikal und erhalten als Nebenprodukt, daß jeder l. k. Ring mit MHL-Radikal spaltbar ist. Hier bedeutet ein MHL-Ring. einen Ring mit Minimalbedingung für Hauptlinksideale.

§ 2. Linear kompakte Ringe mit streng linear kompaktem Radikal

Wir werden in diesem Abschnitt die Verallgemeinerung des Satzes von Kertész und Widiger über artinsche Ringe mit artinschem Radikal für allgemeine linear kompakte Ringe beweisen. Nach dem Satz von Kertész und Widiger ist jeder artinsche Ring mit artinschem Radikal die ringtheoretische direkte Summe endlich vieler voller Matrizenringe über unendlichen Schiefkörpern und eines streng artinschen Ringes. WIDIGER zeigte dies in [11] für den Fall, wenn der Ring und sein Radikal als Ring i. e. S. l. k. mit einer Umgebungsbasis des Nullelementes aus Idealen sind. In [1] haben wir bewiesen, daß die von Widiger erreichten Ergebnisse auch dann gültig bleiben, wenn wir nicht voraussetzen, daß eine aus Idealen bestehende Umgebungsbasis des Nullelementes existiert. Der Preis dafür ist die Bedingung, daß ein sogenanntes vollständiges System idempotenter Elemente mit der Eigenschaft (V) existiere. In diesen Sätzen wird die Struktur l. k. Ringe durch eine Klasseneinteilung der halbeinfachen l. k. Ringe gekennzeichnet. In [1] teilen wir die einfachen l. k. Ringe im zwei Klassen ein; die erste bzw. zweite Klasse besteht aus allen unendlichen bzw. endlichen Ringen. Für artinsche Ringe und l. k. Ringe mit einer Umgebungsbasis des Nullelements aus Idealen ist genau dann ein einfacher Ring der ersten bzw. zweiten Klasse, wenn er ein voller Matrizenring über unendlichem bzw. endlichem Schiefkörper ist. Für einfache l. k. Ringe ist die Situation komplizierter, da der volle Endomorphismenring eines Vektorraumes unendlicher Dimension über endlichem Körper unendlich ist. Dies erklärt, warum wir die Eigenschaft (V) voraussetzen müssen. Es ist bequemer, wenn wir einfache l. k. Ringe nach ihren Grundschiefkörpern klassifizieren. Wir verstehen hier unter einem Grundschiefkörper eines einfachen l. k. Ringes R den bis auf Isomorphie eindeutig bestimmten Schiefkörper, über dem es einen Vektorraum gibt, dessen Endomorphismenring mit R isomorph ist.

Es sei jetzt R ein l. k. Ring mit dem Radikal I, welches bezüglich der von R induzierten Topologie ein streng linear kompakter Ring ist. Nach (XVII) [1] ist $\bar{R}=R/I$ die direkte Summe einfacher l. k. Ringe

$$\bar{R} = \sum_{\mu\in\Gamma^*} R_\mu ,$$

Dann ist jedes $\bar{R}_\mu$ ein voller Endomorphismenring eines Vektorraumes über einem Schiefkörper S_μ. Wir bezeichnen mit $\bar{e}_\mu$ das Einselement von $\bar{R}_\mu$. So ist $e=\sum_{\mu\in\Gamma^*} e_\mu$ das Einselement von $\bar{R}$. Die Menge Γ besteht aus allen Indezes μ in Γ^*, für die die S_μ unendlich sind. Es sei dann $\Gamma'=\Gamma^*-\Gamma$. Wegen (XX) [1] gibt es dann ein System orthogonaler idempotenter Vertreter $\{e_\mu,\ \mu\in\Gamma^*\}$ von $\{\bar{e}_\mu,\ \mu\in\Gamma^*\}$, e von e, so daß

$$e = \sum_{\mu\in\Gamma^*} e_\mu,\ ee_\mu = e_\mu e = e_\mu,\ e_\mu e_\nu = 0,\quad \mu\neq\nu,\ \mu,\nu\in\Gamma^*$$

gilt. Daraus folgt die Summierbarkeit jedes Teilsystemes von $\{e_\mu,\ \mu\in\Gamma^*\}$, insbesondere die Summierbarkeit von $\{e_\mu, \mu\in\Gamma\}$. e^* bezeichne die Summe von $\{e_\mu, \mu\in\Gamma\}e^*= =\sum_{\mu\in\Gamma} e_\mu$. Aus Satz 5 [1] wissen wir, daß $R_\mu=e_\mu Re_\mu$ halbeinfach und $e_\mu I=0$ für alle $\mu\in\Gamma$ sind.

HILFSSATZ 2.1. *Es gilt* $Ie_\mu=0$ *für alle* $\mu\in\Gamma$.

BEWEIS. Wegen der Halbeinfachheit von R_μ ist R_μ ein voller Endomorphismenring eines Vektorraumes über dem unendlichen Schiefkörper S_μ. Daraus folgt $e_\mu = \sum_{\delta \in \Delta\mu} e_{\mu\delta}$, wobei $\{e_{\mu\delta}, \delta \in \Delta_\mu\}$ ein vollständiges System minimaler orthogonaler Idempotente von R_μ bildet und $R_{\mu\delta} = e_{\mu\delta} R_\mu e_{\mu\delta} = e_{\mu\delta} R e_{\mu\delta}$ für jedes $\delta \in \Delta_\mu$ mit S_μ isomorph ist. Für jedes beliebige, aber feste Element a aus I bil det die Abbildung

$$\varphi\colon R_{\mu\delta} \to I, \quad \varphi(x) = ax, \quad x \in R_{\mu\delta}$$

den Schiefkörper $R_{\mu\delta}$ in den $R_{\mu\delta}$-Rechtsmodul I ab. Es zeigt sich unmittelbar, daß φ ein stetiger $R_{\mu\delta}$-Homomorphismus ist. Deshalb ist $\varphi(R_{\mu\delta})$ ein s. l. k. Modul. Da $R_{\mu\delta}$ diskret ist, ist $\varphi(R_{\mu\delta})$ eine Gruppe mit Minimalbedingung für Untergruppen. Weil $R_{\mu\delta}$ ein Schiefkörper ist, ist φ entweder trivial oder injektiv. Das letztere ist unmöglich, da $R_{\mu\delta}$ nach unserer Voraussetzung keine Gruppe mit Minimalbedingung für Untergruppen ist. Also ist $Ie_{\mu\delta}Re_{\mu\delta} = Ie_{\mu\delta} = 0$ für jedes δ aus Δ_μ. Aus $e_\mu = \sum_{\delta \in \Delta\mu} e_{\mu\delta}$ folgt die Gleichung $Ie_\mu = 0$ nach (XXVI) [1].

HILFSSATZ 2.2. *R_μ ist für alle $\mu \in \Gamma$ ein Ideal von R und deshalb abgeschlossen in R.*

BEWEIS. Seien nämlich x, y beliebige Elemente aus R. Dann gilt $\bar{e}_\mu \bar{x} \bar{e}_\mu \bar{y} = \bar{e}_\mu \bar{x} \bar{y} \bar{e}_\mu$, weil $\bar{e}_\mu$ im Zentrum von R liegt. Das besagt aber $e_\mu x e_\mu y - e_\mu x y e_\mu \in I$. Wegen $e_\mu I = 0$ folgt nämlich $e_\mu x e_\mu \bar{y} = e_\mu x y e_\mu \in R_\mu$. Andererseits gilt $\bar{y} \bar{e}_\mu \bar{x} \bar{e}_\mu = \bar{e}_\mu \bar{y} \bar{x} \bar{e}_\mu$, woraus wir $y e_\mu x e_\mu - e_\mu y x e_\mu \in I$ erhalten, und wegen $Ie_\mu = 0$ folgt unmittelbar $y e_\mu x e_\mu = e_\mu y x e_\mu \in R_\mu$. Also ist R_μ ein Ideal von R für jedes $\mu \in \Gamma$. Da R_μ nach (XIX) [1] i. e. S. l. k. ist, ist R_μ in R abgeschlossen. Somit ist der Beweis von Hilfssatz 2.2 erbracht.

Es sei jetzt B das vollständige Urbild von $\sum_{\mu \in \Gamma'} \bar{R}_\mu$ beim natürlichen Epimorphismus von R auf $\bar{R}$. B ist offenbar ein abgeschlossenes Ideal von R. Da für jedes $\mu \in \Gamma e_\mu I = Ie_\mu = 0$ und der Durchschnitt von Γ und Γ' leer sind, gilt $e_\mu B = Be_\mu = 0$ für jedes μ aus Γ. Daraus und aus (XXVI) [1] haben wir $e^*B = Be^* = e^*I = Ie^* = 0$. Wir betrachten den Unterring e^*Re^*. Für irgendzwei Elemente x, y in R haben wir $\bar{e}^* \bar{x} \bar{e}^* \bar{y} = \bar{e}^* \bar{x} \bar{y} \bar{e}^*$ und $\bar{y} \bar{e}^* \bar{x} \bar{e}^* = \bar{e}^* \bar{y} \bar{x} \bar{e}^*$, da $\bar{e}^* = \sum_{\mu \in \Gamma} \bar{e}_\mu$ und die $\bar{e}_\mu$ $(\mu \in \Gamma)$ im Zentrum von $\bar{R}$ liegen. Dies zeigt, daß $e^*xe^*y - e^*xye^*$, $e^*yxe^* - ye^*xe^* \in I$ gilt. Aus $e^*I = Ie^* = 0$ folgen dann $e^*xe^*y = e^*xye^* \in I$ und $ye^*xe^* = e^*yxe^* \in I$. Also ist e^*Re^* ein Ideal von R. Wegen $e^*B = Be^* = Ie^* = e^*I = 0$ sind die Gleichungen $e^*Re^* \cap B = e^*Re^* \cap I = 0$ gültig. Nach (XIX) [1] ist dann e^*Re^* ein halbeinfacher l. k. Ring. Daraus und wegen $e^* = \sum_{\mu \in \Gamma} e_\mu$ ist e^*Re^* die direkte Summe von Ringen R_μ mit $\mu \in \Gamma$.

Ist jetzt $x \in R$ beliebig, so $x = e^*xe^* + (x - e^*xe^*)$ wobei $e^*xe^* \in e^*Re^*$ ist und $x - e^*xe^*$ offensichtlich in B liegt. Dies zeigt, daß R mit der ringtheoretischen direkten Summe $e^*Re^* \boxplus B$ isomorph im algebraischen Sinne ist.

Die Isomorphie ist auch topologisch. Es sei nämlich U_1 bzw. U_2 eine beliebige Umgebung von 0 in e^*Re^* bzw. in B. Es ist klar, dass $U_1 = U \cap e^*Re^*$, $U_2 = U' \cap B$ mit gewissen Umgebungen U, U' aus R erfüllt sind. Deshalb genügt es sich auf den Fall $U_1 = U_0 \cap e^*Re^*$ und $U_2 = U_0 \cap B (U_0 = U \cap U')$ zu beschränken. Wählen wir dann die Umgebungen W, V_1, V_2, sodass die folgenden Bedingungen

$$e^*V_1e^* \subset U_0, \quad V_2 - e^*V_2e^* \subset U_0, \quad W \subseteqq V_1 \cap V_2$$

gelten. Für ein beliebiges Element $w \in W$ haben wir dann

$$e^* we^* \in e^* V_1 e^* \cap e^* Re^* \subseteqq U_0 \cap e^* Re^* = U_1$$

$$w - e^* we^* \in (V_2 - e^* V_2 e^*) \cap B \subseteqq U_0 \cap B = U_2.$$

Folglich gilt $w = e^* we^* + (w - e^* we^*) \in U_1 + U_2$ für alle $w \in W$.

Bisher haben wir also bewiesen, daß es für alle l. k. Ringe R mit s. l. k. Radikal eine ringtheoretische direkte Zerlegung gibt:

(1) $$R = \sum_{\mu \in \Gamma} R_\mu \boxplus R^*,$$

wobei für jedes $\mu \in \Gamma R_\mu$ der volle Endomorphismenring eines Vektorraumes über unendlichem Schiefkörper und der Restklassenring von R^* nach dem Radikal direkte Summe von Endomorphismenringen von Vektorräumen über endlichen Körpern sind.

Sei neben (1) noch die Darstellung

$$R = \sum_{\mu' \in \Gamma'} R'_{\mu'} \boxplus R^{*\prime}$$

mit der obigen Eigenschaft gegeben. Da

$$R/I \cong \sum_{\mu \in \Gamma} R_\mu \boxplus (R^*/I) \cong \sum_{\mu' \in \Gamma'} R'_{\mu'} \boxplus (R^{*\prime}/I)$$

gilt, gibt es wegen der Eindeutigkeit von halbeinfachen l. k. Ringen eine eineindeutige Abbildung $\mu \to \mu'$ zwischen Γ und Γ', so daß $R_\mu \cong R'_{\mu'}$ für jedes $\mu \in \Gamma$ ist. Also gilt

$$R^* \cong R / \sum_{\mu \in \Gamma} R_\mu \cong R / \sum_{\mu' \in \Gamma'} R'_{\mu'} \cong R^{*\prime}.$$

Wir fassen die Ergebnisse dieses Paragraphen zusammen:

Satz. (I) *Jeder l. k. Ring mit s. l. k. Radikal besitzt eine ringtheoretische direkte Darstellung der Form*

$$R = \sum_{\mu \in \Gamma} R_\mu \boxplus R^*,$$

wobei für jedes $\mu \in \Gamma R_\mu$ der volle Endomorphismenring eines Vektorraumes übereinem unendlichen Schiefkörper und der Restklassenring von R^* nach dem Radikal direkte Summe voller Endomorphismenringe von Vektorräumen über endlichen Körpern sind.

(II) *Die obige Darstellung ist eindeutig bis auf Isomorphie.*

Als Folgerungen dieses Satzes betrachten wir folgende zwei Fälle:

1) Es sei R ein linksartinscher Ring mit linksartinschem Radikal. Da jeder linksartinscher Radikalring eine Gruppe mit Minimalbedingung für Untergruppen ist, erhalten wir in diesem Fall nach Satz 1 die Ergebnisse von Kertész und Widiger [3].

2) Es sei R ein l. k. Ring mit einer Umgebungsbasis des Nullelements aus Idealen und mit i. e. S. l. k. Radikal. Da jeder i. e. S. l. k. Radikalring mit einer Umgebungsbasis des Nullelements aus Idealen s. l. k. ist, erhalten wir in diesem Falle die Ergebnisse von Widiger [11].

Es sei nun R ein topologischer Ring mit Radikal J. Wir sagen, daß R die wedderburnsche Zerlegung besitzt, wenn R in der Form $R=S\oplus I$ als Gruppe mit einem passenden abgeschlossenen Unterring S darstellbar ist. Der Unterring S heißt dann Wedderburnscher Faktor von R.

Für kompakte Ringe bewies NUMAKARA [7] den folgenden Satz:

Es sei R ein kompakter Ring und $\bar{e}$ das Einselement des Restklassenringes R/I von R nach dem Radikal I. R hat genau dann die wedderburnsche Zerlegung, wenn $\bar{e}$ einen (nicht notwendig idempotenten) Vertreter e mit folgender Eigenschaft besitzt:

Für jede Umgebung U des Nullelementes in R gibt es endlich viele verschiedene Primzahlen $p_1, \dots, p_n$ mit $(p_1 \dots p_n)e \in U$.

Es sei jetzt R ein l. k. Ring mit s. l. k. Radikal I. Dann ist

$$\bar{R} = R/I = \sum_{\mu\in\Gamma} \bar{R}_\mu = \sum_{\mu\in\Gamma} \bar{e}_\mu \bar{R} \bar{e}_\mu$$

direkte Summe einfacher l. k. Ringe R_μ mit dem Einselement $\bar{e}_\mu$. Es ist bekannt, daß es in R ein System orthogonaler idempotenter Vertreter $\{e_\mu\}$ von $\{\bar{e}_\mu\}$ und e von $\bar{e}=\sum \bar{e}_\mu$ gibt, so daß

$$e = \sum e_\mu, \quad ee_\mu = e_\mu e = e_\mu, \quad e_\mu e_\nu = 0, \quad \mu \neq \nu$$

für jedes μ aus Γ gilt. Wegen Satz 5 [1] ist $e_\mu I=0$ für jeden unendlichen Ring $\bar{R}_\mu$. Dies zeigt, daß $\bar{e}_\mu \bar{R} \bar{e}_\mu$ einfacher l. k. Ring für jeden unendlichen Ring $\bar{R}_\mu$ ist. Es sei $\bar{R}^*=\sum \bar{R}_\mu$ die direkte Summe aller endlicher Ringe $\bar{R}_\mu$. Dann ist $\bar{R}^*$ ein kompakter Ring. Es sei R^* das vollständige Urbild von $\bar{R}^*$ beim natürlichen Homomorphismus von R auf $\bar{R}$. Offensichtlich hat R genau dann die wedderburnsche Zerlegung, wenn R^* diese hat. Da I s. l. k. ist, ist auch R^* s. l. k. Es sei jetzt $\bar{e}^*$ das Einselement von R^*. Dann hat $\bar{e}^*$ einen idempotenten Vertreter e^* in R^*. So gilt $R^*=e^*R^*e^*\oplus e^*R^*(1-e^*)+(1-e^*)R^*e^*\oplus(1-e^*)R^*(1-e^*)$, wobei $e^*R^*(1-e^*)= =\{e^*a-e^*ae^*| \quad a\in R^*\}$, $(1-e^*)R^*e^*=(ae^*-e^*ae^*|a\in R^*\}$, $(1-e^*)R^*(1-e^*)= =\{a-e^*a-ae^*+e^*ae^*|a\in R^*\}$ ist. Offensichtlich liegt $e^*R^*(1-e^*)\oplus(1-e^*)\ R^*e^*\oplus \oplus(1-e^*)R^*(1-e^*)$ im Radikal I. Andererseits ist klar, daß $e^*R^*e^*$ ein s. l. k. Ring mit Einselement ist. Für ein beliebiges, offenes Linksideal L in $K=e^*R^*e^*$ ist K/L eine Gruppe mit Minimalbedingung für Untergruppen. Deshalb gilt $K/L= =D\oplus E$, wobei D bzw. E eine teilbare bzw. endliche Untergruppe von K/L ist. Besteht D nicht nur aus dem Nullelement, so ist D direkte Summe endlich vieler quasi-zyklischer Gruppen. Dies ist ein Widerspruch, da K/L ein zyklischer K-Modul ist. Also ist K/L für jedes offene Linksideal L in K endlich, d. h. K ist ein kompakter Ring. Daraus folgt, daß R genau dann die wedderburnsche Zerlegung besitzt, wenn K diese hat. Also erhalten wir aus dem obigen Satz von Numakura

SATZ 2. *Es sei R ein l. k. Ring mit s. l. k. Radikal I. Es sei $\bar{e}^*$ das Einselement der direkten Summe aller endlichen einfachen direkten Summanden von R/I. R hat genau dann die Wedderburnsche Zerlegung, wenn $\bar{e}^*$ einen (nicht notwendig idempotenten) Vertreter e^* mit folgender Eigenschaft hat:*

Für jede Umgebung U des Nullelementes in R gibt es endlich viele verschiedene Primzahlen $p_1, \dots, p_n$ mit $(p_1 \dots p_n)e^* \in U$.

§ 3. Linear kompakte Ringe mit MHL-Radikal

Wir wollen nun die Struktur l. k. Ringe, deren Radikal als Ring ein MHL-Ring ist, untersuchen. Nach [8] ist jeder MHL-Radikalring ein Torsionsring. Es sei jetzt R ein l. k. Ring mit MHL-Radikal I. Bekanntlich [8] ist jeder MHL-Radikalring transfinit r-nilpotent ist. Deshalb ist R in der gröbsten linear kompakten Topologie i. e. S. l. k. Daher können wir voraussetzen, daß R ein i. e. S. l. k. Ring ist. Nach (XVII) [1] ist

$$\bar{R} = R/I = \sum_{\mu\in\Gamma^*} \bar{e}_\mu \bar{R} \bar{e}_\mu = \sum_{\mu\in\Gamma^*} \bar{R}_\mu$$

direkte Summe einfacher l. k. Ringe $\bar{R}_\mu$ mit Einselement $\bar{e}_\mu \bar{e} = \sum_{\mu\in\Gamma^*} \bar{e}_\mu$ ist das Einselement von $\bar{R}$. Nach (XX) [1] gibt es ein System orthogonaler idempotenter Vertreter $\{e_\mu, \mu\in\Gamma^*\}$ von $\{\bar{e}_\mu, \mu\in\Gamma^*\}$ und e von e^* so daß

$$e = \sum_{\mu\in\Gamma^*} e_\mu, \quad ee_\mu = e_\mu e = e_\mu, \quad e_\mu e_\nu = 0, \quad \mu \neq \nu, \quad \mu, \nu\in\Gamma^*$$

gilt. Weger (XVI) [1] ist dann jedes R_μ ein voller Endomorphismenring eines Vektorraumes über einem Schiefkörper S_μ. Es sei Γ die Menge aller μ aus Γ^*, für die die S_μ Schiefkörper mit der Charakteristik 0 sind.

HILFSSATZ 3.1. *Für alle* $\mu\in\Gamma$ *gilt* $e_\mu I e_\mu = e_\mu I = I e_\mu = 0$.

BEWEIS. Wegen (XIX) [1] ist $R_\mu = e_\mu R e_\mu$ ein primärer i. e. S. l. k. Ring mit dem Radikal $I_\mu = e_\mu I e_\mu$ für jeden beliebigen Index μ in Γ. Erstens nehmen wir an, daß I_μ entgegen unserer Behauptung nicht nur aus dem Nullelement besteht. Wenn $I_\mu^2 \neq (0)$ ist, so betrachten wir den Restklassenring $R_\mu|\mathrm{cl}\,(I_\mu^2)$ nach dem Abschluß $\mathrm{cl}\,(I_\mu^2)$ von I_μ^2. Es gilt natürlich $I_\mu \neq \mathrm{cl}\,(I_\mu^2)$, da sonst $(I_\mu)_\xi = I_\mu$ für alle Ordnungszahlen ξ im Widerspruch steht zur Voraussetzung, daß I_μ transfinit nilpotent und $I_\mu \neq (0)$ sind. Da I ein MHL-Radikalring ist, so ist I ein Torsionsring. Deshalb sind I_μ und so auch $I_\mu/\mathrm{cl}\,(I_\mu^2)$ Torsionsringe. Da $(I_\mu/\mathrm{cl}\,(I_\mu^2))^2 = 0$ ist, genügt es, sich auf den Fall mit $I_\mu^2 = 0$ zu beschränken. Dann wählen wir ein beliebiges, aber festes Element $a\in I_\mu$. Wir bilden den $\bar{R}_\mu$-Modul $\bar{R}_\mu = R_\mu/I_\mu$ vermöge φ in den $\bar{R}_\mu$-Modul I_μ ab:

$$\varphi\colon \bar{R}_\mu \to I_\mu, \quad \varphi(x) = xa, \quad \bar{x}\in\bar{R}_\mu, \quad x\in R_\mu$$

wobei x ein Vertreter der Restklasse $\bar{x}$ in R_μ ist. φ ist unabhängig von der Wahl des Repräsententen aus der Restklasse $\bar{x}$, denn für jedes $b\in I_\mu$ gilt $ba=0$ wegen $I_\mu^2=0$. φ ist evident ein $\bar{R}_\mu$-Homomorphismus. φ ist stetig, weil die Multiplikation in R_μ eine stetige Funktion ist. Der Kern $\ker(\varphi)$ des Homomorphismus φ ist deshalb ein abgeschlossenes Linksideal von $\bar{R}_\mu$. Also ist $\ker(\varphi)$ ein direkter Summand von $\bar{R}_\mu$. Dies zeigt, daß das Bild $\mathrm{im}\,(\varphi)$ von $\bar{R}_\mu$ Torsionsfrei ist. Aber das Radikal I_μ ist ein Torsionsring, so ist φ trivial, d. h. $\varphi(\bar{R}_\mu)=0$. Dies würde bedeuten, daß $xa=0$ für alle x aus R_μ wäre; R_μ hat aber nach seiner Definition das Einselement e_μ. Dieser Widerspruch zeigt, daß $e_\mu I e_\mu = 0$ gilt. Also ist jetzt R_μ ein einfacher l. k. Ring. Nach (XVI) [1] gibt es in R_μ ein vollständiges System $\{e_{\mu\delta}, \delta\in\Delta_\mu\}$ minimaler orthogonaler Idempotente. Für jedes beliebige, aber feste Element $a\in I$ bildet die Abbildung

$$\varphi_l\colon e_{\mu\delta} R e_{\mu\delta} \to I, \quad \varphi(x) = xa, \quad x\in e_{\mu\delta} R_\mu e_{\mu\delta}$$

bzw. die Abbildung

$$\varphi_r: e_{\mu\delta} R_\mu e_{\mu\delta} \to I, \quad \varphi_r(x) = ax, \quad x \in e_{\mu\delta} R_\mu e_{\mu\delta}$$

den mit S_μ isomorphen Schiefkörper $R_{\mu\delta} = e_{\mu\delta} R_\mu e_{\mu\delta}$ in den $R_{\mu\delta}$-Linksbzw. Rechtsmodul I ab. Es zeigt sich unmittelbar, daß φ_l bzw. φ_r ein stetiger $R_{\mu\delta}$-Links-bzw. Rechtshomomorphismus ist. Da $R_{\mu\delta}$ ein Schiefkörper mit der Charakteristik 0 und I ein Torsionsring sind, sind φ_l und φ_r trivial. Also gilt $Ie_{\mu\delta} = e_{\mu\delta} I = 0$. Aus $e_\mu = \sum_{\delta \in \Delta\mu} e_{\mu\delta}$ folgt nach (XXVI) [1] $Ie_\mu = e_\mu I = 0$.

HILFSSATZ 3.2. *$e_\mu R e_\mu$ ist für alle $\mu \in \Gamma$ Ideal von R und deshalb abgeschlossen in R.*

Der Beweis verläuft völlig analog zu dem Beweis von Hilfssatz 2.2.

Nach der obigen Vorbereitung kann man den folgenden Satz ohne Mühe mit Methoden des Beweises vom Satz 1 einsehen:

SATZ 3. *Es sei R ein l. k. Ring mit MHL-Radikal. Dann besitzt R die folgende direkte Zerlegung:*

$$R = \sum R_\mu \boxplus R^*,$$

wobei die R_μ volle Endomorphismenringe der Vektorräume über Schiefkörpern mit der Charakteristik 0 sind und der Restklassenring von R^* nach dem Radikal direkte Summe voller Endomorphismenringe von Vektorräumen über Schiefkörpern mit der Charakteristik p ist, wobei p Primzahl ist.

Diese Zerlegung ist eindeutig bis auf topologische Isomorphie.

Ein Ring R heißt spaltbar, wenn das maximale Torsionsideal T von R ein ringtheoretischer direkter Summand von R ist. Aus Satz 3 folgt unmittelbar

SATZ 4. *Jeder l. k. Ring mit MHL-Radikal ist spaltbar.*

Wir bemerken hier, daß die Frage, ob jeder i. e. S. l. k. Ring spaltbar ist, noch ungelöst bleibt.

Literaturverzeichnis

[1] P. N. ANH, Über die Struktur linear kompakter Ringe, *Acta Math. Acad. Sci. Hungar.* (im Druck).

[2] P. N. ANH, Die additive Struktur der im engeren Sinne linear kompakten Ringe, *Studia Sci. Math. Hung.* (im Druck).

[3] A. KERTÉSZ und A. WIDIGER, Artinsche Ringe mit artinschem Radikal, *J. reine angew. Math.*, **242** (1970), 9—15.

[4] H. LEPTIN, Über eine Klasse linear kompakter abelscher Gruppen, *Abh. Math. Sem. Hamb.*, **19** (1954), 23—40.

[5] H. LEPTIN, Linear kompakte Moduln und Ringe, *Math. Z.*, **62** (1955), 241—267.

[6] H. LEPTIN, Linear kompakte Moduln und Ringe, *Math. Z.*, **66** (1957), 289—327.

[7] K. NUMAKURA, A note on Wedderburn decomposition of compact rings, *Proc. Japan Acad.*, **XXXV** (1959), 313—315.

[8] F. SZÁSZ, Über Ringe mit Minimalbedingung für Hauptrechtsideale. I, *Publ. Math. Debrecen*, **7** (1960), 54—64.

[9] F. SZÁSZ, Über Ringe mit Minimalbedingung für Hauptrechtsideale. II, *Acta Math. Acad. Sci. Hungar.*, **12** (1961), 417—439.

[10] T. SZELE, Nilpotent artinian rings, *Publ. Math. Debrecen*, **4** (1955), 71—78.

[11] A. WIDIGER, Die Struktur einer Klasse linear kompakter Ringe, *Beitäge zur Algb. und Geom.*, **3** (1974), 139—159.

(Received October 17, 1977)

1118 BUDAPEST
BUDAÖRSI ÚT 95—101.

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (3—4), (1979), 253—255.

THE EXTENSION PROPERTY OF ORDERINGS

By
I. CHAJDA (Přerov)

A. DAY states in [1] that the Congruence Extension Property is equivalent to the Principal Congruence Extension Property on every class of algebras closed under the formation of subalgebras. A new proof of Day's theorem that avoids the use of the Axiom of Choice is given in [3]. It is a natural question under which conditions a similar statement can be true also for orderings. Since orderings on a given algebra need not be closed under the formation of upper bounds, the investigated problem is relativized for the case of bounded orderings only.

Let $\mathfrak{A}=(A, F)$ be an algebra and R a binary relation on A, i.e. $R\subseteq A\times A$. Call R *compatible with* $\mathfrak{A}$ provided the implication

$$\langle a_i, b_i\rangle\in R \quad (i = 1, \dots, n) \Rightarrow \langle f(a_1, \dots, a_n), f(b_1, \dots, b_n)\rangle\in R$$

is valid for each n-ary $f\in F$.

By an *ordering on* $\mathfrak{A}$ a partial ordering (i.e. a reflexive, transitive and antisymmetric binary relation) which is compatible with $\mathfrak{A}$ is meant. Denote by I_A the identity relation on A, i.e. $\langle a, b\rangle\in I_A$ if and only if $a=b$. Clearly, I_A is an ordering on $\mathfrak{A}$. If $\mathfrak{B}=(B, F)$ is a subalgebra of $\mathfrak{A}$, the *restriction* of $R\subseteq A\times A$ onto B is denoted by $R|B$ (i.e. $R|B=R\cap(B\times B)$).

The following lemma is clear:

LEMMA 1. *Let* $\mathfrak{B}=(B, F)$ *be a subalgebra of* $\mathfrak{A}$ *and* P *an ordering on* $\mathfrak{A}$. *Then* $P|B$ *is an ordering on* $\mathfrak{B}$.

Let R be a relation on a set A. Recall the known concept of a *transitive hull* R^T of R: $\langle a, b\rangle\in R^T$ if and only if there exist elements $a_0, \dots, a_n\in A$ with $a_0=a$, $a_n=b$ and $\langle a_{i-1}, a_i\rangle\in R$ for each $i=1, \dots, n$. The transitive hull R^T is (in the sense of set inclusion) the least binary relation on A containing R which is transitive. Provided R is reflexive and compatible with $\mathfrak{A}$, also R^T is compatible with $\mathfrak{A}$ (see [4]).

It is well-known the following

LEMMA 2. *Let* $\mathfrak{A}=(A, F)$ *be an algebra and* P *an ordering on* $\mathfrak{A}$. *The set* $\mathscr{P}(\mathfrak{A}, P)$ *of all orderings* Q *on* $\mathfrak{A}$ *with* $Q\subseteq P$ *forms a complete lattice with respect to the set inclusion. The least element in* $\mathscr{P}(\mathfrak{A}, P)$ *is* I_A, *the greatest one is* P; *the lattice meet coincides with the set intersection and the lattice join* $\vee\{R_\gamma;\ \gamma\in\Gamma\}$ *is the transitive hult of the union* $\cup\{R_\gamma;\ \gamma\in\Gamma\}$.

DEFINITION 1. Let P be an ordering on an algebra $\mathfrak{A}=(A, F)$ and $a, b\in A$ with $\langle a, b\rangle\in P$. By $P_{\mathfrak{A}}(a, b)$ we denote the set intersection of all orderings Q on

$\mathfrak{A}$ with $\langle a, b\rangle \in Q$ if such Q exists, and is equal to I_A otherwise. It is called a *principal ordering on* $\mathfrak{A}$ *(generated by* $\langle a, b\rangle$*)*.

REMARK. Like in the Lemma 2, it can be easily shown that $P_{\mathfrak{A}}(a, b)$ is also an ordering on $\mathfrak{A}$. Clearly $P_{\mathfrak{A}}(a, b) \subseteq P$, thus $P_{\mathfrak{A}}(a, b) \in \mathscr{P}(\mathfrak{A}, P)$. Hence, $P_{\mathfrak{A}}(a, b)$ is the meet of all orderings in $\mathscr{P}(\mathfrak{A}, P)$ containing $\langle a, b\rangle$.

LEMMA 3. *Let P be an ordering on an algebra $\mathfrak{A}=(A, F)$. Then $Q = \bigvee \{P_{\mathfrak{A}}(a, b); \langle a, b\rangle \in Q\}$ in $\mathscr{P}(\mathfrak{A}, P)$ for each $Q \in \mathscr{P}(\mathfrak{A}, P)$.*

The proof is easy and is left to the reader.

The following statement is a generalization of GOLDIE's lemma formulated in [2] for congruences:

LEMMA 4. *Let $\mathfrak{B}=(B, F)$ be a subalgebra of $\mathfrak{A}=(A, F)$, R a relation compatible with $\mathfrak{A}$ such that $I_B \subseteq R$ and S a transitive relation compatible with $\mathfrak{B}$ such that $R|B \subseteq S$. Put $D=[B]_R=\{x; x\in A$ and $\langle x, y\rangle \in R$ for some $y \in B\}$. Introduce a binary relation $S(R)$ on D by the rule*

$$(*) \quad \langle u, v\rangle \in S(R) \quad \text{iff} \quad \langle u, x\rangle \in R,\ \langle x, y\rangle \in S,\ \langle y, v\rangle \in R \quad \text{for some} \quad x, y \in B.$$

Then

(a) *$\mathscr{D}=(D, F)$ is a subalgebra of $\mathfrak{A}$,*

(b) *$S(R)$ is compatible with $\mathscr{D}$ and $S(R)|B=S$,*

(c) *if R is transitive, then $S(R)$ is also transitive,*

(d) *if P is an ordering on $\mathfrak{A}$ and $R, S \subseteq P$ are reflexive and transitive, then $S(R)$ is an ordering on $\mathscr{D}$.*

PROOF. (a) Let $f\in F$ be n-ary and $a_1, \dots, a_n \in D$. Hence $\langle a_i, b_i\rangle \in R$ for some $b_1, \dots, b_n \in B$ and, since R is compatible with $\mathfrak{A}$, also $\langle f(a_1, \dots, a_n), f(b_1, \dots, b_n)\rangle \in R$. As $\mathfrak{B}$ is a subalgebra of $\mathfrak{A}$, we have $f(a_1, \dots, a_n) \in D$ proving $\mathscr{D}=(D, F)$ is a subalgebra of $\mathfrak{A}$.

(b) Since $I_B \subseteq R|B \subseteq S$, the transitivity of S implies

$$S = I_B \cdot S \cdot I_B \subseteq R|B \cdot S \cdot R|B \subseteq S \cdot S \cdot S = S.$$

Hence $S = R|B \cdot S \cdot R|B$. By $(*)$, $S(R)|B = R|B \cdot S \cdot R|B$, thus $S(R)|B = S$. The compatibility of $S(R)$ with $\mathscr{D}$ follows directly from this property of R, S.

(c) Suppose the transitivity of R. Let $\langle a, b\rangle \in S(R)$ and $\langle b, c\rangle \in S(R)$. By $(*)$, there exist $x, y, z, w \in B$ with $\langle a, x\rangle \in R$, $\langle x, y\rangle \in S$, $\langle y, b\rangle \in R$ and $\langle b, z\rangle \in R$, $\langle z, w\rangle \in S$, $\langle w, c\rangle \in R$. Since R is transitive, we have $\langle y, z\rangle \in R$. However, $y, z \in B$ and $R|B \subseteq S$, thus $\langle y, z\rangle \in S$. The transitivity of S implies $\langle x, w\rangle \in S$, and by $(*)$, we have $\langle a, c\rangle \in S(R)$ proving statement (c).

(d) By (b), $S(R)$ is compatible with $\mathscr{D}$, by (c) it is transitive. The reflexivity of $S(R)$ is clear. As $R, S \subseteq P$, also $S(R) \subseteq P$, thus $S(R)$ is also antisymmetric.

DEFINITION 2. Let P be an ordering on an algebra $\mathfrak{A}$. We say that $\mathfrak{A}$ satisfies the *P- (Principal) Ordering Extension Property* if for each subalgebra $\mathfrak{B}=(B, F)$ of $\mathfrak{A}$ and every (principal) ordering Q on $\mathfrak{B}$ with $Q \subseteq P$ there exists an ordering $\bar{Q}$ on $\mathfrak{A}$ such that $\bar{Q}|B = Q$ (the so called *extension of* Q onto the whole $\mathfrak{A}$).

LEMMA 5. *Let P be an ordering on an algebra $\mathfrak{A}$. Then $\mathfrak{A}$ satisfies the P-Principal Ordering Extension Property if and only if for each subalgebra $\mathfrak{B}=(B, F)$ of $\mathfrak{A}$ and every $a, b \in B$, $\langle a, b\rangle \in P$, we have $P_{\mathfrak{B}}(a, b) = P_{\mathfrak{A}}(a, b)|B$.*

PROOF. Since $\langle a, b\rangle \in P$, $P_{\mathfrak{A}}(a, b)$ is correctly introduced and, by Lemma 1, $P_{\mathfrak{A}}(a, b)|B$ is an ordering on $\mathfrak{B}$ containing the pair $\langle a, b\rangle$. Hence $P_{\mathfrak{B}}(a, b) \subseteqq P_{\mathfrak{A}}(a, b)|B$. As $\mathfrak{A}$ satisfies the P-Principal Ordering Extension Property, there exists an ordering $\bar{Q}$ on $\mathfrak{A}$ with $\bar{Q}|B = P_{\mathfrak{B}}(a, b)$. As $P_{\mathfrak{A}}(a, b)$ is the least ordering on $\mathfrak{A}$ containing $\langle a, b\rangle$, we conclude $P_{\mathfrak{A}}(a, b) \subseteqq \bar{Q}$, thus $P_{\mathfrak{A}}(a, b)|B \subseteqq \bar{Q}|B = P_{\mathfrak{B}}(a, b)$ proving the converse inclusion.

THEOREM. *Let P be an ordering on an algebra $\mathfrak{A}$. The following conditions are equivalent:*

(a) *$\mathfrak{A}$ satisfies the P-Ordering Extension Property,*

(b) *$\mathfrak{A}$ satisfies the P-Principal Ordering Extension Property.*

PROOF. Clearly (a)⇒(b). Prove the converse implication. If $\mathscr{C} = (C, F)$ is a subalgebra of $\mathfrak{A} = (A, F)$ and Q an ordering on $\mathscr{C}$ with $Q \subseteqq P$, then by Lemma 2 there exists the least ordering $\bar{Q}$ on $\mathfrak{A}$ satisfying $\bar{Q}|_C \supseteqq Q$. By Lemma 3, $\bar{Q} = \bigvee \{P_{\mathfrak{A}}(x, y); \langle x, y\rangle \in \bar{Q}\}$. However, also $Q^* = \bigvee \{P_{\mathfrak{A}}(x, y); \langle x, y\rangle \in Q\}$ satisfies clearly $Q^*|_C \supseteqq Q$ and $Q^* \subseteqq \bar{Q}$, thus, by the minimality of $\bar{Q}$, we have

$$\bar{Q} = \bigvee \{P_{\mathfrak{A}}(x, y); \langle x, y\rangle \in Q\}.$$

We want to show $\bar{Q}|_C = Q$. In the wiew of foregoing, (b)⇒(a) is equivalent to

$(**)$ For every subalgebra $\mathscr{C} = (C, F)$ of $\mathfrak{A}$ and an arbitrary ordering Q on $\mathscr{C}$ with $Q \subseteqq P$, if $a, b \in C$, $\langle a_i, b_i\rangle \in Q$ for $a_i, b_i \in C$ $(i = 1, \dots, n)$ and $x_0, \dots, x_n \in A$ such that $x_0 = a$, $x_n = b$ and $\langle x_{i-1}, x_i\rangle \in P_{\mathfrak{A}}(a_i, b_i)$ for $i = 1, \dots, n$, then $\langle a, b\rangle \in Q$.

We prove $(**)$ by induction on n. For $n = 1$ it is obvious by Lemma 5 since $\mathfrak{A}$ satisfies the P-Principal Ordering Extension Property. Assume $n > 1$ and $(**)$ is valid for $n-1$. Set $D = [C]_{P_{\mathfrak{A}}(a_n, b_n)}$, $Q_0 = \bigvee \{P_{\mathscr{C}}(a_i, b_i); i = 1, \dots, n\}$ in $\mathscr{P}(\mathscr{C}, P|_C)$, $T = Q_0(P_{\mathfrak{A}}(a_n, b_n))$ (the symbols like in Lemma 4). Since $\mathfrak{A}$ satisfies P-Principal Ordering Extension Property, we obtain $P_{\mathfrak{A}}(a_n, b_n)|_C \subseteqq Q_0$. Thus, by Lemma 4 (a), $\mathscr{D} = (D, F)$ is a subalgebra of $\mathfrak{A}$ and, by Lemma 4 (d), T is an ordering on $\mathscr{D}$. Further, by Lemma 4 (b), $T|_C = Q_0 \subseteqq Q$. Now we observe that $\mathscr{D}$, T and $a = x_0$, $x_1, \dots, x_{n-1}$, and $a_1, b_1, \dots, a_{n-1}, b_{n-1}$ satisfy assumption $(**)$ for $n-1$, hence $\langle a, x_{n-1}\rangle \in T$. By Lemma 4 (c), T is transitive, i.e. $\langle a, b\rangle \in T$. As $a, b \in C$ and $T|_C = Q_0 \subseteqq Q$, we conclude $\langle a, b\rangle \in Q$ proving $(**)$ for all integers. The proof is complete.

References

[1] A. DAY, A note on the congruence extension property, *Algebra Univ.*, **1** (1971), 234—235.

[2] A. W. GOLDIE, The Jordan—Hölder theorem for general abstract algebras, *Proc. London Math. Soc.*, **52** (1950), 107—131.

[3] G. GRÄTZER, H. LAKSER, *Two observations on the Congruence Extension Property* (preprint).

[4] I. CHAJDA, B. ZELINKA, Compatible relations on algebras, *Časop. pěst. matem.*, **100** (1975), 355—360.

(Received October 27, 1977)

TŘÍDA LIDOVÝCH MILICÍ 22
750 00 PŘEROV
CZECHOSLOVAKIA

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (3—4), (1979), 257—260.

ON THE COMMUTATIVITY OF CERTAIN RINGS

By

B. FELZENSZWALB (Rio de Janeiro)[1]

In [3] HERSTEIN have proved that if R is a ring satisfying the identity $(xy)^n = x^n y^n$ for a fixed integer $n>1$, then the commutator ideal of R is nil. Recently [6, Theorem 2], A. KAYA have proved that a ring R with unit element such that for every $x, y \in R$

(1) $$(xy)^k = x^k y^k, \quad k = n(x, y),\ n(x, y)+1,\ n(x, y)+2$$

where $n(x, y)$ is an integer >1 which depends on x and y, is commutative if it is a primary ring or if it is semiprime. Moreover [6, Theorem 1] if R is a semisimple ring (not necessarily with 1) such that for every $x, y \in R$ the identities (1) hold for just two consecutive integers, then R must be commutative. In this paper we shall prove the following:

THEOREM 1. *Let R be a ring with unit element in which for every $x, y \in R$ there exists an integer $n(x, y)>1$ such that*

$$(xy)^k = x^k y^k, \quad k = n(x, y),\ n(x, y)+1.$$

Then the commutator ideal of R is nil. Equivalently, if R has no nonzero nil ideals, then R must be commutative.

Localizing the problem we shall also prove:

THEOREM 2. *Let R be a ring with no nonzero nil right ideals. Suppose that $a \in R$ is such that for every $x \in R$*

(2) $$(ax)^k = a^k x^k, \quad k = n(x),\ n(x)+1,\ n(x)+2$$

where $n(x)$ is an integer >1 which depends on x. Then a is a central element.

In the proof of Theorem 2 we also show that if R is a semisimple ring and $a \in R$ is such that for every $x \in R$ the identities (2) hold for just two consecutive integers, the conclusion that a is central remains valid.

PROOF OF THEOREM 1. Let $a \in R$ be non-nilpotent and let U_a be an ideal of R maximal with respect to the exclusion of all powers of a. Then $R_a = R/U_a$ is a ring with the property that every nonzero ideal in R_a contains some power of $\bar{a} = a + U_a$ (so R_a has no nonzero nil ideals). Moreover R_a inherits the hypothesis placed on R. Since, as a runs over the non-nilpotent elements of R, $\cap U_a$ is a nil

[1] This work was supported by FINEP, BRAZIL.

ideal, the result will follow if we show that each U contains the commutator ideal of R (i.e., that each R_a is commutative).

Thus, we may assume, henceforth, that R is a ring with no nonzero nil ideals and containing a non-nilpotent element a such that every nonzero ideal of R contains some power of a. We have to show R is commutative.

Let $b, d\in R$ with $bd=0$. Let $n>1$ such that

$$[d(1+db)]^k = d^k(1+db)^k, \quad k = n, n+1. \tag{3}$$

Since $bd=0$, on expansion of (3) we get $(n-1)d^{n+1}b=0=n(d^{n+2}b)$. Hence $(n-1)d^{n+2}b=0=nd^{n+2}b$ and so $d^{n+2}b=0$. Similarly one shows that $db^s=0$ for some s. Thus for every $x\in R$

$$V_x = \{r\in R | rx^m = 0 \text{ for some } m\}$$

is an ideal of R.

Since every nonzero ideal of R contains some power of a and a is not nilpotent, $V_a=0$. Hence a is regular. If $b\in R$ is a zero-divisor, then $V_b\neq 0$; so $a^t\in V_b$ for some t. Since a is regular, b must be nilpotent. That is, every zero-divisor in R is nilpotent.

We claim that for every $x, y\in R$ there exists $m=m(x, y)\geqq 1$ such that $xy^m=y^m x$.

Case I. If x and y are regular elements, let $m>1$ such that $(xy)^m=x^m y^m$ and $(xy)^{m+1}=x^{m+1}y^{m+1}$. Then $x^m y^m xy=x^{m+1}y^{m+1}$; so $y^m x=xy^m$.

Case II. If y is a zero-divisor, then, by what we deduced before, y is nilpotent. Hence $xy^m=0=y^m x$ for some m.

Case III. If x is a zero-divisor, then x is nilpotent. So $(1+x)$ is regular and, by the first two cases, $(1+x)y^m=y^m(1+x)$ for some m. Hence $xy^m=y^m x$. The claim now follows. Since R has no nonzero nil ideals, by a result of Herstein [4] we conclude that R is commutative. This establishes the theorem.

We proced now to prove Theorem 2.

Notation. Let R be a ring. If $x, y\in R$ let $[x, y]=xy-yx$. If $c\in R$ let $r(c)=\{x\in R | cx=0\}$ and $l(c)=\{x\in R | xc=0\}$, the right and left annihilator of c in R respectively.

Lemma 1. *Let R be a ring with no nonzero nil right ideals. Suppose that $a\in R$ is such that for every $x\in R$ there exists an integer $n=n(x)>1$ with $(ax)^n=a^n x^n$. Then*

(i) *if $x, y\in R$ and $xy=0$, then $xay=0$,*
(ii) *$r(a)=\{x\in R | a^m x=0$ for some $m\geqq 1\}$,*
(iii) *$l(a)=\{x\in R | xa^m=0$ for some $m\geqq 1\}$,*
(iv) *$l(a)=r(a)$.*

Proof. (i) Let $x, y\in R$ with $xy=0$. Then $(ayx)^n=a^n(yx)^n=0$ for a suitable $n>1$. Thus $xar(x)$ is a nil right ideal of R. By our assumptions we must have $xar(x)=0$. In particular, $xay=0$.

(ii) It is enough to show that if $b\in R$ and $a^2b=0$, then $ab=0$. But $a^2b=0$ implies that for every $x\in R$, $(abx)^n=a^n(bx)^n=0$ for a suitable $n=n(bx)>1$. That is abR is nil. Since R has no nonzero nil right ideals we get $ab=0$.

(iii) Let $b\in R$ with $ba^2=0$. Then for every $x\in R$, $b(axb)^n=ba^n(xb)^n=0$ for a suitable $n=n(xb)>1$. Thus baR is nil. Since R has no nonzero nil right ideals we get $ba=0$ and (iii) follows.

(iv) If $x\in l(a)$ there exists $n=n(x)>1$ such that $a^nx^n=(ax)^n=0$, so, by (ii), $ax^n=0$. Since R has no nonzero nil right ideals, by [2, Theorem 2], $al(a)=0$. That is $l(a)\subset r(a)$. It remains now to show $r(a)\subset l(a)$.

Let $b\in r(a)$. Then $(ba)^2=0$ and $c=(1+ba)a(1-ba)=a+ba^2$ has the same property as a: $(cx)^{m(x)}=c^{m(x)}x^{m(x)}$, $m(x)>1$, all $x\in R$. Thus, by (i), if $x, y\in R$ and $xy=0$ we obtain $0=xcy=xay+xba^2y=xba^2y$.

Now, if $x\in R$ there exists $n=n(x)>1$ such that $\big(a^nx^{n-1}-(ax)^{n-1}a\big)x=0$. Hence $a^nx^{n-1}ba^2x-(ax)^{n-1}aba^2x=(a^nx^{n-1}-(ax)^{n-1}a)ba^2x=0$. Since $b\in r(a)$ we obtain $a^nx^{n-1}ba^2x=0$; so $(a^nx^{n-1}b)^2=0$. Thus $a^nx^{n-1}r(a)$ is a nil right ideal of R. By our assumptions on R we must have $a^nx^{n-1}r(a)=0$, and consequently, by (ii), $ax^{n-1}r(a)=0$. By [1, Lemma 1] it follows that $r(a)a=0$, i.e., $r(a)\subset l(a)$.

LEMMA 2. *Let R be a semisimple ring. Suppose that $a\in R$ is such that for every $x\in R$ there exists an integer $n(x)>1$ with*

$$(ax)^k = a^kx^k, \quad k = n(x),\ n(x)+1.$$

Then a is a central element.

PROOF. Since R is semisimple it is a subdirect sum of primitive rings R_α. The image of a in each R_α has the same property in R_α as a in R (for each R_α is a homomorphic image of R). If we knew that the image of a in each R_α is central we would get a in the centre of R. Hence without loss of generality we may assume that R is primitive. Therefore R is a dense ring of linear transformations on a vector space V over a division ring D. Moreover, since a primitive ring is prime, by Lemma 1 (iv) it follows that a is regular. Thus, since (by our assumptions on a) $(a^nx^n)ax=a^{n+1}x^{n+1}$, $n=n(x)>1$, all $x\in R$, we obtain

$$(4) \qquad x^nax = ax^{n+1}, \quad n = n(x) > 1, \quad \text{all} \quad x\in R.$$

If $\dim_D V=1$, then $R=D$ is a division ring, so by (4), $x^{n(x)}a=ax^{n(x)}$, $n(x)>1$, all $x\in R$. By the hypercentre theorem [4] it follows that a is central. Hence we may assume $\dim_D V>1$.

Suppose that for some $v\in V$, v and va are linearly independent over D. By the density of the action of R on V, there exists an x in R with $vx=0$ and $vax=va$. Thus, by (4), $0=v(x^nax)=v(ax^{n+1})=va$, a contradiction. In other words, for every $v\in V$, v and va are linearly dependent over D. As in [4, Lemma 2] this implies that a is central.

LEMMA 3. *Let R be a prime ring with no nonzero nil right ideals. Suppose that $a\in R$ is such that for every $x\in R$ there exists an integer $n(x)>1$ with*

$$(ax)^k = a^kx^k, \quad k = n(x),\ n(x)+1,\ n(x)+2.$$

Then a is a central element.

PROOF. By Lemma 2 we may assume that $J(R)$, the Jacobson radical of R, is nonzero. Let $x\in J(R)$. We claim that $[a, x]=0$. Since R is prime, by Lemma 1 (iv) a is regular, so $y=a(1+x)$ is regular. Now, by hypothesis, there exists

$n=n(y)>1$ such that $(ay)^k=a^k y^k$, $k=n(y), n(y)+1, n(y)+2$. Hence $a^n y^n ay=$ $=a^{n+1}y^{n+1}$ and $a^{n+1}y^{n+1}ay=a^{n+2}y^{n+2}$. Since a and y are regular we obtain $ay^n=y^n a$ and $ay^{n+1}=y^{n+1}a$, and, consequently, $[a, y]=0$. Thus $0=[a, a(1+x)]=$ $=a[a, x]$. Since a is regular, $[a, x]=0$.

In other words, a centralizes the nonzero ideal $J(R)$. Since R is prime it follows that a is a central element.

PROOF OF THEOREM 2. Let $x\in R$ and let $n=n(x)>1$ such that $(ax)^k=a^k x^k$, $k=$ $=n, n+1$. Then $a^n x^n ax=a^{n+1}x^{n+1}$, so $a^n[a, x^n]x^n=0$. By Lemma 1 (ii) we get $a[a, x^n]x^n=0$. Hence, by Lemma 1 (i) and (iv), $a[a, x^n]ax^n=0$ and $[a, x^n]x^n a=0$. Thus $[a, x^n]^3=0$. Let $s\geqq 1$ be minimal such that $[a, x^n]^s=0$ (we know $1\leqq s\leqq 3$).

Suppose $s>1$. We claim that if P is a prime ideal of R, then $[a, x^n]^{s-1}\in P$.

Since, by Lemma 2, $[a, x^n]\in J(R)$, the Jacobson radical of R, we are done in case $J(R)\subset P$. Also if $\bar{R}=R/P$ has no nonzero nil right ideals we are done by Lemma 3. Thus we may assume $J(R)\not\subset P$ and $\bar{R}$ has nil right ideals $\neq 0$.

Let K be the set theoretic union of all nil right ideals of $\bar{R}$. Since $\bar{R}$ is prime and $\overline{J(R)}$, the image of $J(R)$ in $\bar{R}$, is not zero, we have that $K\cap\overline{J(R)}\neq 0$.

Let $U=\{u\in R\,|\,xuy=0$ for every $x, y\in R$ with $xy=0\}$. Then U is a subring of R invariant under automorphisms of R and, by Lemma 1 (i), $a\in U$.

Let $\bar{U}$ be the image of U in $\bar{R}$. Since U is invariant under automorphisms of R it follows that $(1+\bar{y})\bar{U}(1+\bar{y})^{-1}\subset\bar{U}$ for every $\bar{y}\in K\cap\overline{J(R)}$. As in [5] we conclude that either $\bar{U}$ is central or contains a nonzero ideal of $\bar{R}$.

If $\bar{U}$ is central it is clear that $\overline{[a, x^n]^{s-1}}\in P$, for $a\in U$. Suppose $\bar{U}$ contains an ideal $A\neq 0$ of $\bar{R}$. Then $\overline{[a, x^n]}A\overline{[a, x^n]^{s-1}}\subset\overline{[a, x^n]}\bar{U}\overline{[a, x^n]^{s-1}}=0$. Hence, since $\bar{R}$ is prime, $[a, x^n]^{s-1}\in P$. With this the claim is established.

Since R has no nonzero nil right ideals, the intersection of its prime ideals is zero. By what we deduced above it follows that $[a, x^n]^{s-1}=0$. This contradicts the minimal nature of s. Therefore $[a, x^n]=0$. In other words, a commutes with a power of every element of R. By HERSTEIN's hypercenter theorem [4] a is a central element.

References

[1] B. FELZENSZWALB, Rings radical over subrings, *Israel J. Math.*, **23 (2)** (1976), 156—164.
[2] B. FELZENSZWALB, On a result of Levitzki, to appear in *Can. Math. Bull.*
[3] I. N. HERSTEIN, Power maps in rings, *Michigan Math. J.*, **8** (1961), 29—32.
[4] I. N. HERSTEIN, On the hypercenter of a ring, *Journal of Algebra*, **36** (1975), 151—157.
[5] I. N. HERSTEIN, Invariant subrings of a certain kind, *Isreal J. Math.*, **26 (2)** (1977), 205—208.
[6] A. KAYA, On a commutative theorem of Luh, *Acta Math. Acad. Sci. Hungar.*, **22** (1971), 211—213.

(Received October 27, 1977)

INSTITUTO DE MATEMÁTICA
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
C. P. 1835—ZC—00
20.000 — RIO DE JANEIRO, RJ, BRAZIL

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (3—4), (1979), 261—265.

CONDITIONS FOR ELEMENTS TO BE CENTRAL IN A RING

By
L. O. CHUNG[1] and J. LUH (Raleigh)

Let R be an associative ring with unity element 1. In 1971, LUH [3] proved that if R is primary and if there exists a fixed non-negative integer n such that $(xy)^k=x^k y^k$, $k=n, n+1, n+2$, for all $x, y\in R$, then R is necessarily commutative. Since then this result has subsequently been generalized. KAYA [1] weakens the condition to allow n to be a function of x and y. LIGH and RICHOUX [2], [4] show by an elementary and elegant method that the result remains true without assuming that R being primary.

In this paper, we shall provide a "pointwise" version of these results. Let R be an associative ring with 1 and $x\in R$. We shall say that x *possesses the property* P if, for each $y\in R$, there exists non-negative integers n and m, both depend upon y, such that

$$(1)\qquad (xy)^k = x^k y^k, \quad k = n, n+1, n+2;$$

$$(2)\qquad (yx)^h = y^h x^h, \quad h = m, m+1, m+2.$$

We shall show that if x possesses the property P then, for each $y\in R$, there is a positive integer l which depends upon y such that $x^l y=x^{l-1}yx=x^{l-2}yx^2=\dots=yx^l$. If, in addition, R is semisimple or $1+x$ also possesses the property P, then x lies in the centre C of R. Therefore the results of [1], [2], [3], [4] immediately follow.

In what follows R will be an associative ring with 1, x will be an element in R possessing the property P. Moreover, $\mathcal{J}(R)$ will denote the Jacobson radical of R, C the centre of R and $\mathbf{Z}^+$ the set of positive integers.

We begin with the following

LEMMA 1. *Let $y\in R$ and $i, j\in \mathbf{Z}^+$.*
(i) *If $x^i y^j\in\mathcal{J}(R)$ then $xy^j\in\mathcal{J}(R)$;*
(ii) *If $y^j x^i\in\mathcal{J}(R)$ then $y^j x\in\mathcal{J}(R)$.*

PROOF. By left-right symmetry we need only prove (i). Assume that $x^i y^j\in\mathcal{J}(R)$. Suppose $xy^j\notin\mathcal{J}(R)$. Let t be the least positive integer such that $x^t y^j\in\mathcal{J}(R)$. Then $t\geqq 2$. Since x possesses the property P, for each $z\in R$, there is an integer $s\geqq 2$ such that

$$[x(x^{t-2}y^j z)]^s = x^s(x^{t-2}y^j z)^s.$$

By noting that $s+t-2\geqq t$, we have $x^s(x^{t-2}y^j z)^s\in\mathcal{J}(R)$. So $x^{t-1}y^j R$ is a quasi-regular right ideal of R. Hence $x^{t-1}y^j\in\mathcal{J}(R)$ which contradicts the minimality of t.

[1] This author is partially supported by the Engineering Foundations of NCSU, Raleigh, N. C. 27 605, USA.

LEMMA 2. *Suppose* $\mathcal{I}(R)=0$ *and* $y\in R$. *Then* $xy=0$ *if and only if* $yx=0$.

PROOF. Assume $xy=0$. Then $(yx)^2=0$. By (2), $y^h x^h=0$ for some $h\in \mathbf{Z}^+$. By Lemma 1, $y^h x=0$. Suppose to the contrary that $yx\neq 0$. Let t be the least positive integer such that $y^t x=0$. By (2) there is $k\in\mathbf{Z}^+$ such that

$$(3) \qquad (1+y)^k x^k = [(1+y)x]^k$$

and

$$(4) \qquad (1+y)^{k+1} x^{k+1} = [(1+y)x]^{k+1}.$$

By noting that $xy=0$, (3) yields

$$\left(1+ky+\frac{k(k-1)}{2!}y^2+\ldots+y^k\right)x^k = x^k+yx^k.$$

So $\left[(k-1)y+\frac{k(k-1)}{2!}y^2+\ldots+y^k\right]x^k=0$. By Lemma 1, we have

$$(5) \qquad \left[(k-1)y+\frac{k(k-1)}{2!}y^2+\ldots+y^k\right]x=0.$$

Likewise (4) yields

$$(6) \qquad \left[ky+\frac{(k+1)k}{2!}y^2+\ldots+y^{k+1}\right]x=0.$$

Subtracting (5) from (6) side-by-side, we get

$$(7) \qquad (y+a_2y^2+a_3y^3+\ldots+y^{k+1})x=0,$$

where

$$a_2=\frac{(k+1)k}{2!}-\frac{k(k-1)}{2!}, \quad a_3=\frac{(k+1)k(k-1)}{3!}-\frac{k(k-1)(k-2)}{3!},$$

etc.

Premultiplying both sides of (7) by y^{t-2} yields $y^{t-1}x=0$. This condtradicts the minimality of t. Hence $yx=0$. That $yx=0$ implies $xy=0$ can be proved similarly.

THEOREM 3. *If* $\mathcal{I}(R)=0$ *and* $x\in R$ *possesses the property P, then* $x\in C$.

PROOF. Let $y\in R$. From the first two identities of (1),

$$x^{n+1}y^{n+1}=(xy)^{n+1}=(xy)^n(xy)=x^ny^nxy,$$

and hence $x^n(y^nx-xy^n)y=0$. By Lemma 1, $x(y^nx-xy^n)y=0$. By Lemma 2, $(y^nx-xy^n)yx=0$.

Similarly, from the last two identities of (1), we obtain $(y^{n+1}x-xy^{n+1})yx=0$. Thus, $0=(y^{n+1}x-xy^{n+1})yx-y(y^nx-xy^n)yx=(yx-xy)y^{n+1}x$. Again by Lemma 2,

$$(8) \qquad x(yx-xy)y^{n+1}=0.$$

Likewise by considering $1+y$ instead of y, we obtain

$$(9) \qquad x(yx-xy)(1+y)^{q+1}=0,$$

for some $q \in \mathbf{Z}^+$. Now we claim that $x(yx-xy)=0$. Suppose not. Let t be the least positive integer such that $x(yx-xy)y^t=0$. By postmultiplying both sides of (9) by y^{t-1}, we get $x(yx-xy)y^{t-1}=0$, which contradicts the minimality of t. Thus, $x(yx-xy)=0$. Consequently, $x^2y=xyx$. Applying this we can see that for any $u, v \in R$,

$$[(xy-yx)u][(xy-yx)v] = x(yu)xyv - x(yuy)xv - yxuxyv + yx(uy)xv =$$

$$= x^2(yu)yv - x^2(yuy)v - yx^2uyv + yx^2(uy)v = 0.$$

This means that $(xy-yx)R$ is a nilpotent right ideal of R. Since $\mathscr{I}(R)=0$, $xy-yx=0$, i.e. $yx=xy$. Hence $x \in C$.

Now we proceed to the following

LEMMA 4. *If* $y \in \mathscr{I}(R)$ *then* $x^2y=xyx=yx^2$.

PROOF. As the first part in the proof of Theorem 3, we consider $1+y$ instead of y. We obtain $x(xy-yx)(1+y)^{t+1}=0$ for some $t \in \mathbf{Z}^+$. Since $1+y$ is invertible, $x(xy-yx)=0$. The left-right symmetry implies that $(xy-yx)x=0$. Thus, $x^2y=xyx=yx^2$ as we desired.

LEMMA 5. *Let* $y \in R$. *Assume that* y *satisfies the conditions* (1) *and* (2). *Then, for any* $k \in \mathbf{Z}^+$,

(i) $$x(x^{nk}y-yx^{nk})y^n = 0 = x(x^{nk+1}y-yx^{nk+1})y^{n+1};$$

(ii) $$y^m(x^{mk}y-yx^{mk})x = 0 = y^{m+1}(x^{mk+1}y-yx^{mk+1})x.$$

PROOF. From the first two identities of (1), $x^{n+1}y^{n+1}=(xy)^{n+1}=xy(xy)^n=xyx^ny^n$. So $x(x^ny-yx^n)y^n=0$. Similarly, from the last two identities of (1), we have $x(x^{n+1}y-yx^{n+1})y^{n+1}=0$. Thus (i) holds for $k=1$.

Now, we assume $k \geqq 2$. Notice that $x^ny-yx^n \in \mathscr{I}(R)$ by Theorem 3. Using Lemma 4, we obtain

$$x(x^{nk}y-yx^{nk}) = x[x^{n(k-1)}(x^ny-yx^n)+x^{n(k-2)}(x^ny-yx^n)x^n+$$

$$+\dots+x^n(x^ny-yx^n)x^{n(k-2)}+(x^ny-yx^n)x^{n(k-1)}]y^n =$$

$$= x[x^{n(k-1)}(x^ny-yx^n)+x^{n(k-1)}(x^ny-yx^n)+\dots+x^{n(k-1)}(x^ny-yx^n)]y^n = 0.$$

Thus, using Lemma 4 again, we get

$$x(x^{nk+1}y-yx^{nk+1})y^{n+1} =$$

$$= [x^{n(k-1)}(x^{n+1}y-yx^{n+1})+(x^{n(k-1)}y-yx^{n(k-1)})x^{n+1}]y^{n+1} =$$

$$= x[x^{n(k-1)}(x^{n+1}y-yx^{n+1})+x^{n+1}(x^{n(k-1)}y-yx^{n(k-1)})]y^{n+1} = 0.$$

This completes the proof of (i). (ii) can be proved similarly.

LEMMA 6. *For each* $y \in R$, *there exists* $t \in \mathbf{Z}^+$ *such that*

$$x(x^ty-yx^t) = 0 \quad \textit{and} \quad x(x^{t+1}y-yx^{t+1}) = 0.$$

PROOF. Suppose y satisfies the identities (1) and (2) and $1+y$ satisfies

(10) $$[x(1+y)]^k = x^k(1+y)^k, \quad k = p, p+1, p+2;$$

and

(11) $$[(1+y)x]^h = (1+y)^h x^h, \quad h = q, q+1, q+2,$$

where $p, q \in \mathbf{Z}^+$. By Lemma 5,

(12) $$x(x^{np}y - yx^{np})y^n = 0;$$

(13) $$x(x^{np+1}y - yx^{np+1})y^{n+1} = 0;$$

(14) $$x(x^{np}y - yx^{np})(1+y)^p = 0;$$

(15) $$x(x^{np+1}y - yx^{np+1})(1+y)^{p+1} = 0.$$

We assert that $x(x^{np}y-yx^{np})=0$. Suppose not. Then by (12) there is a least positive integer r such that $x(x^{np}y-yx^{np})y^r=0$.
Postmultiplying both sides of (14) by y^{r-1} yields $x(x^{np}y-yx^{np})y^{r-1}=0$, a contradiction. So $x(x^{np}y-yx^{np})=0$. That $x(x^{np+1}-yx^{np+1})=0$ can be proved similarly.

Now we are in the position to prove our main result.

THEOREM 7. *Let R be a ring with 1. If $x\in R$ possesses the property P, then, for each $y\in R$, there exists $l\in \mathbf{Z}^+$ such that* $x^l y = x^{l-1}yx = x^{l-2}yx^2 = \dots = yx^l$.

PROOF. According to Lemma 6, there is $t\in\mathbf{Z}^+$ such that $x(x^t y - yx^t)=0$ and $x(x^{t+1}y-yx^{t+1})=0$. Since $x^t y - yx^t \in \mathcal{J}(R)$,

$$(x^t y - yx^t)x^2 = x(x^t y - yx^t)x = x^2(x^t y - yx^t) = 0$$

and, similarly,

$$(x^{t+1}y - yx^{t+1})x^2 = x(x^{t+1}y - yx^{t+1})x = x^2(x^{t+1}y - yx^{t+1}) = 0$$

by Lemma 4. Thus $x^t yx^2 = yx^{t+2}$, $x^{t+1}yx = xyx^{t+1}$, $x^{t+2}y = x^2yx^t$, $x^{t+1}yx^2 = yx^{t+3}$, $x^{t+2}yx = xyx^{t+2}$, and $x^{t+3}y = x^2yx^{t+1}$. It follows that $x^{t+3}y = x^2yx^{t+1} = x\cdot xyx^{t+1} = x\cdot x^{t+1}yx = x^{t+2}yx$ and $yx^{t+3} = x^{t+1}yx^2 = x^{t+1}yx\cdot x = xyx^{t+1}\cdot x = xyx^{t+2}$. Now let $l=2t+4$.

For $0\leqq i\leqq t+1$, $x^i yx^{l-i} = x^i yx^{2t+4-i} = x^i yx^{t+3}\cdot x^{t+1-i} = x^i xyx^{t+2}\cdot x^{t+1-i} = x^{i+1}yx^{2t+3-i} = x^{i+1}yx^{l-i-1}$.

For $t+1<i<2t+4=l$, $x^i yx^{l-i} = x^i yx^{2t+4-i} = x^{i-t-2}\cdot x^{t+2}yx\cdot x^{2t+3-i} = x^{i-t-2}\cdot x^{t+3}yx^{2t+3-i} = x^{i+1}yx^{2t+3-i} = x^{i+1}yx^{l-i-1}$. This completes the proof.

COROLLARY. *Let $x\in R$. If x and $1+x$ both possess the property P, then $x\in C$.*

PROOF. Suppose to the contrary that $x\notin C$. Then there exists $y\in R$ such that $xy\neq yx$. By Theorem 7, there exist $l, r\in\mathbf{Z}^+$ such that $x^l y = x^{l-1}yx$ and $(1+x)^r y = (1+x)^{r-1}y(1+x)$. That is,

(16) $$x^{l-1}(xy-yx) = 0,$$

and

(17) $$(1+x)^{r-1}(xy-yx) = 0.$$

Since $xy-yx\neq 0$, there is a least positive integer t such that $x^t(xy-yx)=0$. Premultiplying both sides of (17) by x^{t-1} yields $x^{t-1}(xy-yx)=0$ which contradicts the minimality of t. Hence $xy=yx$.

In conclusion, we should note that if $x\in R$ satisfies *only* one of the conditions (1) and (2) the conclusions of Theorem 7 need not be true. This can be seen easily from the example that

$$R=\left\{\begin{bmatrix} a & b\\ 0 & c\end{bmatrix}\middle|\, a, b, c\in \mathbf{Z}_2\right\}$$

being the ring of upper triangular matrices over the ring $\mathbf{Z}_2$ of integers and $x=\begin{bmatrix} 1 & 0\\ 0 & 0\end{bmatrix}$.

References

[1] A. Kaya, On a commutativity theorem of Luh, *Acta Math. Acad. Sci. Hungar.*, **28** (1976), 33—36.

[2] S. Ligh and A. Richoux, A commutativity theorem for rings, *Bull. Austral. Math. Soc.*, **16** (1977), 75—77.

[3] J. Luh, A commutativity theorem for primary rings, *Acta Math. Acad. Sci. Hungar.*, **22** (1971), 211—213.

[4] A. Richoux, On a commutativity theorem of Luh (to appear).

(Received October 1, 1977)

DEPARTMENT OF MATHEMATICS
NORTH CAROLINA STATE UNIVERSITY
RALEIGH, N. C. 27650
USA

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (3—4), (1979), 267—277.

ϱ-NORMAL POINT SYSTEMS

By
P. VÉRTESI (Budapest)

1. Introduction

The aim of this paper is to construct some new ϱ-normal point systems. For this aim we establish some new estimations for the distance of the Jacobi roots, too.

2.1. Definitions and preliminary results

2.1. The Hermite—Fejér interpolatory polynomial $H_n(f; x)$ of degree $\leqq 2n-1$ for a function $f(x)$ on the nodes

(2.1) $$-1 \leqq x_{nn} < x_{n-1,n} < \dots < x_{1n} \leqq 1$$

is uniquely defined by

(2.2) $$H_n(f;\, x_{kn}) = f(x_{kn}), \quad H_n'(f;\, x_{kn}) = 0 \quad (k = 1, 2, \dots, n).$$

If

(2.3) $$\omega_n(x) = c_n(x-x_{1n})(x-x_{2n})\dots(x-x_{nn}),$$

(2.4) $$l_{k,n}(x) = \frac{\omega_n(x)}{\omega_n'(x_{kn})(x-x_{kn})} \quad (k = 1, 2, \dots, n),$$

finally

(2.5) $$v_{kn}(x) = 1 - \frac{\omega_n''(x_{kn})}{\omega_n'(x_{kn})}(x-x_{kn}) \quad (k = 1, 2, \dots, n),$$

we have

(2.6) $$H_n(f;\, x) = \sum_{k=1}^{n} f(x_{kn})\, v_{kn}(x)\, l_{kn}^2(x)$$

(see, e.g. [1]).

2.2. As G. GRÜNWALD [2] proved, $H_n(f; x)$ uniformly tends to $f(x)$ in $[-1, 1]$ (if $n\to\infty$), whenever f is continuous on $[-1, 1]$, moreover the point system (2.1) is ϱ-normal, i.e. for a certain $\varrho > 0$ and $n > n_0$

(2.7) $$v_{kn}(x) \geqq \varrho \quad (k = 1, 2, \dots, n;\ x\in[-1, 1]).$$

If the nodes form a *normal* system, i.e. (2.7) is true with $\varrho = 0$, the statement holds for arbitrary $[a, b]\subset(-1, 1)$.

2.3. Until quite recently the only known ϱ-normal and normal systems were the roots $\{x_{kn}^{(\alpha,\beta)}\}$ of the Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ supposing $-1\leqq\alpha, \beta<0$ (here

$\varrho=\min(-\alpha, -\beta))$ and the roots of the Legendre polynomials $P_n(x)=P_n^{(0,0)}(x)$ (see, e.g. [2]).

In 1972, G. FREUD [2] proved that the roots of

$$P_n(x)+A_nP_{n-1}(x) \quad (|A_n|\leqq 1) \tag{2.8}$$

form a normal point system, and later in [3], I found that the nodes

$$\{x_{kn}^{(\alpha,\beta)}\}_{k=1}^{n}\cup\{1\} \quad (1\leqq\alpha\leqq 2;\ -1\leqq\beta\leqq 0), \tag{2.9}$$

$$\{x_{kn}^{(\alpha,\beta)}\}_{k=1}^{n}\cup\{-1\} \quad (-1\leqq\alpha\leqq 0,\ 1\leqq\beta\leqq 2), \tag{2.10}$$

$$\{x_{kn}^{(\alpha,\beta)}\}_{k=1}^{n}\cup\{-1, 1\} \quad (1\leqq\alpha\leqq 2,\ 1\leqq\beta\leqq 2) \tag{2.11}$$

form ϱ-normal systems with $\varrho=\min[-(\alpha-2), -\beta]$, $\varrho=\min[-\alpha, -(\beta-2)]$ and $\varrho=\min[-(\alpha-2), -(\beta-2)]$, respectively (if $\varrho=0$, we obtain normal system).

3. Results

3.1. We wish to construct some new ϱ-normal systems, using the older ones. First we may try to move the original roots. Obviously, the new nodes again form a $\varrho/2$-normal system if the shift is small enough. Now we characterize this shift. Using the notations $x=\cos\vartheta$, $x_{kn}=\cos\vartheta_{kn}$ $(k=1, 2, \dots, n)$, $x_{0n}=\cos\vartheta_{0n}=1$ and $x_{n+1,n}=\cos\vartheta_{n+1,n}=-1$, we prove

THEOREM 3.1. *Let us suppose that the matrix of nodes $X=\{x_{kn}\}$ $(k=1, 2, \dots, n;$ $n=1, 2, \dots)$ form a ϱ-normal system. Then, substituting x_{sn} by y_{sn} $(n=1, 2, \dots,$ $1\leqq s=s(n)\leqq n,$ $|y_{sn}|\leqq 1)$, the new matrix is a $\varrho/2$-normal system, supposing that with a suitable $c=c(X)>0$ we have*

$$|x_{sn}-y_{sn}|\leqq c\max\left(\frac{\sin^2\vartheta_{sn}}{n^2}, \frac{1}{n^4}\right) \quad (n=1, 2, \dots). \tag{3.1}$$

Here the order of the right-hand side, for special X, is the best possible for arbitrary sequence $\{s(n)\}$ provided $2\leqq s(n)\leqq n-1$.

3.2. Using special nodes, we can get more, at least for certain values of s. For this aim we verify a statement which is interesting in itself.

LEMMA 3.2. *For the roots $x_{kn}^{(\alpha,\beta)}=\cos\vartheta_{kn}^{(\alpha,\beta)}$ of $P_n^{(\alpha,\beta)}(x)$ we have the inequalities*

$$\frac{k-1}{n-1}\pi\leqq\vartheta_{kn}^{(\alpha,\alpha)}\leqq\frac{k-1/2}{n}\pi \tag{3.2}$$

if $-1\leqq\alpha\leqq-0.5$, $(k=1, 2, \dots, [n/2])$, *further*

$$\frac{k-1}{n+1/2}\pi\leqq\vartheta_{kn}^{(\alpha,\beta)}\leqq\frac{k-1/2}{n}\pi \tag{3.3}$$

if $-1 \leqq \alpha \leqq -0.5$ *and* $-0.5 \leqq \beta \leqq 1$,

$$\frac{k-1}{n+1/2}\pi \leqq \vartheta_{kn}^{(\alpha,\beta)} \leqq \frac{k-1/2}{n+1/2}\pi \tag{3.4}$$

if $-1 \leqq \alpha \leqq -0.5$ *and* $0.5 \leqq \beta \leqq 1$, *for* $k=1, 2, \ldots, n$.
Using the lemma, we obtain

STATEMENT 3.3. *For fixed values* α *and* β *the nodes*

$$\{1, x_{2n}^{(\alpha,\beta)}, x_{3n}^{(\alpha,\beta)}, \ldots, x_{n-1,n}^{(\alpha,\beta)}, x_{n,n}^{(\alpha,\beta)}\} \quad (n=1, 2, \ldots) \tag{3.5}$$

form a ϱ*-normal system if*

$$-1 \leqq \alpha < -\frac{2}{3} \quad \text{and} \quad -1 \leqq \beta < 0. \tag{3.6}$$

3.3. For the roots of $P_n^{(-1/2,\beta)}(x)$ (especially for the Chebyshev-nodes), we get the following

STATEMENT 3.4. *The nodes*

$$\{y_{1n} = \cos\eta_{1n}, x_{kn}^{(-1/2,\beta)} \quad (k=2, 3, \ldots, n)\} \quad (n=1, 2, \ldots) \tag{3.7}$$

form a ϱ*-normal system, whenever* $-1 \leqq \beta < 0$ *and*

$$0 \leqq \eta_{1n} < d\vartheta_{1n}^{(-1/2,\beta)} \quad (n=1, 2, \ldots). \tag{3.8}$$

Here $d=1.3672\ldots$ *and it cannot be replaced by* $d+\varepsilon$ $(\varepsilon>0)$.

3.4. By similar method we can get

STATEMENT 3.5. *If* $-1 \leqq \alpha = \beta < -2/3$, *then the nodes*

$$\{1, x_{2n}^{(\alpha,\alpha)}, x_{3n}^{(\alpha,\alpha)}, \ldots, x_{n-1,n}^{(\alpha,\alpha)}, -1\} \quad (n=1, 2, \ldots) \tag{3.9}$$

form a ϱ*-normal system.*
Finally, for the Chebyshev roots we obtain

STATEMENT 3.6. *The system*

(3.10)

$$\left\{y_{1n} = \cos\eta_{1n}, \cos\frac{2k-1}{2n}\pi \quad (k=2, 3, \ldots, n-1), \; y_{nn} = \cos\eta_{nn}\right\} \quad (n=1, 2, \ldots)$$

is ϱ*-normal whenever*

$$0 \leqq \eta_{1n} = -\eta_{nn} < d\frac{\pi}{2n} \quad (n=1, 2, \ldots). \tag{3.11}$$

Here $d=1.3672\ldots$ *and it cannot be replaced by* $d+\varepsilon$ $(\varepsilon>0)$.

3.5. REMARKS 1. According to the corresponding consideration, the systems (3.7) and (3.10) are ϱ_1-normal for $\varrho_1 < 1/4$ if $y_{1n}=1$ and $\alpha=\beta=-1/2$.

2. Statements analogous to Statements 3.3 and 3.4 can be obtained shifting x_n instead of x_1.

3. A simple calculation shows that the usual estimations for the Jacobi roots (see [5], 6.21 and 6.3) do not lead to results like in 3.2—3.4.

4. Proofs

4.1. Proof of Theorem 3.1. For the matrix X we have by (2.4)

$$\omega_n(x)=(x-x_{sn})\,l_{sn}(x). \tag{4.1}$$

Using this, we have, sometimes omitting the superfluous notations,

$$-\frac{\omega''(x_k)}{\omega'(x_k)}=\frac{2}{x_s-x_k}-\frac{l_s''(x_k)}{l_s'(x_k)}\quad (k\neq s), \tag{4.2}$$

$$-\frac{\omega''(x_s)}{\omega'(x_s)}=-2\frac{l_s'(x_s)}{l_s(x_s)}=-2l_s'(x_s). \tag{4.3}$$

For the new system we can write

$$\Omega_n(x)\overset{\text{def}}{=}c_n(x-y_s)\prod_{k\neq s}(x-x_k)=(x-y_s)\,l_s(x), \tag{4.4}$$

i.e.

$$-\frac{\Omega''(x_k)}{\Omega'(x_k)}=\frac{2}{y_s-x_k}-\frac{l_s''(x_k)}{l_s'(x_k)}\quad (k\neq s), \tag{4.5}$$

$$-\frac{\Omega''(y_s)}{\Omega'(y_s)}=-\frac{2l_s'(y_s)}{l_s(y_s)}. \tag{4.6}$$

By (4.1)—(4.6) we obtain

$$v_k(\Omega;\,x)=v_k(\omega;\,x)+\frac{2(x_s-y_s)}{(y_s-x_k)(x_s-x_k)}(x-x_k)\quad (k\neq s), \tag{4.7}$$

$$v_s(\Omega;\,x)=v_s(\omega;\,x)+2\left[l_s'(x_s)(x-x_s)-\frac{l_s'(y_s)}{l_s(y_s)}(x-y_s)\right]. \tag{4.8}$$

Now we show that $v_k(\Omega;\,\pm 1)\geqq \varrho/2$ $(k=1, 2, \dots, n)$ which proves the theorem. First we quote a lemma established by P. Erdős and P. Turán [4].

Lemma 4.1 ([4], Part 4). *If for the matrix* $X=\{x_{kn}\}$

$$|v_{kn}(x)|\,l_{kn}^2(x)\leqq K\quad (k=1, 2, \dots, n;\; n=1, 2, \dots), \tag{4.9}$$

then for certain $M=M(X)\geqq m=m(X)>0$ *we have*

$$\frac{m}{n}\leqq \vartheta_{k+1,n}-\vartheta_{kn}\leqq \frac{M}{n}, \tag{4.10}$$

where the upper estimate holds for $k=0, 1, 2, \dots, n,$ *and the lower one for* $k=1, 2, \dots, n-1$.

Remark that the lemma can be applied because $\sum_{k=1}^{n} v_k(x)\,l_k^2(x)\equiv 1$ and $v_k(x)\geqq\varrho$.

By this lemma it is easy to obtain

$$m_1 \frac{\sin \vartheta_k}{n} \leqq x_k - x_{k+1} \leqq M_1 \frac{\sin \vartheta_k}{n} \quad (k = 1, 2, \dots, n-1) \tag{4.11}$$

where $M_1 \geqq m_1 > 0$.

Henceforward suppose $n \sin \vartheta_s \geqq 1$ and $x_s < y_s$. Then, by (4.7) and (4.11), we get for $k \neq s$ and $s \neq n$

$$v_k(\Omega; \pm 1) \geqq \varrho - 2 \frac{|x_s - y_s|}{|y_s - x_k||x_s - x_k|} |x_k \pm 1| \geqq \varrho - \frac{8|x_s - y_s|}{|y_s - x_s + x_s - x_{s+1}||x_s - x_{s+1}|} \geqq$$

$$\geqq \varrho - 8c \frac{\sin^2 \vartheta_s}{n^2} \frac{1}{(x_s - x_{s+1})^2} \geqq \varrho - \frac{8c}{m_1^2} \frac{\sin^2 \vartheta_s}{n^2} \frac{n^2}{\sin^2 \vartheta_s} \geqq \varrho/2$$

if c is small enough. Analogous argument holds for $s = n$, when $v_k(\Omega; \pm 1) \geqq \geqq v_{n-1}(\Omega; \pm 1)$ $(k \neq n)$.

Let now $k = s$. By (4.8) we get, after a little calculation,

$$v_s(\Omega; \pm 1) \geqq \varrho - 2 \{ [|y_s \pm 1| |l_s'(y_s)| |l_s(x_s) - l_s(y_s)|] \frac{1}{|l_s(y_s)|} +$$

$$+ |x_s \pm 1| |l_s'(x_s) - l_s'(y_s)| + |y_s - x_s| |l_s'(y_s)| \}.$$

Using the ϱ-normality we have from $\sum_{k=1}^{n} v_k(x) l_k^2(x) \equiv 1$

$$|l_s(x)| \leqq \frac{1}{\sqrt{\varrho}}. \tag{4.12}$$

Further by (4.12), using the Bernstein-inequality, we get

$$|l_s'(x)| < \frac{1}{\sqrt{\varrho}} \frac{n}{\sin \vartheta}, \quad |l_s''(x)| < \frac{1}{\varrho} \frac{n^2}{\sin^2 \vartheta}, \tag{4.13}$$

from where $0.5 \leqq l_s(y_s) \leqq 2$ because of $|l_s(y_s) - l_s(x_s)| = |l_s'(\xi_1)||y_s - x_s| \leqq \leqq \frac{c}{\sqrt{\varrho}} \frac{n}{\sin \vartheta_s} \frac{\sin^2 \vartheta_s}{n^2} = \frac{c \sin \vartheta_s}{\sqrt{\varrho} n}$.

Moreover, using

$$|l_s'(y_s) - l_s'(x_s)| = |l_s''(\xi_2)||y_s - x_s| < \frac{c}{\varrho} \frac{n^2}{\sin^2 \vartheta_s} \frac{\sin^2 \vartheta_s}{n^2} = \frac{c}{\varrho},$$

we obtain with $c_1 = c_1(c, \varrho)$

$$v_s(\Omega; \pm 1) \geqq \varrho - 4 \left[c_1 \frac{n}{\sin \vartheta_s} \frac{\sin \vartheta_s}{n} + c_1 + c_1 \frac{\sin^2 \vartheta_s}{n^2} \frac{n}{\sin \vartheta_s} \right] \geqq \varrho/2$$

for suitable $c > 0$. The cases $y_s < x_s$ and $n \sin \vartheta_s < 1$ can be treated similarly, considering the Markov-inequality, too. So we proved the estimation (3.1).

To obtain the second part of our statement we consider the matrix $X=\{x_{kn}^{(-1,-1)}\}$. Then as it is well-known $v_k(\omega; x)\equiv 1$ if $2\leqq k\leqq n-1$ (see e.g. [5], 4.22, (4.2.1) and (14.5.1)). Let e.g., $2\leqq s\leqq n/2$ and $x_s>y_s+K\sin^2\vartheta_s n^{-2}$. Then by (4.7) and (4.10)

$$v_{s+1}(\Omega; -1)\leqq 1-2\frac{K\sin^2\vartheta_s}{n^2}K_1(m, M)\frac{n}{\sin\vartheta_s}\frac{n}{\sin\vartheta_s}<0$$

if K is large enough. The case $n/2<s\leqq n-1$ is similar, Q.E.D.

4.2. Proof of Lemma 3.2. Using continuity and [5], 6.21, one can obtain $\vartheta_{kn}^{(\alpha,\alpha)}>\vartheta_{kn}^{(\alpha^*,\alpha^*)}$ if $\alpha>\alpha^*\geqq -1$ and $k=1, 2, \dots [n/2]$. Further, by $P_{n+2}^{(-1,-1)}(x)=$ $=c_n(1-x^2)P_n^{(1,1)}(x)$ we can write

$$\frac{k\pi}{n+1}\equiv\vartheta_{kn}^{(1/2,1/2)}<\vartheta_{k,n}^{(1,1)}\equiv\vartheta_{k+1,n+2}^{(-1,-1)}\leqq\vartheta_{k+1,n+2}^{(\alpha,\alpha)}\leqq\vartheta_{k+1,n+2}^{(-1/2,-1/2)}\equiv\frac{2k+1}{2(n+2)}\pi$$

if $k=1, 2, \dots, [n/2]$, $-1\leqq\alpha\leqq -1/2$. If $k=0$ then, for the same α's,

$$0=\vartheta_{1,n+2}^{(-1,-1)}\leqq\vartheta_{1,n+2}^{(\alpha,\alpha)}\leqq\vartheta_{1,n+2}^{(-1/2,-1/2)}\equiv\frac{\pi}{2(n+2)}.$$

By these we obtain (3.2).

To get (3.3) and (3.4) we write, using similar arguments,

$$\frac{2k}{2n+1}\pi\equiv\vartheta_{kn}^{(1/2,-1/2)}<\vartheta_{kn}^{(1,-1)}\equiv\vartheta_{k+1,n}^{(-1,1)}\leqq\vartheta_{k+1,n}^{(\alpha,\beta)}\leqq\vartheta_{k+1,n}^{(-1/2,-1/2)}\equiv\frac{2k+1}{2n}\pi$$

if $-1\leqq\alpha\leqq -1/2$, $-1/2\leqq\beta\leqq 1$ and $k=1, 2, \dots, n-1$. or

$$\frac{2k\pi}{2n+1}\equiv\vartheta_{kn}^{(1/2,-1/2)}<\vartheta_{kn}^{(1,-1)}\equiv\vartheta_{k+1,n}^{(-1,1)}\leqq\vartheta_{k+1,n}^{(\alpha,\beta)}\leqq\vartheta_{k+1,n}^{(-1/2,1/2)}\equiv\frac{2k+1}{2n+1}\pi$$

if $-1\leqq\alpha\leqq -1/2$, $1/2\leqq\beta\leqq 1$ and $k=1, 2, \dots, n-1$,
The case $k=0$ can be treated as above. So we verified the whole lemma.

4.3. Proof of Statement 3.3. For the system (3.5) we have

$$\Omega(x)\stackrel{\text{def}}{=\!=}c_n(x-1)\prod_{k=2}^{n}(x-x_{kn}^{(\alpha,\beta)})=P_n^{(\alpha,\beta)}(x)\frac{x-1}{x-x_{1n}^{(\alpha,\beta)}}. \tag{4.15}$$

Then, by a simple computation

$$\frac{\Omega''(1)}{\Omega'(1)}=2\left[\frac{P_n'(1)}{P_n(1)}-\frac{1}{1-x_1}\right]. \tag{4.16}$$

Obviously $v_0(\Omega; 1)=1$ (where, as usual, $x_0\equiv 1$) so to prove that $v_0(\Omega; x)\geqq\varrho$, it is enough to verify $v_0(\Omega; -1)\geqq\varrho$. To this end we state that

$$\frac{P_n'(1)}{P_n(1)}-\frac{1}{1-x_1}>0. \tag{4.17}$$

Indeed, we have

$$\frac{1}{1-x_1} = \frac{P_n'(\xi)}{P_n(1)-P_n(x_1)} < \frac{P_n'(1)}{P_n(1)} \quad (x_1 < \xi < 1)$$

because, by [5], (4.21.7) and Part 7.32 (2),

$$P_n'^{(\alpha,\beta)}(\xi) = 0.5(n+\alpha+\beta+1)P_{n-1}^{(\alpha+1,\beta+1)}(\xi) < 0.5(n+\alpha+\beta+1)P_{n-1}^{(\alpha+1,\beta+1)}(1) =$$
$$= P_n'^{(\alpha,\beta)}(1).$$

By (4.17) $v_0(\Omega;\,-1)=1-2[...](-1-1)>1$ which was stated.

To estimate the remaining $v_k(\Omega;\,x)$ $(k\geqq 2)$ we use (4.7). Obviously with $P_n^{(\alpha,\beta)}(x)=\omega(x)$ and [5], (14.5.3)

$$v_k(\Omega;\,-1) = v_k(\omega;\,-1)+2\frac{(x_1-1)(-1-x_k)}{(1-x_k)(x_1-x_k)} \geqq v_k(\omega;\,-1) \geqq -\beta.$$

On the other hand, if $x=1$, then we have from (4.7)

$$v_k(\Omega;\,1) = v_k(\omega;\,1)-2\frac{1-x_1}{x_1-x_k} \quad (k=2,3,\ldots,n). \tag{4.18}$$

First let $\alpha=\beta$. By (3.2) $\vartheta_1 \leqq \frac{\pi}{2n}$ and $\vartheta_2 \geqq \frac{\pi}{n-1}$, so

$$2\frac{1-x_1}{x_1-x_k} \leqq 2\frac{2\left(\sin\frac{\pi}{4n}\right)^2}{2\sin\left(\frac{\pi}{4n}+\frac{\pi}{2n-2}\right)\sin\left(\frac{\pi}{2n-2}-\frac{\pi}{4n}\right)} \leqq \frac{2}{3-\frac{c}{n^2}},$$

i.e. by (4.18)

$$v_k(\Omega;\,1) > -\alpha - \frac{2}{3-\frac{c}{n^2}} > \varrho > 0 \quad (n \geqq n_0)$$

if $-1<\alpha<-2/3$, which was stated.

If $-1\leqq\alpha<-2/3$ and $-1\leqq\beta<0$, then using that $\vartheta_i=(j_i^{(\alpha)}+\varepsilon_n)n^{-1}$ $(i=1,2)$, where $j_i^{(\alpha)}$ denote the positive zeros of $J_\alpha(z)$ (the Bessel function of the first kind of order α) and $\varepsilon_n\to 0$ (see [5], (8.1.1) and (8.1.3)) we obtain $\vartheta_1\leqq\frac{\pi+|\varepsilon_n|}{2n}$ and $\vartheta_2\geqq\frac{\pi-|\varepsilon_n|}{n-1}$, from where we get the statement.

4.4. PROOF OF STATEMENT 3.4. First let $\alpha=\beta=-1/2$. As in 4.3, we obtain with $T_n(x)=\cos n\vartheta$, excluding the trivial case $y_1=\cos\frac{\pi}{2n}$,

$$v_1(\Omega;\,x) = 1-2\left[\frac{T_n'(y_1)}{T_n(y_1)}-\frac{1}{y_0-x_1}\right](x-y_1). \tag{4.19}$$

Here the expression $[...]\geqq 0$ if $y_1\geqq\cos\frac{\pi}{n}$. (This can be verified as (4.17); $\cos\frac{\pi}{n}$ is the largest root of $T_n'(x)$.) So $v_1(\Omega;\,-1)\geqq 1$, i.e. we have to investigate $v_1(\Omega;\,1)$.

Denoting $y_{1n}=\cos c\frac{\pi}{2n}$ $(0<c\leqq 2;\ c\neq 1)$ we get by simple calculation

$$v_1(\Omega;\,1)=1-2\left[\frac{n\sin c\dfrac{\pi}{2}}{\sin c\dfrac{\pi}{2n}\cos c\dfrac{2}{\pi}}-\frac{1}{2\sin(c+1)\dfrac{\pi}{4n}\cdot\sin(1-c)\dfrac{\pi}{4n}}\right]2\sin^2 c\frac{\pi}{4n},$$

from where $v_1(\Omega;\,1)\geqq\varrho>0$ $(n\geqq n_0)$ whenever

$$F(c)=\left[\frac{1}{c}\frac{\sin c\dfrac{\pi}{2}}{\cos c\dfrac{\pi}{2}}-\frac{4}{\pi(1-c^2)}\right]\frac{c^2\pi}{2}<1-\varrho. \tag{4.20}$$

Let $I_1=(0,0.58)$, $I_2=[0.58,0.67)$, $I_3=[0.67,0.75)$, $I_4=[0.75,0.82)$, $I_5=[0.82,0.83)$, $I_6=[0.83,0.96)$, $I_7=[0.97,1)$ and finally $J=(1,1.3672]$.
To estimate (4.20) we apply

$$\cos c\frac{\pi}{2}-\sin(1-c)\frac{\pi}{2}>(1-c)\frac{\pi}{2}\left\{1-\left[(1-c)\frac{\pi}{2}\right]^2\frac{1}{6}\right\}\quad\text{if}\quad 0<c<1$$

further the estimations $\sin c\frac{\pi}{2}\leqq c\frac{\pi}{2}$ or $\sin c\frac{\pi}{2}\leqq 1$ if $c\in I_1$, or $c\notin I_1$, respectively.

Using these estimations in the intervalls $I_1, I_2, \dots, I_7$ and the analogous ones for $c\in J$ we obtain (4.20) if

$$0<c\leqq 1.3672\quad(c\neq 1). \tag{4.21}$$

Further, considering that the case $c=0$ was investigated in Statement 3.3, we proved

$$v_1(\Omega;\,x)\geqq\varrho>0\quad\text{if}\quad 0\leqq c\leqq 1.3672. \tag{4.22}$$

Remark that $F(1.3673)>1$, i.e. $F(d)=1$ for $1.3672<d<1.3673$.

Now we investigate $v_k(\Omega;\,x)$ $(k\geqq 2)$. First let $0\leqq c\leqq 1$. We obtain by (4.7) (with $\omega_n(x)=T_n(x)$ and $s=1$) that $v_k(\Omega;\,-1)\geqq v_k(T;\,-1)>1/2$ $(k\geqq 2)$. Further, for $v_k(\Omega;\,1)$, we have

$$\frac{2(y_1-x_1)(1-x_k)}{(y_1-x_k)(x_1-x_k)}=$$

$$=2\,\frac{2\sin(1+c)\dfrac{\pi}{4n}\sin(1-c)\dfrac{\pi}{4n}\cdot 2\sin^2\left[(2k-1)\dfrac{\pi}{4n}\right]}{2\sin(2k-1+c)\dfrac{\pi}{4n}\sin(2k-1-c)\dfrac{\pi}{4n}\cdot 2\sin 2k\dfrac{\pi}{4n}\cdot\sin(2k-2)\dfrac{\pi}{4n}}=$$

$$=\frac{(1-c^2)(2k-1)^2}{2[(2k-1)^2-c^2]k(k-1)}[1+O(n^{-2})]\leqq\frac{9}{4}\frac{1-c^2}{9-c^2}[1+O(n^{-2})]\leqq\frac{1+O(n^{-2})}{4},$$

which means $v_k(\Omega;\,1)>\frac{1}{2}-\frac{1}{4}+O(n^{-2})=\frac{1}{4}+O(n^{-2})$ $(n\geqq n_0,\ 0\leqq c\leqq 1)$.

On the other hand if $c>1$, by (4.7) we get $v_k(\Omega; 1) \geqq v_k(T; 1) > \frac{1}{2}$. So we have to investigate $v_k(\Omega; -1)$. By (4.7), [5], 14.5.3 and the above argument we get

$$v_k(\Omega; -1) = v_k(T; -1) - 2\frac{(x_1-y_1)(1+x_k)}{(y_1-x_k)(x_1-x_k)} =$$

$$= \frac{1}{1-x_k} - 2\frac{(x_1-y_1)(1+x_k)}{(y_1-x_k)(x_1-x_k)} = \frac{1}{1-x_k}\left[1-2\frac{(x_1-y_1)(1+x_k)(1-x_k)}{(y_1-x_k)(x_1-x_k)}\right].$$

Here

$$[\ldots] \geqq [\ldots]_{x_k=x_2} \geqq$$

$$\geqq (1-x_2)\left\{\frac{1}{1-x^2} - 4\frac{x_1-y_1}{(y_1-x_2)(x_1-x_2)}\right\} > (1-x_2)\frac{8n^2}{\pi^2}\left\{\frac{1}{9} - \frac{c^2-1}{2(9-c^2)}\right\}\{1+O(n^{-2})\} > \varrho,$$

if, e.g. $0<c\leqq 1.5$.

By (4.22) and the succeeding arguments we get our statement if $\alpha=\beta=-0.5$.

Now let $-1\leqq\beta<0$. Denoting for fixed β and c $(0\leqq c<d,\ c\neq 1)$ $\tilde{y}_1 = \cos c\tilde{\vartheta}_1$, where $\tilde{\vartheta}_1 = \vartheta_{1n}^{(-1/2,\beta)}$, we investigate

$$v_1(\Omega; 1) = 1-2\left[\frac{P_n'^{(-1/2,\beta)}(\tilde{y}_1)}{P_n^{(-1/2,\beta)}(\tilde{y}_1)} - \frac{1}{\tilde{y}_1-\tilde{x}_1}\right](1-\tilde{y}_1)$$

(here $\tilde{x}_1 = x_1^{(-1/2,\beta)}$) and verify

$$\left[\frac{P_n'(\tilde{y}_1)}{P_n(\tilde{y}_1)} - \frac{1}{\tilde{y}_1-\tilde{x}_1}\right](1-\tilde{y}_1) = (1\pm\xi_n)\left[\frac{T_n'(y_1)}{T_n(y_1)} - \frac{1}{y_1-x_1}\right](1-y_1)+o(1), \tag{4.23}$$

$(\xi_n \searrow 0)$ from where $v_1(\Omega; 1)\geqq\varrho>0$ if $n\geqq n_0$ (see (4.19) and the succeeding arguments.) To obtain (4.23) we first apply [5], (8.1.1) and (1.71.2) from where

$$\left|\sqrt{n}\,P_n^{(-1/2,\beta)}(\tilde{y}_1) - \frac{\cos c\tilde{\vartheta}_1 n}{\sqrt{\pi}}\right| \leqq \delta_n \quad (\delta_n \to 0). \tag{4.24}$$

By (4.24), using $\tilde{\vartheta}_1 = (\pi+\varepsilon_n)/(2n)$ $(\varepsilon_n\to 0)$ we obtain, by $P_n(\tilde{y}_1)\sim T_n(y_1)\sim 1$,

$$P_n^{(-1/2,\beta)}(\tilde{y}_1) = \frac{T_n(y_1)}{\sqrt{n\pi}}(1+\delta_n') \quad (|\delta_n'|\to 0). \tag{4.25}$$

By similar arguments

$$P_n'^{(-1/2,\beta)}(\tilde{y}_1) = \frac{T_n'(y_1)}{\sqrt{n\pi}}(1+\delta_n'') \quad (|\delta_n''|\to 0). \tag{4.26}$$

If we use the simple relations

$$\tilde{y}_1-\tilde{x}_1 = (y_1-x_1)(1+\delta_n''') \quad 1-\tilde{y}_1 = (1-y_1)(1+\delta_n'''') \tag{4.27}$$

($|\delta_n''''|\to 0$, $|\delta_n'''''|\to 0$), we can write by (4.24)–(4.27)

$$\left[\frac{P_n'(\tilde y_1)}{P_n(\tilde y_1)}-\frac{1}{\tilde y_1-\tilde x_1}\right](1-\tilde y_1)=\left[\frac{T_n'(y_1)}{T_n(y_1)}(1+\varepsilon_n')-\frac{1}{y_1-x_1}(1+\varepsilon_n'')\right]\cdot$$

$$\cdot(1-y_1)(1+\delta_n'''')=\left[\frac{T_n'(y_1)}{T_n(y_1)}-\frac{1}{y_1-x_1}\right](1-y_1)(1+\delta_n''')+$$

$$+\varepsilon_n'''(1-y_1)(1+\delta_n'''')\left[\frac{T_n'(y_1)}{T_n(y_1)}+\frac{1}{y_1-x_1}\right]\quad (|\varepsilon_n'|,|\varepsilon_n''|,|\varepsilon_n'''|\to 0),$$

from where we obtain (4.23).

The remaining parts can be treated similarly. We omit the details.

4.5. PROOF OF STATEMENT 3.5. Let

$$\Omega(x)\overset{\text{def}}{=\!=}c_nP_n^{(\alpha,\alpha)}(x)\frac{x^2-1}{x^2-[x_1^{(\alpha,\alpha)}]^2}. \tag{4.28}$$

By (4.28) and (4.17)

$$\frac{\Omega''(1)}{\Omega'(1)}=2\left[\frac{P_n'(1)}{P_n(1)}-\frac{1}{1-x_1}+\frac{1}{2}-\frac{1}{1+x_1}\right]>-1/4\quad (n\geqq n_0)$$

because $(1+x_{1n})-2\to 0$ if $n\to\infty$. I.e.,

$$v_0(\Omega;-1)=1+2\frac{\Omega''(1)}{\Omega'(1)}\geqq\varrho>0\quad (n\geqq n_0).$$

By $v_0(\Omega;1)=1$ we obtain $v_0(\Omega;x)\geqq\varrho$. Similar argument holds for v_{n+1} (where $x_{n+1}\equiv 1$).

Now we investigate $v_k(\Omega;x)$ $(2\leqq k\leqq n-1)$. By Statement 3.3 and $P_n^{(\alpha,\alpha)}(x)=(-1)^nP_n^{(\alpha,\alpha)}(-x)$ the nodes $\{x_{1n},x_{2n},\dots,x_{n-1,n},-1\}$ $(n=1,2,\dots)$ form a ϱ-normal system if $-1\leqq\alpha<-2/3$. Now using the argument of Section 4.3 with $\omega(x)=(x+1)\prod_{k=1}^{n-1}(x-x_k)$ and $\Omega(x)$ defined by (4.28) we obtain the desired result.

4.6. PROOF OF STATEMENT 3.6. Let $\Omega_n(x)=(x-y_1)(x-x_1)^{-1}T_n(x)$ and $\Delta_n(x)=(x-y_n)(x-x_n)^{-1}\Omega_n(x)$.
By usual calculation

$$\frac{\Delta''(y_1)}{\Delta'(y_1)}=2\left[\frac{T_n'(y_1)}{T_n(y_1)}-\frac{1}{y_1-x_1}+\frac{1}{y_1-y_n}-\frac{1}{y_1+x_1}\right]. \tag{4.29}$$

Now, using again $(y_1+x_1)\to(y_1-y_n)\to 2$ $(n\to\infty)$, we get by (4.19), (4.29) and (4.22)

$$v_1(\Delta;x)=v_1(\Omega;x)-2\left(\frac{1}{y_1-y_n}-\frac{1}{y_1+x_1}\right)(x-y_1)\geqq\varrho/2 \tag{4.30}$$

if $0\leqq c\leqq 1.3672$ and $n\geqq n_0$. The estimation for $v_n(\Delta;x)$ is similar.

To consider $v_k(\Delta; -1)$ $(2 \leqq k \leqq n-1)$ we get, by repeatedly applying (4.7),

$$v_k(\Delta; 1) = v_k(T; 1) - \left[\frac{4x_k(1-x_k)(y_1^2-x_1^2)}{(y_1^2-x_k^2)(x_1^2-x_k^2)}\right]$$

and

$$v_k(\Delta; -1) = v_k(T; -1) - \left\{\frac{4x_k(1+x_k)(x_1^2-y_1^2)}{(y_1^2-x_k^2)(x_1^2-x_k^2)}\right\}.$$

First let $0 \leqq c < 1$. Then $v_k(\Delta; -1) \geqq v_k(T; -1)$, further for $v_k(\Delta; 1)$ we remark that

$$[\dots] \leqq 2\frac{(y_1-x_1)}{(y_1-x_k)(x_1-x_k)} \quad \text{if} \quad x_k \geqq 0,$$

which was treated in Section 4.4. If $x_k < 0$, then $v_k(\Delta; 1) = v_k(T; 1) + O(n^{-2})$. The case $1 < c < d$ is analogous to Section 4.4.

References

[1] G. Grünwald, On the theory of interpolation, *Acta Math.*, **75** (1942), 219—245.

[2] G. Freud, On Hermite—Fejér interpolation processes, *Studia Sci. Math. Hungar.*, **7** (1972), 307—310.

[3] P. Vértesi, Hermite—Fejér type interpolations. III., *Acta Math. Acad. Sci. Hungar.*, **34** (1979), 67—84.

[4] P. Erdős and P. Turán, On interpolation. II., *Ann. of Math.*, **39** (1938), 703—724.

[5] G. Szegő, *Orthogonal polynomials*, AMS Coll. Publ. Vol. XXIII. (New York, 1959).

(Received Decemer 5, 1977)

MATHEMATICAL INSTITUTE OF THE
HUNGARIAN ACADEMY OF SCIENCES
H—1053 BUDAPEST
REÁLTANODA U. 13—15.

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (3—4), (1979), 279—286.

ON A PROBLEM ABOUT DARBOUX POINTS

By

J. CEDER (Santa Barbara), M. LACZKOVICH and G. PETRUSKA (Budapest)

It is well known that a subset of the reals is a G_δ set if and only if it is the set of continuity points, $C(f)$, of a real function f (see [5]). ROSEN [6] showed that the set of Darboux points, $D(f)$, of a real function f is also a G_δ. Hence, if f is any real function then both $C(f)$ and $D(f)$ are G_δ sets with $C(f)\subset D(f)$.

In CEDER [2] the converse of this result was conjectured, namely: given two G_δ sets C and D with $C\subset D$ there exists a Baire 2 function f such that $C(f)=C$ and $D(f)=D$. This conjecture was settled affirmatively for a number of special cases in CEDER [2] and CEDER and PEARSON [3].

In this paper we validate the conjecture in the general case with the following

THEOREM. *Suppose C and D are G_δ subsets of a real interval I with $C\subset D$. Then, there exists a Baire 2 function h such that $C(h)=C$ and $D(h)=D$. Moreover, if C is dense in I then h can be taken to be an upper semi-continuous function.*

Notation and terminology. All functions considered in the sequel will have domain and range subsets of the real numbers. The set of right (left) *cluster values* of f at x is defined to be the set of all real numbers y such that there exists a sequence $\{x_n\}_{n=1}^{\infty}$ in $\operatorname{dom}(f)\cap(x,\infty)$ $\big(\operatorname{dom}(f)\cap(-\infty,x)\big)$ such that $\{x_n\}_{n=1}^{\infty}$ converges to x and the sequence $\{f(x_n)\}_{n=1}^{\infty}$ converges to y. The set of right (left) cluster values of f at x is denoted by $K^+(f,x)$ $\big($respectively $K^-(f,x)\big)$. Note that f does not have to be defined at x for $K^+(f,x)$ or $K^-(f,x)$ to be non-empty.

If I is a real interval then a function f from I into R is said to be Darboux if it has the intermediate value property. The intermediate value property can be localized as follows: a point $x\in\operatorname{dom}(f)$ is a *Darboux point* of f provided for each $\delta>0$ and right (resp. left) cluster value a, if b is strictly between a and $f(x)$ then there exists a $y\in(x,x+\delta)\cap\operatorname{dom}(f)$ (resp. $(x-\delta,x)\cap\operatorname{dom}(f)$) such that $f(y)=b$. It turns out that f is Darboux in I iff $D(f)=I$ (see CSÁSZÁR [4]). Clearly $C(f)\subset D(f)$ for all function f.

The proof of the main theorem consists of first proving it for the special case when C is dense (Lemma 2), then using that to prove the general case.

LEMMA 1. *Let C be a G_δ set dense in an interval I. Then there exists a bounded function g defined on C such that*

(1) *g is continuous on C,*

(2) $\limsup\limits_{t\to x+0} g(t)=\limsup\limits_{t\to x-0} g(t)>\liminf\limits_{t\to x+0} g(t)=\liminf\limits_{t\to x-0} g(t)$ *for* $x\notin C$,

(3) *g has the intermediate value property on C.*

PROOF. Since C is a G_δ dense in I we may put $I\setminus C=\bigcup_{k=1}^{\infty} F_k$ where $\{F_k\}_{k=1}^{\infty}$ is a sequence of disjoint nowhere dense closed sets. (For a proof see [7].) For each k let $\{(\alpha_n^k, \beta_n^k)\}_{n=1}^{\infty}$ be an enumeration of the family of components of $I\setminus F_k$. For each k define f_k on $I\setminus F_k$ by

$$f_k(x) = 3^{-k}\sin[(x-\alpha_n^k)^{-1}(\beta_n^k-x)^{-1}]$$

whenever $x\in(\alpha_n^k, \beta_n^k)$. Then define $g(x)=\sum_{k=1}^{\infty} f_k(x)$ if $x\in C$. The facts that g is bounded and continuous on C are obvious by the uniform convergence of $\sum_{k=1}^{\infty} f_k$. If $x\notin C$, then $x\in F_k$ for some k. Then $\sum_{j\neq k} f_j$ is continuous at x and statement (2) is true when g is replaced by f_k. Hence (2) is true for g.

Now we proceed to prove (3). Let c be any real number. Let us call a closed interval J with end points in C a c-interval if there exist points $a, b\in J$ for which $g(a)>c>g(b)$. To prove (3) we need to show each c-interval J_0 contains a point $x\in C$ such that $g(x)=c$.

By induction suppose we have found c-intervals $J_0, J_1, \dots, J_{k-1}$ such that $J_0\supset J_1\supset \dots \supset J_{k-1}$, $J_j\cap\left(\bigcup_{i=1}^{j} F_i\right)=\emptyset$ and $|J_j|\leqq\frac{1}{j+1}|I|$ for $j\leqq k-1$. From the construction the function $h=\sum_{i=1}^{k-1} f_i$ is continuous on J_{k-1} (when $k=1$ we put $h=0$). Since C is everywhere dense we can find points $x_1, \dots, x_{n-1}$ in C such that $x_0<x_1<\dots<x_n$, $J_{k-1}=[x_0, x_n]$, $x_i-x_{i-1}\leqq\frac{1}{k+1}|I|$ for $i\leqq n$ and the oscillation of h in $[x_{i-1}, x_i]$ satisfies

$$\omega(h, [x_{i-1}, x_i]) < 2^{-1}3^{-k} \quad \text{for} \quad i \leqq n.$$

If one of the values $g(x_i)$ for $i\leqq n-1$ equals c, we stop the construction and the proof is complete. Otherwise, at least one of the intervals, say $J^*=[x_{i-1}, x_i]$ is a c-interval again. Select $a, b\in J^*\cap C$ such that $g(a)>c>g(b)$. (We may assume that $a<b$ without loss of generality.) Let (α, β) and (α', β') denote the components of $I\setminus F_k$ containing a and b, respectively. If $(\alpha, \beta)=(\alpha', \beta')$, that is $[a, b]\subset I\setminus F_k$ then we put $J_k=[a, b]$. Otherwise the end points β and α' lie in (a, b).

Let $s(x)=\sum_{i\neq k} f_i(x)$ when $x\in C\cup F_k$.

Suppose $s(\beta)<c+3^{-k}$. Then $\liminf_{\substack{t\to\beta-0\\ t\in C}} f_k(t)=-3^{-k}$ and the continuity of s at $\beta\in F_k$ imply that

$$\liminf_{\substack{t\to\beta-0\\ t\in C}} g(t) < c.$$

Therefore, there exists $\varepsilon>0$ such that $[a, \beta-\varepsilon]$ is a c-interval. In this case we put $J_k=[a, \beta-\varepsilon]$.

On the other hand if $s(\beta) \geqq c+3^{-k}$ we have

$$|s(\beta)-s(\alpha')| \leqq \omega(s, [a, b]) \leqq \omega(s, J^*) \leqq \omega(h, J^*)+\omega\left(\sum_{i=k+1}^{\infty} f_i, J^*\right) \leqq$$

$$\leqq \frac{1}{2} 3^{-k}+2 \sup\left|\sum_{i=k+1}^{\infty} f_i\right| \leqq \frac{1}{2} 3^{-k}+2 \cdot 3^{-k-1} \frac{3}{2} = \frac{3}{2} 3^{-k}.$$

Hence, $s(\alpha') \geqq c-\frac{1}{2} 3^{-k} > c-3^{-k}$. Therefore, $\limsup_{\substack{t \to \alpha'+0 \\ t \in C}} f_k(t) = 3^{-k}$ and this together with the continuity of s at $\alpha' \in F_k$ imply that $\limsup_{\substack{t \to \alpha'+0 \\ t \in C}} g(t) > c$. Hence there exists $\varepsilon > 0$ such that $[\alpha'+\varepsilon, b]$ is a c-interval. Then put $J_k = [\alpha'+\varepsilon, b]$.

In any case, we have found a c-interval J_k of length $\leqq \frac{1}{k+1}|I|$ for which $J_k \subset J_{k-1} \setminus F_k$.

We continue this induction process and we stop it whenever we obtain x_i for which $g(x_i)=c$ as remarked above. If the process is never stopped, then we obtain a nested sequeuce of c-intervals $\{J_k\}_{k=1}^{\infty}$ for which there exists an $x_0 \in C$ such that $\bigcap_{k=1}^{\infty} J_k = \{x_0\}$ and $g(x_0)=c$. This completes the proof.

LEMMA 2. *Suppose C and D are G_δ sets dense in I with $C \subset D$. Then there exists a bounded upper semi-continuous function f such that $C(f)=C$, $D(f)=D$ and $K^-(f, x)=K^+(f, x)$ whenever $x \in D$.*

Moreover, for every $x_0 \notin D$, $\limsup_{\substack{x \to x_0 \\ x \in C}} f(x) < f(x_0)$ and for any $\varepsilon > 0$ there exists $\delta > 0$ such that all the irrational numbers $\lambda \in \left(\limsup_{\substack{x \to x_0 \\ x \in C}} f(x)+\varepsilon, f(x_0)\right)$ are omitted by f in $(x_0-\delta, x_0+\delta)$.

PROOF. Let g be a function specified by Lemma 1. Define F on I by $F(x) = \limsup_{t \to x} g(t)$. Let $I-D = \bigcup_{n=1}^{\infty} H_n$ where $\{H_n\}_{n=1}^{\infty}$ is a disjoint sequence of closed sets. Now we define f on I by

$$f(x) = \begin{cases} F(x) & \text{if } x \in D \\ \dfrac{[nF(x)]+1}{n} & \text{if } x \in H_n \end{cases}$$

where $[y]$ denotes the greatest integer in y. It is obvious that F is Darboux, upper semi-continuous, that $f(x)=g(x)$ whenever $x \in C$, and $C(F)=C$. Moreover, the definition implies

$$F(x) < f(x) \leqq F(x)+n^{-1} \quad \text{if} \quad x \in H_n. \tag{4}$$

Let us now prove that f, too, is upper semi-continuous. Let $x_0 \in D$, $\varepsilon > 0$ and N be an integer greater than $\frac{1}{\varepsilon}$. Then there exists $\delta > 0$ such that $(x_0-\delta, x_0+\delta) \cap$

$\cap \bigcup_{i=1}^{N} H_i = \emptyset$. Therefore,

$$\limsup_{x\to x_0} f(x) \leqq \limsup_{x\to x_0}\left(F(x)+\frac{1}{N}\right) = F(x_0)+\frac{1}{N} < f(x_0)+\varepsilon$$

so that f is u.s.c. at points of D. If $x_0 \notin D$ then $x_0 \in H_n$ for some n. Since the composition of an u.s.c. function with the greatest integer function is again u.s.c. it follows that $f|_{H_n}$ is u.s.c. Again, let $\varepsilon > 0$ and choose N so that $N > \frac{1}{\varepsilon}$. Then there exists $\delta > 0$ such that $(x_0-\delta, x_0+\delta) \cap \bigcup_{\substack{i\leqq N\\ i\neq n}} H_i = \emptyset$. Hence,

$$\limsup_{x\to x_0} f(x) = \max\left(\limsup_{x\to x_0,\, x\notin H_n} f(x), \limsup_{x\to x_0,\, x\in H_n} f(x)\right) \leqq$$

$$\leqq \max\left(F(x_0)+\frac{1}{N}, f(x_0)\right) < f(x_0)+\frac{1}{N} < f(x_0)+\varepsilon$$

so that f is u.s.c. at points of $I-D$ too.

From conditions (2) and (3) of Lemma 1 it follows that $K^-(g, x) = K^+(g, x)$ for all $x \in I$. From the continuity of g and the definition of F it follows that

(5) $$K^-(g, x) = K^+(g, x) = K^-(F, x) = K^+(F, x) \quad \text{for all} \quad x \in I.$$

Conditions (4) and (5) then imply that

(6) $$K^-(f, x) \cap K^+(f, x) \supset K^+(F, x) = K^-(F, x) \quad \text{for all} \quad x \in I, \quad \text{and}$$

$$K^-(f, x) = K^+(f, x) = K^+(F, x) = K^-(F, x) \quad \text{for} \quad x \in D.$$

Conditions (5) and (6) immediately imply that $C \subset C(f)$. If $x \in C(f)$, then $K^-(F, x) = K^+(F, x) = \{f(x)\}$ by (6) and hence, by (5), we have $K^+(g, x) = K^-(g, x) = \{f(x)\}$. By virtue of (2) of lemma 1 we must have $x \in C$. Thus $C = C(f)$.

Next we show that $D = D(f)$. The inclusion $D \subset D(f)$ follows from the facts that the cluster sets of f and g are identical on D (conditions (5) and (6)) and that g has the intermediate value property on C. It remains to show that $D(f) \subset D$. For this it suffices to show that $x_0 \in H_n$ implies $x_0 \notin D(f)$. Let $x_0 \notin D$ and $\varepsilon > 0$ be given. Then $\limsup_{\substack{x\to x_0\\ x\in C}} f(x) = \limsup_{\substack{x\to x_0\\ x\in C}} g(x) = F(x_0) < f(x_0)$. By the upper semi-continuity of F we can find $\delta > 0$ such that $0 < |x-x_0| < \delta$ implies

$$F(x) \leqq F(x_0)+\varepsilon = \limsup_{\substack{x\to x_0\\ x\in C}} f(x)+\varepsilon \stackrel{\text{def}}{=} \Lambda.$$

In particular $f(x) \leqq \Lambda$ for $x \in D \cap (x_0-\delta, x_0+\delta)$. By the construction of f, all the values taken on $I \setminus D$ are rational numbers hence if $\Lambda < \lambda < f(x_0)$ and λ is irrational then it cannot be a value of f on $(x_0-\delta, x_0+\delta)$. This obviously implies $x_0 \notin D(f)$ as well and the proof is complete.

LEMMA 3. *Let G be an everywhere dense G_δ subset of an open interval J. Suppose $f_1<f_2$ are continuous functions on J. Then there exists a Baire 2 function q defined on G such that*

(7) $$f_1(x) \leqq q(x) \leqq f_2(x) \quad \text{for all} \quad x\in G,$$

(8) *for each $x\in J$, $\delta>0$ and $\lambda\in(f_1(x), f_2(x))$ the inverse image $q^{-1}(\lambda)$ intersects each of the intervals $(x-\delta, x)$ and $(x, x+\delta)$.*

PROOF. Let $J_1, J_2, \dots$ be an enumeration of such subintervals of J with rational end-points for which $\sup_{x\in J_k} f_1(x) < \inf_{x\in J_k} f_2(x)$. By induction we can easily select a sequence $P_1, P_2, \dots$ of pairwise disjoint and nowhere dense perfect sets with $P_k\subset J_k\cap G$. (Referring to Baire's category theorem an everywhere dense G_δ subset cannot be countable and hence by [5], p. 355 it contains a perfect subset.)

Let h_k be a continuous function defined on P_k such that $h_k(P_k)=[\sup_{x\in J_k} f_1(x), \inf_{x\in J_k} f_2(x)]$ (Cantor-type function.) Now we define

$$q(x) = \begin{cases} h_k(x), & x\in P_k \\ f_2(x), & x\in G\setminus \bigcup_{j=1}^{\infty} P_j. \end{cases}$$

The reader can easily check (7) and (8) using routine arguments.

LEMMA 4. *Let D be a G_δ subset of an open interval $J=(a, b)$. Suppose f, β_1, β_2 are continuous functions on J with $0<\beta_1(x)\leqq 1$, $0<\beta_2(x)\leqq 1$ for all $x\in J$. Then there exists a Baire 2 function s on J with the following properties:*

(9) $$f(x)-\beta_1(x) \leqq s(x) \leqq f(x)+\beta_2(x) \quad (x\in J),$$

(10) *for any open subinterval $J'\subset J$, $s(J')\supset f(J')$ (in particular $s(J)\supset f(J)$),*

(11) $$C(s) = \emptyset,$$

(12) $$D(s) = D.$$

PROOF. We can find an ascending sequence of compact sets $\emptyset=K_0\subset K_1\subset\dots$ such that $\bigcup_{n=0}^{\infty} K_n=J\setminus D$. Next, pick H_n to be a countable dense subset of $K_n\setminus K_{n-1}$. Put $H=\bigcup_{n=1}^{\infty} H_n$, $G=J\setminus H$. Now we apply Lemma 3 with this dense G_δ set G and $f_1=f-\beta_1$, $f_2=f+\frac{1}{2}\beta_2$ to obtain a Baire 2 function q with (7) and (8) of Lemma 3. Define

$$s(x) = \begin{cases} q(x), & x\in G, \\ f(x)+\frac{1}{2}\beta_2(x)+\frac{1}{2n}\min\{\beta_2(y), y\in K_n\}, & x\in H_n. \end{cases}$$

s is obviously a Baire 2 function and assertion (9) is clear by (7) of Lemma 3. (8) of Lemma 3 and $f(x)\in\left(f(x)-\beta_1(x), f(x)+\frac{1}{2}\beta_2(x)\right)$ immediately imply (10) and (11).

If $x\in D$ and N is fixed then $(x-\delta, x+\delta)\cap K_n=\emptyset$ for $n\leqq N$ and $\delta>0$ small enough. Therefore $K^-(q, x)=K^-(s, x)=K^+(s, x)=K^+(q, x)=\left[f(x)-\beta_1(x), f(x)+\frac{1}{2}\beta_2(x)\right]$ hence by $s(x)=q(x)$, (8) of Lemma 3 implies the Darboux property. If $x\notin D$ then $x\in K_N\setminus K_{N-1}$. Take first the case $x\in H_N$. Then by the construction of s we have $s(x)>f(x)+\frac{1}{2}\beta_2(x)$ and $K^+(s, x)\supset K^+(q, x)=\left[f(x)-\beta_1(x), f(x)+\frac{1}{2}\beta_2(x)\right]$. Let $\lambda\in\left(f(x)+\frac{1}{2}\beta_2(x), s(x)\right)\setminus s(H)$ ($s(H)$ is countable). Then we get $\lambda\notin s((x-\delta, x+\delta))$ if δ is small enough and the Darboux property at x is violated. On the other hand if $x\in(K_N\setminus K_{N-1})\setminus H_N$ then $z=f(x)+\frac{1}{2}\beta_2(x)+ +\frac{1}{2N}\min\{\beta_2(y), y\in K_N\}\in K^-(s, x)\cup K^+(s, x)$ since H_N is everywhere dense in $K_N\setminus K_{N-1}$. Furthermore $z>f(x)+\frac{1}{2}\beta_2(x)\geqq q(x)=s(x)$ and selecting $\lambda\in\left(f(x)+ +\frac{1}{2}\beta_2(x), z\right)\setminus s(H)$ we get $x\notin D(s)$ as above. Hence $D(s)=D$ and the proof is complete.

PROOF OF THE THEOREM. Without loss of generality we may assume that I is a bounded interval. Let C and D be G_δ sets with $C\subset D\subset I$. Put $C^*=C\cup(I\setminus\bar{C})$ and $D^*=D\cup C^*$. Then C^* and D^* are dense G_δ sets with $C^*\subset D^*$ and we may apply Lemma 2 to obtain a bounded Baire 1 function f such that

(13) $$C(f) = C^* \quad\text{and}\quad D(f) = D^*,$$

(14) $$K^-(f, x) = K^+(f, x) \quad\text{if}\quad x\in D^*,$$

(15) $$\text{for every } x_0\notin D^* \text{ we have } \limsup_{\substack{x\to x_0\\ x\in C^*}} f(x) < f(x_0)$$

and for any given $\varepsilon>0$ there exists $\delta>0$ such that all the irrational numbers $\lambda\in\left(\limsup_{\substack{x\to x_0\\ x\in C^*}} f(x)+\varepsilon, f(x_0)\right)$ are omitted by f in $(x_0-\delta, x_0+\delta)$.

Let $\{J_n=(a_n, b_n);\ n\in\Gamma\}$ be an enumeration of the components of $I\setminus\bar{C}$ where Γ is an ordinal $\leqq\omega_0$. For each n we apply Lemma 4 with $J=J_n$, $\beta_1(x)=\beta_2(x)= =(x-a_n)(b_n-x)/n(b_n-a_n)^2=\eta_n(x)$, $f=f|_{J_n}$, $D=D\cap J_n$ and obtain a Baire 2 function s_n on J_n with (9)—(12) as stated in Lemma 4.

Now we are ready to define

$$h(x)=\begin{cases} f(x), & x\in\bar{C},\\ s_n(x), & x\in J_n \quad (n\in\Gamma).\end{cases}$$

h is obviously a Baire 2 function. We prove $C(h)=C$.

Let $x_0\in C$. We show that h is continuous from the right hand side at x_0. If $x_0=a_n$ then $|h(x)-f(x)|\leqq\eta_n(x)$ $(x\in J_n)$ and the continuity of f at x_0 assures $h(x)\to h(x_0)$ $(x\to x_0+0)$.

If $[x_0, x_0+\delta]\subset\overline{C}$ for some $\delta>0$ then the continuity is obvious by $h(x)=f(x)$ $(x\in\overline{C})$. Otherwise, if $\varepsilon>0$ is given we can find δ so small that $(x_0, x_0+\delta)\cap J_n=\emptyset$ for $n\leqq\frac{1}{\varepsilon}$. Hence $|h(x)-f(x)|<\varepsilon$ for every $x\in(x_0, x_0+\delta)$ and the continuity is immediate again.

Let $x_0\notin C$. If $x_0\notin\overline{C}$ then $x_0\notin C(h)$ by (11) of Lemma 4. If $x_0\in\overline{C}\setminus C$ then $x_0\notin C^*$ and hence $x_0\notin C(f)$. (10) of Lemma 4 implies $\omega(h, x_0)\geqq\omega(f, x_0)>0$ where $\omega(\cdot, x_0)$ denotes the oscillation at x_0. Thus we have $x_0\notin C(h)$ and $C(h)=C$.

Now we turn to prove $D(h)=D$.

$$[C\cup(I\setminus\overline{C})]\cap D = [C\cup(I\setminus\overline{C})]\cap D(h)$$

is obvious by the construction. Therefore $x_0\in\overline{C}$ is to be considered only.

For every $x_0\in\overline{C}$, $K^{\pm}(h, x_0)=K^{\pm}(f, x_0)$ easily follows from $\eta_n(x)\leqq\frac{1}{n}$ if $x_0\notin\{a_n, b_n;\ n\in\Gamma\}$ and from $\lim\limits_{x\to a_n+0}\eta_n(x)=\lim\limits_{x\to b_n-0}\eta_n(x)=0$ if $x_0=a_n$ or $x_0=b_n$. Let $x_0\in D\cap(\overline{C}\setminus C)$. Then $h(x_0)=f(x_0)$, $K^{\pm}(f, x_0)=K^{\pm}(h, x_0)$, $x_0\in D(f)$ and (10) of Lemma 4 imply Darboux property at x_0.

Finally consider $x_0\in[\overline{C}\setminus C]\setminus D$. Then obviously $x_0\notin D^*$. If $x_0=a_n$ then, by (15), $\lim\limits_{x\to a_n+0}\eta_n(x)=0$ implies $\limsup\limits_{x\to x_0+0}h(x)\leqq\limsup\limits_{x\to x_0+0}f(x)<f(x_0)=h(x_0)$ and hence $x_0\notin D(h)$ is clear. The case $x_0=b_n$ is similar.

If $x_0\notin\{a_n, b_n;\ n\in\Gamma\}$ then for

$$\varepsilon=\frac{1}{3}\Big(f(x_0)-\limsup_{\substack{x\to x_0\\ x\in C^*}}f(x)\Big)$$

we choose δ_1 according to (15). Let $\delta\leqq\delta_1$ be so small that $x\in(x_0-\delta,\ x_0+\delta)\cap C^*$ implies

$$f(x)<\limsup_{\substack{x\to x_0\\ x\in C^*}}f(x)+\varepsilon$$

furthermore $n\leqq\frac{1}{\varepsilon}$ implies $(x_0-\delta,\ x_0+\delta)\cap J_n=\emptyset$.

By

$$|f(x)-h(x)|\leqq\eta_n(x)\leqq\frac{1}{n}\quad\big(x\in(x_0-\delta, x_0+\delta)\cap J_n\big)$$

we have

$$h(x)<\limsup_{\substack{x\to x_0\\ x\in C^*}}f(x)+2\varepsilon\quad\text{if}\quad x\in(x_0-\delta, x_0+\delta)\cap C^*.$$

Hence if

$$\lambda\in\Big(\limsup_{\substack{x\to x_0\\ x\in C^*}}f(x)+2\varepsilon, f(x_0)\Big)$$

is an irrational number then

$$h^{-1}(\lambda)\cap(x_0-\delta, x_0+\delta)\subset\big[f^{-1}(\lambda)\cap\big((x_0-\delta, x_0+\delta)\setminus C^*\big)\big]\cup$$
$$\cup\big[h^{-1}(\lambda)\cap\big((x_0-\delta, x_0+\delta)\cap C^*\big)=\emptyset\cup\emptyset=\emptyset.$$

This means $x_0\notin D(h)$ and the proof is complete.

References

[1] J. CEDER, On maximally Borel resolvable spaces, *Rev. Roumaine Math. Pures et Appl.*, **11** (1966), 89—94.

[2] J. CEDER, On Darboux points of real functions, to appear in *Rev. Roumaine Math. Pures et Appl.*

[3] J. CEDER and T. PEARSON, More on Darboux Points of Real Functions, to appear in *Periodica Math. Hung.*

[4] Á. CSÁSZÁR, Sur la propriété de Darboux, *C. R., Premier Congrès des Mathématiciens Hongrois*, Akadémiai Kiadó (Budapest, 1952), 551—560.

[5] C. KURATOWSKI, *Topologie I* (Warszawa, 1958).

[6] H. ROSEN, Connectivity points and Darboux points of real functions, *Fund. Math.*, **89** (1975), 265—269.

[7] W. SIERPIŃSKI, Sur une propriété des ensembles F_σ linéaires, *Fund. Math.*, **14** (1929), 216—220.

(Received December 16, 1977; revised June 28, 1978)

UNIVERSITY OF CALIFORNIA
SANTA BARBARA, CALIFORNIA 93106
USA

AND

EÖTVÖS LORÁND UNIVERSITY
DEPARTMENT I OF ANALYSIS
1088 BUDAPEST, MÚZEUM KRT. 6—8.

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (3—4), (1979), 287—292.

ON REGULAR PRONORMAL SUBGROUPS OF SYMMETRIC GROUPS

By

P. P. PÁLFY (Budapest)

In this paper we shall investigate regular pronormal subgroups of symmetric groups. These subgroups are closely related to a combinatorial problem as pointed out by L. BABAI [2]. All the groups considered are finite. We need the following

DEFINITION. A subgroup H of G is said to be *pronormal* if for any $g \in G$, $g^{-1}Hg$ is conjugate to H in $\langle H, g^{-1}Hg \rangle$.

The other notion important for us was introduced by L. BABAI [2].

DEFINITION. G is a *CI-group* (abbreviation of Cayley isomorphism property) if between any two isomorphic relational structures on the group G as underlying set which admit all right translations as automorphisms, there exists an isomorphism which is at the same time an automorphism of G.

The following simple observation characterizes CI-groups:

PROPOSITION [2, Corollary 3.2]. *G is a CI-group if and only if the right regular representation of G is a pronormal subgroup of the symmetric group on the elements of G.*

The following groups have been found to be CI-groups: the cyclic groups of prime order, Z_p see [2, Theorem 2.3], as well as Z_4 and $Z_2 \times Z_2$ [2, Corollary 3.3]. Our aim is to extend this list, thus giving an affirmative answer to [2, Problem 3.4].

THEOREM 1. *If* (i) $p_1 < p_2 < \ldots < p_k$ $(k \geqq 0)$ *are prime numbers,* $n = \prod_{i=1}^{k} p_i$, (ii) $\varphi(n)$ *is prime to n and* (iii) $\prod_{i=1}^{l-1} p_i < p_l$ *for any l, $2 \leqq l \leqq k$, then the cyclic group of order n is a CI-group. (φ denotes Euler's φ function.)*

We do not know whether these conditions are necessary. Nevertheless, we are able to prove that some of them are:

THEOREM 2. *If G is a finite CI-group, then either G is a cyclic group of order n such that $\varphi(n)$ is prime to n, or the order of G is four.*

REMARK. Since the CI property involves the case of (directed) graphs, we obtain — as a corollary — positive answer to a problem of A. ÁDÁM [1], whenever n has the forms as in Theorem 1.

Acknowledgement. I am grateful to LÁSZLÓ BABAI for his useful advices.

1.

This section is devoted to proving Theorem 1. We shall proceed by induction on the number of prime factors of n. By the Proposition, we have to show that for any two n-cycles x, y in S_n, $\langle x\rangle$ and $\langle y\rangle$ are conjugate subgroups in $\langle x, y\rangle$.

In order to emphasize where the assumptions are used, the proof is divided into four lemmas. Throughout this section let $n=mp$, $x=(1\ 2\ \dots\ n)$, $z=(i\ i+m\ \dots$ $\dots\ i+(p-1)m)$, $P=\langle z_i \mid 1\leqq i\leqq m\rangle$ and y an arbitrary n-cycle. So $x^m=z_1z_2\dots z_m$, $x^{-1}z_ix=z_{i+1}$ (subscripts must be considered mod m).

LEMMA 1.1. *If $m<p$ and p is a prime number, then there exists a $y'\in S_n$ such that y' is conjugate to y in $\langle x, y\rangle$ and $y'^m\in P$.*

PROOF. Since $m<p$, P is the unique Sylow p-subgroup of S_n containing x^m. So by Sylow's theorem an appropriate conjugate of y^m by an element of $\langle x, y\rangle$ belongs to P.

LEMMA 1.2. *If Z_m is a CI-group and $y^m\in P$, then there exists a $y'\in S_n$ such that $\langle y'\rangle$ is conjugate to $\langle y\rangle$ in $\langle x, y\rangle$, $y'^{-1}\langle z_i\rangle y'=\langle z_{i+1}\rangle$ and $y'^m\in P$.*

PROOF. As x and y are n-cycles and $x^m, y^m\in P$, they normalize P and the effect of conjugation by both of them on the set of the subgroups $\langle z_i\rangle$ $(1\leqq i\leqq m)$ is an m-cycle. Since Z_m is a CI-group, this forces the existence of a y' with the required properties.

LEMMA 1.3. *If p is a prime number, m is prime to both p and $p-1$, $y^m=\prod_{i=1}^{m} z_i^{a_i}$ $(0<a_i<p)$ and $y^{-1}z_iy=z_{i+1}^{b_i}$ $(0<b_i<p)$, then there exists a $y'\in S_n$ such that $\langle y'\rangle$ is conjugate to $\langle y\rangle$ in $\langle x, y\rangle$, $y'^m=x^m$ and $y'^{-1}z_iy'=z_{i+1}$.*

PROOF. As m is prime to $p-1$ we can choose an r such that $mr\equiv -1 \bmod p-1$. Let

$$u=\prod_{j=1}^{m}\left((yx^{-1})^{r(m-j)}x\right).$$

As one can check immediately

$$u^{-1}z_iu=z_i^{\prod_{j=1}^{m} b_{i+j-1}^{r(m-j)}}.$$

For brevity let $c_i=\prod_{j=1}^{m} b_{i+j-1}^{r(m-j)}$, so $u^{-1}z_iu=z_i^{c_i}$.

Since $y^m=y^{-1}y^my=\prod_{i=1}^{m} z_{i+1}^{a_ib_i}$, thus $a_{i+1}\equiv a_ib_i \bmod p$. Now

$$b_ic_i\equiv b_ic_i\left(\prod_{l=1}^{m} b_l\right)^r=b_i\prod_{j=1}^{m} b_{i+j-1}^{r(m-j+1)}=b_i^{1+rm}c_{i+1}\equiv c_{i+1} \bmod p.$$

Thus there exists an s for which $c_i\equiv sa_i \bmod p$ and we can suppose also that $s\equiv 1 \bmod m$, because m is prime to p. So $u^{-1}z_iu=z_i^{sa_i}$. Set $y'=uy^su^{-1}$, then

$$y'^m=uy^{ms}u^{-1}=\prod_{i=1}^{m} uz_i^{sa_i}u^{-1}=\prod_{i=1}^{m} uz_i^{c_i}u^{-1}=\prod_{i=1}^{m} z_i=x^m$$

furthermore

$$y'^{-1}z_iy' = uy^{-s}u^{-1}z_iuy^su^{-1} = uy^{-s}z_i^{c_i}y^su^{-1} =$$

$$= uz_{i+s}^{c_i\prod_{j=1}^{s}b_{i+j-1}}u^{-1} = uz_{i+s}^{c_{i+s}}u^{-1} = z_{i+s} = z_{i+1}.$$

This completes the proof.

LEMMA 1.4. *If m is prime to p and $y^m=x^m$, $y^{-1}z_iy=z_{i+1}$ then y is conjugate to x in $\langle x, y\rangle$.*

PROOF. We can choose a natural number t such that $mt\equiv -1 \bmod p$, since m is prime to p. Set

$$w = \prod_{j=1}^{m}\left((yx^{-1})^{t(m-j)}x\right).$$

We shall prove that $y=w^{-1}xw$.

We have supposed, that $(l)x=l+1$ (of course mod n but we do not mark it), so

$$(l+m)y = (l)x^m y = (l)y^{m+1} = (l)yx^m = (l)y+m$$

and

$$(l)yz_{l+1} = (l)z_ly = (l+m)y = (l)y+m,$$

hence $(l)y\equiv l+1 \bmod m$, and $(l)y-l-1$ depends only on the residue class of $l \bmod m$. Therefore we are justified to write $(l)y=l+1+md_l$. Now

$$l+m = (l)x^m = (l)y^m = l+m+m\sum_{j=1}^{m}d_j,$$

so $m\sum_{j=1}^{m}d_j=0$ (more precisely $\equiv 0 \bmod n$).

Let us compute now $(l)w$.

$$(l)w = (l)\prod_{j=1}^{m}\left((yx^{-1})^{t(m-j)}x\right) = l+m+mt\sum_{j=1}^{m}(m-j)d_{l+j-1}.$$

From this it follows that

$$(l)wy = \left(l+m+mt\sum_{j=1}^{m}(m-j)d_{l+j-1}\right)y = l+m+1+md_l+mt\sum_{j=1}^{m}(m-j)d_{l+j-1} =$$

$$= l+m+1+md_l+mt\sum_{j=1}^{m}(m-j+1)d_{l+j-1} =$$

$$= l+1+m+mt\sum_{i=1}^{m}(m-i)d_{l+i}+md_l(1+mt) = (l+1)w = (l)xw.$$

That is $wy=xw$, so we have $y=w^{-1}xw$.

Now we are able to prove Theorem 1.

We proceed by induction on the number of prime factors. Z_1 is obviously a CI-group. Let n be as in our theorem, $m=\prod_{i=1}^{k-1} p_i$ and $p=p_k$. So p is a prime number, $p>m$ by (iii), m is prime to $p-1$ by (ii) and by the induction hypothesis Z_m is a CI-group. Notice that if $\langle y'\rangle$ is conjugate to $\langle y\rangle$ in $\langle x, y\rangle$, then $\langle x, y'\rangle\subseteq\subseteq\langle x, y\rangle$. So the successive application of the lemmas yields the result.

2.

We shall denote by G_R the right regular representation of a group G. The right translation by $g\in G$ will be denoted by g_R. S_G stands for the symmetric group acting on G.

LEMMA 2.1. *If G is a CI-group and H is a subgroup of G, then H is also a CI-group.*

PROOF. Let $G=\bigcup_{t\in T} Ht$ be the right coset decomposition of G by H. So each $g\in G$ has a unique representation in the form $g=h(g)t(g)$, where $h(g)\in H$, $t(g)\in T$. Let $f\in S_H$ be arbitrary, and define $f^*\in S_G$ in the following way: for $g\in G$ let $gf^*=h(g)ft(g)_R$.

Since G is a CI-group there exists a $d_0\in\langle f^{*-1}G_Rf^*, G_R\rangle$ such that $f^{*-1}G_Rf^*==d_0^{-1}G_Rd_0$. Let $d^*=(1d_0^{-1})_Rd_0\in\langle f^{*-1}G_Rf^*, G_R\rangle$, it satisfies $d^{*-1}G_Rd^*==d_0^{-1}G_Rd_0=f^{*-1}G_Rf^*$ and $1d^*=1(1d_0^{-1})_Rd_0=1d_0^{-1}d_0=1$.

The right cosets by H form a system of blocks of imprimitivity for G_R and for f^*, and thus for d^* too. On the other hand $1d^*=1$, therefore $Hd^*=H$. Let us denote by d the restriction of d^* to H. The subgroup of $d^{*-1}G_Rd^*=f^{*-1}G_Rf^*$ which leaves H invariant is $d^{*-1}H_Rd^*=f^{*-1}H_Rf^*$. Restricting this equation to H we obtain $d^{-1}H_Rd=f^{-1}H_Rf$. (We do not distinguish $H_R\leqq S_H$ and $H_R\leqq\leqq G_R\leqq S_G$ in notation.)

As we know $d^*\in\langle f^{*-1}G_Rf^*, G_R\rangle$, therefore it has the form

$$d^* = g_{1R}f^{*-1}g_{2R}f^*g_{3R}f^{*-1}g_{4R}f^*\dots.$$

Then for any element $a\in H$,

$$ad^* = ah(g_1)_Rf^{-1}\big(h(g_1)^{-1}h(g_1g_2)\big)_Rf\big(h(g_1g_2)^{-1}\times$$

$$\times h(g_1g_2g_3)\big)_Rf^{-1}\big(h(g_1g_2g_3)^{-1}h(g_1g_2g_3g_4)\big)_Rf\dots.$$

This proves that $d\in\langle f^{-1}H_Rf, H_R\rangle$, so we are done.

LEMMA 2.2. *If G is not commutative, then G is not a CI-group.*

PROOF. Let $gf=g^{-1}$ for each $g\in G$, so $f\in S_G$. For $h\in G$ we have $gf^{-1}h_Rf==(g^{-1}h)^{-1}=h^{-1}g$, that is $f^{-1}h_Rf=h_L^{-1}$ and so $f^{-1}G_Rf=G_L$ the left regular representation of G. It is well-known, that G_L centralizes G_R, thus $G_R\triangleleft\triangleleft\langle G_R, f^{-1}G_Rf\rangle$. On the other hand the non-commutativity of G implies $G_R\neq G_L$ so G is not a CI-group.

LEMMA 2.3. *If $n=pq$, p and q are prime numbers and q divides $p-1$, then Z_n is not a CI-group.*

PROOF. Let $G=\langle a, b|a^p=b^q=1,\ ab=ba\rangle\cong Z_n$. Since $q|p-1$, we can choose an r such that $r\not\equiv 1$ but $r^q\equiv 1 \bmod p$. Let us define f as follows: $(a^k b^l)f=a^{kr^l} b^l$. Now

$$(a^k b^l)f^{-1}a_R f = (a^{kr^{q-l}} b^l)a_R f = (a^{kr^{q-l}+1} b^l)f = a^{k+r^l} b^l$$

and similarly $(a^k b^l)f^{-1}b_R f=a^{kr} b^{l+1}$. Easy computations show that

$$(f^{-1}a_R f)^{-1}a_R(f^{-1}a_R f) = a_R \quad \text{and} \quad (f^{-1}b_R f)^{-1}a_R(f^{-1}b_R f) = a_R^r$$

so $\langle a_R\rangle \lhd \langle G_R, f^{-1}G_R f\rangle$, but $\langle a_R\rangle \neq f^{-1}\langle a_R\rangle f$ are the unique Sylow p-subgroups of G_R and $f^{-1}G_R f$, respectively. Thus G is not a CI-group.

LEMMA 2.4. *If p is an odd prime, then Z_{p^2} is not a CI-group.*

PROOF. Let $G=\langle a\,|\,a^{p^2}=1\rangle\cong Z_{p^2}$ and $a^k b=a^{k(p+1)+1}$. Clearly $b\in S_G$. If $1b^m=1$ then $a^0=a^0 b^m=a^{\sum_{j=0}^{m-1}(p+1)^j}$, thus

$$0\equiv \sum_{j=0}^{m-1}(p+1)^j \equiv \sum_{j=0}^{m-1}(jp+1) = \frac{m(m-1)}{2}p+m \bmod p^2$$

hence $m\equiv 0 \bmod p$, and since p is odd $\frac{m(m-1)}{2}p\equiv 0 \bmod p^2$, so $m\equiv 0 \bmod p^2$. We have obtained that b is a p^2-cycle, so it is conjugate to a_R in S_G. Now $b^{-1}a_R b=a_R^{p+1}$, therefore $a_R\lhd\langle a_R, b\rangle$. Obviously $\langle b\rangle\neq\langle a_R\rangle$, thus G is not a CI-group.

LEMMA 2.5. *If p is an odd prime, then $Z_p\times Z_p$ is not a CI-group.*

PROOF. Let $G=\langle a, b\,|\,a^p=b^p=1,\ ab=ba\rangle\cong Z_p\times Z_p$, and let $(a^k b^l)f=a^{k+\frac{l(l-1)}{2}} b^l$, so $f\in S_G$. It is easy to check that $f^{-1}a_R f=a_R$, and for $c=f^{-1}b_R f$ we have $(a^k b^l)c=a^{k+l}b^{l+1}$. Now $c^{-1}a_R c=a_R$ and $c^{-1}b_R c=a_R b_R$, therefore $G_R\lhd\langle G_R, f^{-1}G_R f\rangle$, but $f^{-1}G_R f\neq G_R$ proving that G is not a CI-group.

LEMMA 2.6. *$Z_4\times Z_2$ is not a CI-group.*

PROOF. Let $G=\langle a, b\,|\,a^4=b^2=1,\ ab=ba\rangle\cong Z_4\times Z_2$ and let $(a^k b^l)f=a^{k(-1)^l} b^l$. Now for $c=f^{-1}a_R f$ and $d=f^{-1}b_R f$ we get $(a^k b^l)c=a^{k+(-1)^l} b^l$, $(a^k b^l)d=a^{-k}b^{l+1}$. Moreover $c^{-1}a_R c=a_R$, $c^{-1}b_R c=(a^2 b)_R$, $d^{-1}a_R d=a_R^{-1}$, $d^{-1}b_R d=b_R$, that is $\langle c, d\rangle\neq\langle a_R, b_R\rangle\lhd\langle a, b, c, d\rangle=\langle G_R, f^{-1}G_R f\rangle$. (Compare with the proof of Lemma 2.3.)

LEMMA 2.7. *Z_8 is not a CI-group.*

PROOF. Let $G=\langle a\,|\,a^8=1\rangle\cong Z_8$ and let $b\in S_G$: $a^k b=a^{5k+1}$. Similarly to the proof of Lemma 2.4 we can show that b is an 8-cycle and $b^{-1}a_R b=a_R^5$ and conclude that G is not a CI-group.

LEMMA 2.8. *$Z_2\times Z_2\times Z_2$ is not a CI-group.*

PROOF. Let $G=\langle a, b, c\,|\,a^2=b^2=c^2=1,\ ab=ba,\ ac=ca,\ bc=cb\rangle\cong Z_2\times Z_2\times Z_2$, and $(a^k b^l c^m)f=a^{k+lm} b^l c^m$, $f\in S_G$. Now $f^{-1}a_R f=a_R$ and for $d=f^{-1}b_R f$, $e=f^{-1}c_R f$ $(a^k b^l c^m)d=a^{k+m} b^{l+1} c^m$, $(a^k b^l c^m)e=a^{k+l} b^l c^{m+1}$ hold. Moreover $d^{-1}a_R d=$

$=a_R$, $d^{-1}b_Rd=b_R$, $d^{-1}c_Rd=(ac)_R$, $e^{-1}a_Re=a_R$, $e^{-1}b_Re=a_Rb_R$, $e^{-1}c_Re=c_R$. Thus $G_R\lhd\langle G_R, f^{-1}G_Rf\rangle$ but $G_R\neq f^{-1}G_Rf$. (Compare with the proof of Lemma 2.5.)

PROOF OF THEOREM 2. Let G be a finite CI-group. Then by Lemma 2.2 G is commutative. If the order of G is even, then Lemmas 2.1 and 2.3 imply that G is a 2-group, and moreover by Lemmas 2.1, 2.6, 2.7 and 2.8 the order of G is at most 4. If n, the order of G, is odd, then by Lemmas 2.1, 2.4 and 2.5 n is square-free. Finally Lemmas 2.1 and 2.3 imply that $\varphi(n)$ is prime to n.

References

[1] A. ÁDÁM, Research Problem 2—10, *Journal of Comb. Theory*, **2** (1967), 393.
[2] L. BABAI, Isomorphism problem for a class of point-symmetric structures, *Acta Math. Acad. Sci. Hungar.*, **29** (1977), 329—336.

(Received December 21, 1977)

EÖTVÖS LORÁND UNIVESITY
DEPARTMENT OF ALGEBRA AND NUMBER THEORY
1088 BUDAPEST, MÚZEUM KRT. 6—8

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (3—4), (1979), 293—296.

AN APPROXIMATION TO THE FOURIER TRANSFORM

By

A. McD. MERCER (Guelph)

§ 1. Let $\mathscr{F}(f, x)$ be the Fourier transform of f, namely

$$\mathscr{F}(f, x) = \int_0^\infty f(t) e^{-ixt}\, dt.$$

In [1] C. BALÁZS has given a formula which approximates $\mathscr{F}(f, x)$ when $f(t)$ is known at $n+1$ equidistant points and is known to tend to zero as $t \to +\infty$ at least as fast as some exponential function. Her theorem is

THEOREM A. *Let $f \in C[0, +\infty)$ and writing $\varphi(t) = f(t)e^{ct}$, suppose $\lim\limits_{t\to+\infty} \varphi(t)$ exists for some $c > 0$. Then as $n \to +\infty$*

$$\mathscr{F}(f, x) = \sum_{k=0}^{n} \varphi\left(\frac{k}{a_n}\right) \frac{a_n^k}{(c+a_n+ix)^{k+1}} + O\left\{\omega_\varphi\left(\sqrt{\frac{\log n}{n}}\right)\right\}.$$

Here $\omega_\varphi(\delta)$ is the modulus of continuity of φ, $a_n = \frac{c}{2}\frac{\log n}{n}$ and the constant of the O notation depends on c and $\sup\limits_{[0,+\infty)} |\varphi(t)|$.

Now it is not necessary for the definition of $\mathscr{F}(f, x)$ that such a constant $c > 0$ should exist and our first object is to examine the approximation of $\mathscr{F}(f, x)$ when a weaker condition at ∞ is imposed. The result of this is Theorem 1 below. Once this theorem is obtained we shall see that its proof can be modified to give a result for the case in which we *do* have the existence of $\lim\limits_{t\to+\infty} f(t)e^{ct}$ for some $c > 0$. This result appears as Theorem 2.

It is the behaviour of $f(t)$ as $t \to +\infty$ rather than its smoothness properties which seems to be the interesting feature. So to fix ideas we shall assume throughout that $f \in \text{Lip}\,[0, +\infty)$. This of course corresponds to the case $\omega_\varphi(\delta) = K\delta$ of Theorem A. We shall prove the following two theorems.

THEOREM 1. *Let $f \in \text{Lip}\,[0, +\infty)$ and $f \in L(0, +\infty)$. Writing $F(t) = \int_t^\infty |f(u)|\,du$ suppose that $t^\varepsilon F(t) = O(1)$ as $t \to +\infty$ for some $\varepsilon > 0$. Then as $n \to +\infty$*

$$\mathscr{F}(f, x) = \sum_{k=0}^{n} f\left(\frac{k}{a_n}\right) \frac{a_n^k}{(a_n+ix)^{k+1}} + O\{n^{-\frac{\varepsilon}{4+2\varepsilon}}\}.$$

Here $a_n = n^{\frac{3+2\varepsilon}{4+2\varepsilon}}$ and the constant of the O notation depends on $F(0)$ and the Lipschitz constant.

THEOREM 2. *Let $f \in \mathrm{Lip}\,[0, +\infty)$ and let $\lim_{t\to+\infty} f(t)e^{ct}$ exist for some $c>0$. Let $\{b_n\}$ be any strictly increasing sequence of positive numbers such that $b_n \to +\infty$, $b_n^2 n^{-1/2} \to 0$ and $b_n^{2+\varepsilon} n^{-1/2} \to \infty$ for some $\varepsilon > 0$. Then as $n \to +\infty$*

$$\mathcal{F}(f, x) = \sum_{k=0}^{n} f\left(\frac{k}{a_n}\right) \frac{a_n^k}{(a_n+ix)^{k+1}} + O\left\{\frac{b_n^2}{\sqrt{n}}\right\}.$$

Here $a_n = \dfrac{n}{b_n}$ and the constant of the O notation depends on $F(0)$ and the Lipschitz constant but not on c.

§ 2. Our analysis proceeds along lines generally similar to those in [1] but due to the absence of the convergence factor e^{-ct} in the integrals different inequalities have to be found. First we prove the following result

LEMMA.
$$J_n \equiv \int_0^\infty \sum_{k=0}^{n} |u-k| \frac{u^k}{k!} e^{-u}\, du = O(n^{3/2})$$
as $n \to +\infty$.

PROOF. For $k \geqq 0$ we have

$$\int_0^\infty |u-k| \frac{u^k}{k!} e^{-u}\, du = \left\{\int_k^\infty - \int_0^k\right\} \frac{u^{k+1}}{k!} e^{-u}\, du + \left\{\int_0^k - \int_k^\infty\right\} \frac{ku^k}{k!} e^{-u}\, du.$$

We integrate the first two integrals here by parts and the right hand side becomes

$$\left\{\int_k^\infty - \int_0^k\right\} \frac{u^k}{k!} e^{-u}\, du + 2\frac{k^{k+1}}{k!} e^{-k} = 2\int_k^\infty \frac{u^k}{k!} e^{-u}\, du + 2\frac{k^{k+1}}{k!} e^{-k} - 1 \leqq 1 + 2\frac{k^{k+1}}{k!} e^{-k}.$$

Summing over $0 \leqq k \leqq n$ we get

$$J_n \leqq n+1+2\sum_{k=1}^{n} \frac{k^{k+1}}{k!} e^{-k}.$$

If we write $a_k = \dfrac{k^{k+1}}{k!} e^{-k}$ we see that

$$\frac{a_{k-1}}{a_k} = \left(1-\frac{1}{k}\right)^k e < 1$$

and so $\{a_k\}$ is non-decreasing. Hence

$$J_n \leqq n+1+2n\frac{n^{n+1}}{n!} e^{-n}.$$

Since $n! \sim \sqrt{2\pi}\, n^{n+1/2} e^{-n}$ the last expression is seen to be $O(n^{3/2})$ and the lemma is proved.

We now turn to the proof of Theorem 1. It should be noted that a_n is not chosen till the end of the proof. All that is assumed of a_n before the final stage is that $\dfrac{a_n}{n} \to 0$.

PROOF OF THEOREM 1. We write

$$I_n(f, x) = \sum_{k=0}^{n} f\left(\frac{k}{a_n}\right) \frac{a_n^k}{(a_n+ix)^{k+1}} = \int_0^\infty e^{-a_n t} \sum_{k=0}^{n} f\left(\frac{k}{a_n}\right) \frac{(a_n t)^k}{k!} e^{-ixt}\, dt.$$

Then

$$\mathscr{F}(f,x)-I_n(f,x)=\int_0^\infty\left\{f(t)-e^{-a_nt}\sum_{k=0}^n f\left(\frac{k}{a_n}\right)\frac{(a_nt)^k}{k!}\right\}e^{-ixt}\,dt=$$

$$=\int_0^\infty e^{-a_nt}\left\{\sum_{k=0}^n\left[f(t)-f\left(\frac{k}{a_n}\right)\right]\frac{(a_nt)^k}{k!}+f(t)\sum_{k=n+1}^{n}\frac{(a_nt)^k}{k!}\right\}e^{-ixt}\,dt=A_n+B_n\text{ (say)}.$$

Consider A_n. Since $f\in\mathrm{Lip}\,[0,+\infty)$ then for some constant K_1 we have

(2.1)
$$|A_n|\leqq K_1\int_0^\infty\sum_{k=0}^n\left|t-\frac{k}{a_n}\right|\frac{(a_nt)^k}{k!}e^{-a_nt}\,dt=$$

$$=\frac{K_1}{a_n^2}\int_0^\infty\sum_{k=0}^n|u-k|\frac{u^k}{k!}e^{-u}\,du=O\left(\frac{n^{3/2}}{a_n^2}\right)$$

by the lemma. Next we consider the integral B_n. Clearly

$$|B_n|\leqq\int_0^\infty|f(t)|e^{-a_nt}\left\{\sum_{k=n+1}^{n}\frac{(a_nt)^k}{k!}\right\}dt.$$

If we write $R_n(a_nt)$ for the expression in brackets it is a simple matter to verify that

$$\frac{d}{dx}\{e^{-x}R_n(x)\}=e^{-x}\frac{x^n}{n!}.$$

So, integrating by parts, we get

$$|B_n|\leqq\left[-F(t)e^{-a_nt}\sum_{k=n+1}^\infty\frac{(a_nt)^k}{k!}\right]_0^\infty+a_n\int_0^\infty F(t)e^{-a_nt}\frac{(a_nt)^n}{n!}\,dt$$

where $F(t)=\int_t^\infty|f(u)|\,du$. The square bracket vanishes at $t=0$ and since $\lim_{t\to+\infty}F(t)=0$ and

$$0\leqq e^{-a_nt}\sum_{k=n+1}^\infty\frac{(a_nt)^k}{k!}\leqq 1$$

it vanishes at $t=+\infty$ also. Hence

$$|B_n|\leqq a_n\int_0^\infty F(t)e^{-a_nt}\frac{(a_nt)^n}{n!}\,dt.$$

Since $t^\varepsilon F(t)=O(1)$ at $t\to+\infty$ then we will have $t^\varepsilon F(t)\leqq K_2$ for all sufficiently large t, say $t\geqq M$. We write this last integral as

$$a_n\left\{\int_M^\infty+\int_0^M\right\}F(t)e^{-a_nt}\frac{(a_nt)^n}{n!}\,dt=C_n+D_n\text{ (say)}.$$

Consider C_n. In it we replace $F(t)$ by $K_2t^{-\varepsilon}$ and then put $u=a_nt$. We get

$$C_n\leqq K_2a_n^\varepsilon\int_{a_nM}^\infty e^{-u}\frac{u^{n-\varepsilon}}{n!}\,du\leqq K_2a_n^\varepsilon\int_0^\infty e^{-u}\frac{u^{n-\varepsilon}}{n!}\,du.$$

Hence

(2.2) $$C_n \leqq K_2 a_n^{\varepsilon} \frac{\Gamma(n-\varepsilon+1)}{n!} \sim K_2 \left(\frac{a_n}{n}\right)^{\varepsilon}.$$

Finally we turn to D_n. The function $e^{-a_n t}\frac{(a_n t)^n}{n!}$ has its maximum value in $[0, +\infty)$ when $t=\frac{n}{a_n}$. Since $\frac{a_n}{n}\to 0$ as $n\to+\infty$ we may suppose that n is so large that $\frac{n}{a_n}>M$. Then for such n we will have

$$D_n \leqq a_n M F(0) e^{-a_n M} \frac{(a_n M)^n}{n!}.$$

Again using the asymptotic formula for $n!$ we see that this term is $O\left\{\left(\frac{a_n}{n}\right)^{\gamma}\right\}$ for *all* $\gamma>0$. In particular we have

(2.3) $$D_n = O\left\{\left(\frac{a_n}{n}\right)^{\varepsilon}\right\}.$$

Collecting the results (2.1), (2.2), (2.3) we have

(2.4) $$\mathscr{F}(f, x) - I_n(f, x) = O\left\{\frac{n^{3/2}}{a_n^2}\right\} + O\left\{\left(\frac{a_n}{n}\right)^{\varepsilon}\right\}.$$

We now choose a particular value for a_n. We choose it to make these orders the same and we find that $a_n = n^{\frac{3+2\varepsilon}{4+2\varepsilon}}$ when each order term is $O\{n^{-\varepsilon/(4+2\varepsilon)}\}$ and this completes the proof of Theorem 1.

§ 3. We conclude by showing how the above proof can be modified to give the proof of Theorem 2. In the case of Theorem 2 the hypotheses of Theorem 1 are satisfied for *all* $\varepsilon>0$. Let the sequence $\{b_n\}$ be as stated in Theorem 2. With $a_n=\frac{n}{b_n}$ the proof goes through as before but now in (2.4) the right hand side is simply $O\left\{\frac{n^{3/2}}{a_n^2}\right\}$, that is $O\left\{\frac{b_n^2}{\sqrt{n}}\right\}$. This gives the result of Theorem 2 with $a_n=\frac{n}{b_n}$. The important feature of this theorem is that the numbers a_n in the expression for $I_n(f, x)$ can be chosen once and for all as they do not depend on c as in Balázs case nor on ε as in Theorem 1.

Reference

[1] C. BALÁZS, Approximative representation of Fourier transforms, *Acta. Math. Acad. Sci. Hungar.*, **28** (1976), 153—155.

(Received January 2, 1978)

UNIVERSITY OF GUELPH
DEPARTMENT OF MATHEMATICS AND STATISTICS
GUELPH, ONTARIO
CANADA N1G 2W1

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (3—4), (1979), 297—305.

A PSEUDONORM FOR UNBOUNDED TRANSFORMATIONS

By
I. Z. RUZSA (Budapest)

Concepts and notations

H, K, L, H_1, H' ...	Hilbert spaces
dim H	usual (geometrical) dimension
Dim H	algebraical dimension
transformation	linear transformation
manifold	subset, closed under the linear operations
subspace	(topologically) closed manifold
ON	orthonormal
basis	ON complete system

1. Introduction

Problem 39 of HALMOS [1] consists in constructing an unbounded linear transformation from a Hilbert space into another which is bounded on an ON basis. The problem is not a too difficult one; our aim is to show the more surprising fact that every unbounded transformation is a solution. More exactly,

THEOREM 1. *Let H, K be Hilbert spaces and $A\colon H\to K$ a transformation. Then there exists an ON basis E in H such that* $\sup_{e\in E}\|Ae\|<\infty$.

The quantity

$$\psi(A)=\inf_E\big(\sup_{e\in E}\|Ae\|\big),$$

where E runs over all the ON bases, is a kind of pseudonorm, which has some paradoxical properties. For example, in finite dimension it is subadditive, in countable dimension it is supermultiplicative, and in continuum dimension it is neither.

PROBLEM. Given a transformation A, does there exist necessarily an ON basis E such that A is bounded on the linear hull of E, that is, on the set of all the finite linear combinations

$$\sum_{j=1}^{n} a_j e_j,\ a_j\in \underline{R},\ e_j\in E?$$

2. The existence of $\psi(A)$

In this section we prove Theorem 1. Define the *conorm* of a transformation $A;\ H\to K$ by

$$\operatorname{cn} A = \inf_{\|e\|=1} \|Ae\|.$$

Obviously $\operatorname{cn} A = \|A^{-1}\|^{-1}$ if A has a bounded inverse (whose domain need not be the whole space K) and 0 otherwise. Now let E_1 be a maximal ON system in H such that $\|Ae\| \leqq 1$ for all $e \in E_1$. (Such an E_1 exists by Zorn's lemma.) Let $H_1 = E_1^\perp$. Evidently we have $\operatorname{cn} A|_{H_1} \geqq 1$.

Having defined E_n and H_n, let E_{n+1} be a maximal ON system in H_n satisfying $\|Ae\| \leqq n+1$ for all $e \in E_{n+1}$ and let $H_{n+1} = H_n \cap E_{n+1}^\perp$. If this sequence breaks off (that is $H_n = 0$ for some n), we are ready, $E = \cup E_n$ is the desired basis. Now we prove that it must break off.

(2.1) LEMMA. *Let H, K be Hilbert-spaces, $0 \neq H_k \subset H$ a decreasing sequence of subspaces and $A\colon H \to K$ a transformation. Then*

$$\lim_{k\to\infty} \operatorname{cn} A|_{H_k} < \infty.$$

The limit above exists, since

$$H' \supset H'' \Rightarrow \operatorname{cn} A|_{H'} \leqq \operatorname{cn} A|_{H''}.$$

PROOF. Suppose the contrary. Let $e_n \in H_n$ be a unit vector and consider the series $x = \sum_{j=1}^{\infty} 3^{-j} e_{n_j}$, $n_1 < n_2 < \ldots$. This series is convergent for an arbitrary sequence $\{n_j\}$; we shall define n_j recursively.

Writing $x_k = \sum_{j=k}^{\infty} 3^{-j} e_{n_j}$ we have

$$Ax = \sum_{j=1}^{k-1} 3^{-j} A e_{n_j} + A x_k. \tag{2.2}$$

Since $\|e_{n_j}\| = 1$, we get

$$\|x_k\| \geqq 3^{-k} - \sum_{j=k+1}^{\infty} 3^{-j} = \frac{1}{2} 3^{-k},$$

and $x_k \in H_{n_k}$ implies

$$\|Ax_k\| \geqq \frac{1}{2} 3^{-k} \operatorname{cn} A|_{H_{n_k}}. \tag{2.3}$$

Now let $n_1 = 1$, and given $n_1, \ldots n_{k-1}$, choose n_k so large that

$$\operatorname{cn} A|_{H_{n_k}} > 3^{k+1}\left(k + \sum_{j=1}^{k-1} 3^{-j} \|Ae_{n_j}\|\right).$$

This can be done by the indirect hypothesis, and together with (2.2) and (2.3) this implies $\|Ax\| > k$ for all k, which is impossible. Q.e.d.

3. Additive and multiplicative properties

To state our results exactly, we need some concepts. Let H, K, L be Hilbert spaces. We define the sets

$$\sum(H, K) \quad \text{and} \quad \prod(H, K, L) \subset R^3_+$$

in the following way: $(a, b, c) \in \sum(H, K)$ if there are transformations $A_1, A_2 \colon H \to K$ such that

$$\psi(A_1) = a, \quad \psi(A_2) = b, \quad \psi(A_1 + A_2) = c,$$

and $(a, b, c) \in \prod(H, K, L)$ if there are transformations $A \colon H \to K$ and $B \colon K \to L$ such that

$$\psi(A) = a, \quad \psi(B) = b, \quad \psi(BA) = c.$$

In case $H = K = L$ we shall write simply $\sum(H)$ and $\prod(H)$.

The complete description of these sets would need a rather lengthy discussion, therefore we shall confine ourselves to the main cases a) finite dimension and b) $H = K = L$.

THEOREM 2. (Additive properties.) a) *If* $\dim H < \aleph_0$, *then* ψ *is subadditive on the transformations* $A \colon H \to K$ *(K being arbitrary) and this is its only additive property, that is,*

$$\sum(H, K) = \{(a, b, c) \colon |a - b| \leqq c \leqq a + b\}.$$

b) *If* $\dim H \geqq \aleph_0$, *then* ψ *has no additive property on the transformations* $A \colon H \to H$, *that is,*

$$\sum(H) = R^3_+.$$

The multiplicative behaviour is more varied. We split the spaces H into two classes:

I. $\dim H' < \dim H$ implies $\operatorname{Dim} H' < \operatorname{Dim} H$;

II. there is an H' for which $\dim H' < \dim H$, but $\operatorname{Dim} H' = \operatorname{Dim} H$.

(Of course, this classification depends only on $\dim H = \alpha$. H belongs to the second class if and only if either $\alpha = \beta^{\aleph_0}$ for some $\beta < \alpha$ or $\aleph_0 < \alpha < 2^{\aleph_0}$. $\aleph_0$ belongs to class I, $\aleph_1$ and $2^{\aleph_0}$ belong to class II, and both classes are easily shown to be cofinal.)

THEOREM 3. (Multiplicative properties.) a) *If* $\dim H < \aleph_0$ *and* $2 \leqq n = \dim K < \aleph_0$, *then the only multiplicative property of* ψ *is*

$$\psi(BA) \leqq \sqrt{n}\, \psi(B) \psi(A),$$

that is

$$\prod(H, K, L) = \{(a, b, c) \colon c \leqq ab\sqrt{n}\}.$$

(The case $n = 1$ *is obvious.)*

b) *If* $\dim H \geqq \aleph_0$, *then either* ψ *is supermultiplicative or it has no multiplicative property at all. More precisely:*

(I) $\prod(H) = \{(a, b, c) \colon c \geqq ab\}$ *or*

(II) $\prod(H) = R^3_+$,

according as H *belongs to class* I *or* II *defined above.*

4. The finite dimensional case

If $A\colon H\to K$ is a transformation, $\dim H<\aleph_0$, then $\psi(A)$ is closely related to the Hilbert—Schmidt norm of A, defined by

$$h(A)=\left(\sum_{j=1}^{n}\|Ae_j\|^2\right)^{1/2}, \tag{4.1}$$

where $\{e_j\}$ is a basis. (4.1) is easily shown to be independent from the particular choice of the basis. The Hilbert—Schmidt norm has meaning for some operators in infinite dimension as well, but is no longer connected with our pseudonorm. Or to be more honest, it has a trivial connection: all Hilbert—Schmidt operators have a pseudonorm 0.

THEOREM 4. *If* $A\colon H\to K$ *is a transformation,* $\dim H=n<\aleph_0$, *then we have*

$$\psi(A)=\frac{h(A)}{\sqrt{\dim H}}=\left(\frac{1}{n}\sum_{j=1}^{n}\|Ae_j\|^2\right)^{1/2}. \tag{4.2}$$

PROOF. Denote the right-hand side of (4.2) temporarily by r. The inequality $\psi(A)\geqq r$ is obvious, a maximum is always greater than or equal to the quadratic mean. To prove the reverse inequality, we must find a basis for which $\|Ae_j\|=r$.

Regard a basis $\{f_j\}$. By (4.1) there must be a j and a k for which $\|Af_j\|\geqq r$, $\|Af_k\|\leqq r$. Since the unit sphere is connected and $f(x)=\|Ax\|$ is continuous, there must be a unit vector e_1 satisfying $\|Ae_1\|=r$.

Let $H_1=e_1^{\perp}$. Applying the definition (4.1) to a basis containing our e_1 we see that

$$h^2(A)=h^2(A_1)+r^2\quad (A_1=A|_{H_1}).$$

Therefore $\dfrac{h(A_1)}{\sqrt{n-1}}=r$, and iterating the argument we get the required basis. Q.e.d.

The positive part of Theorems 2a and 3a is merely a reformulation of the sub-additive and submultiplicative property of the Hilbert—Schmidt norm. Now we show that there is no other multiplicative property. It is sufficient to prove that

$$\{h(BA)\colon h(B)=h(A)=1\}=[0,1]. \tag{4.3}$$

Since the unit sphere of operators in the Hilbert—Schmidt norm is connected, the left-hand side of (4.3) must be a connected set and it is sufficient to show that it contains both 0 and 1. Let $e\in H$, $f_1, f_2\in K$ and $g\in L$ be unit vectors, $f_1\perp f_2$. For the operators

$$A_1\colon H\to K,\quad A_1x=\langle e,x\rangle f_1,$$

$$A_2\colon H\to K,\quad A_2x=\langle e,x\rangle f_2,$$

$$B\colon K\to L,\quad By=\langle f_1,y\rangle g$$

we have $h(A_1)=h(A_2)=h(B)=1$, $h(BA_1)=1$ and $h(BA_2)=0$, so we are ready. The assertion concerning additive properties can be proved similarly.

5. The infinite dimensional case

Let H_1 be a subspace of the Hilbert space H, $\dim H \geqq \aleph_0$. Call H_1 *small* if $\dim H_1 < \dim H$ and *large* if $H_1^\perp$ is small.

THEOREM 5. *Let* $A\colon H \to K$ *be a transformation,* $\dim H \geqq \aleph_0$. *We have* $\psi(A) = \sup \operatorname{cn} A|_{H_1}$, *where* H_1 *runs over the large subspaces of* H.

(5.1) LEMMA. *Let* $\{e_j : j \in J\}$ *be an ON system in a Hilbert space* H, $|J| = \dim H \geqq \aleph_0$ *and let* H_1 *be a small subspace of* H. *Let* e_j^1 *be the projection of* e_j *onto* H_1. *We have*

a) $e_j^1 \to 0$ *if* $\dim H = \aleph_0$,

b) $|\{j : e_j^1 \neq 0\}| < \dim H$ *if* $\dim H > \aleph_0$.

PROOF. a) is well-known (and very easy). To prove b) let $\{f_m : m \in M\}$ be a basis of H_1. By Bessel's inequality we have $\sum_j |\langle f_m, e_j\rangle|^2 \leqq \|f_m\|^2 = 1$, therefore $|\{j : \langle f_m, e_j\rangle \neq 0\}| \leqq \aleph_0$. This implies

$$\forall m \, \langle f_m, e_j\rangle = 0 \tag{5.2}$$

for all j, with at most $\aleph_0 \cdot |M| = \aleph_0 \cdot \dim H_1 < \dim H$ exceptions. (5.2) just means $e_j \perp H_1$, that is $e_j^1 = 0$. Q.e.d.

Let $\{a_j : j \in J\}$ be a set of real numbers, $|J| \geqq \aleph_0$. We call b an essential lower bound for this set if $|\{j : a_j < b\}| < |J|$ and we define the essential infimum by

$$\inf \operatorname{ess} \{a_j : j \in J\} = \sup \{b : b \text{ is an ess. lower bound}\}.$$

(For a countable J the latter coincides with the limit inferior.)

(5.3) LEMMA. *Let* $\{e_j : j \in J\}$ *be an ON system in a Hilbert space* H, $|J| = \dim H \geqq \aleph_0$, $A\colon H \to K$ *a transformation and* H_1 *a large subspace of* H. *We have*

$$\operatorname{cn} A|_{H_1} \leqq \inf \operatorname{ess} \|Ae_j\|.$$

PROOF. Let $H_2 = H_1^\perp$ (H_2 is small) and let $e_j = e_j^1 + e_j^2$, $e_j^i \in H_i$. If $\dim H = \aleph_0$, we have $e_j^2 \to 0$ by Lemma (5.1). Since $\dim H_2$ is finite, $A|_{H_2}$ is continuous, therefore $Ae_j^2 \to 0$. This implies

$$\operatorname{cn} A|_{H_1} \leqq \liminf \frac{\|Ae_j^1\|}{\|e_j^1\|} \leqq \liminf \frac{\|Ae_j\| + \|Ae_j^2\|}{1 - \|e_j^2\|} = \liminf \|Ae_j\|.$$

If $\dim H > \aleph_0$, by Lemma (5.1) we have $e_j \in H_1$ with less than $\dim H$ exceptions, which do not affect the essential infimum. Therefore

$$\inf \operatorname{ess} \|Ae_j\| = \inf \operatorname{ess} \{\|Ae_j\| : e_j \in H_1\} \geqq \operatorname{cn} A|_{H_1}.$$

Q.e.d.

PROOF OF THEOREM 5. Temporarily introduce the notation

$$\varphi(A) = \sup \{\operatorname{cn} A|_{H_1} : H_1 \text{ is a large subspace}\}.$$

$\varphi(A) \leqq \psi(A)$ follows from Lemma (5.3); now we prove $\varphi(A) \geqq \psi(A)$.

Let $\varepsilon>0$ be arbitrary. We shall construct a basis $\{e_j\}$ satisfying

(5.4) $$\|Ae_j\| \leqq \varphi(A)+\varepsilon.$$

Let α be the initial ordinal of cardinality $\dim H$ and let $\{f_\beta\colon \beta<\alpha\}$ be a basis in H. We shall define recursively a sequence $\{E_\beta\colon \beta<\alpha\}$ where E_β is a finite set of unit vectors such that

a) $\bigcup E_\beta$ is an ON system,

b) $f_\beta \in \operatorname{span} \bigcup_{\gamma\leqq\beta} E_\gamma$

c) $\|Ae\| \leqq \varphi(A)+\varepsilon$ for $e\in E_\beta$.

$\bigcup E_\beta$ will be the required basis; it is ON by a), complete by b) and satisfies (5.4) by c).

Now suppose that E_γ has been defined for $\gamma<\beta$. Let $H_1=\operatorname{span} \bigcup_{\gamma<\beta} E_\gamma$ (it is a small subspace) and $S_0=H_1^\perp$. Let f^* be the projection of f_β onto S_0. If $f^*=0$, we may set $E_\beta=\emptyset$. If $f^*\neq 0$, let $g_0=f^*/\|f^*\|$ and $S_1=S_0\cap g_0^\perp$.

S_1 is a large subspace, therefore $\operatorname{cn} A|_{S_1} \leqq \varphi(A)$. Let $g_1\in S_1$ be a unit vector for which $\|Ag_1\|\leqq\varphi(A)+\varepsilon$. Let $S_2=S_1\cap g_1^\perp$ and so on. Repeating the argument we get an ON sequence $g_1, \dots g_n$ in S_1 such that $\|Ag_i\|\leqq\varphi(A)+\varepsilon$, $i=1, \dots, n$. Let $S=\operatorname{span}\{g_0, g_1, \dots g_n\}$. By Theorem 2,

$$\psi(A|_S) = \left(\frac{1}{n+1}\sum_{j=0}^{n}\|Ag_j\|^2\right)^{1/2} \leqq \left(\frac{\|Ag_0\|^2+n\big(\varphi(A)+\varepsilon\big)^2}{n+1}\right)^{1/2} \leqq \varphi(A)+2\varepsilon$$

if n is large enough. Choose a large n and an ON basis $\{e_0, \dots e_n\}$ in S such that $\|Ae_i\|=\psi(A|_S)\leqq\varphi(A)+2\varepsilon$.

Then $E_\beta=\{e_0, \dots e_n\}$ will do (with 2ε instead of ε). Q.e.d.

(5.5) Corollary. *If H_1 is a large subspace of H, then* $\psi(A|_{H_1})=\psi(A)$.

(5.6) Corollary. *If $A\colon H\to K$ is a transformation,* $\dim H\geqq\aleph_0$ *and* $\dim K<\dim H$, *then* $\psi(A)=0$.

Proof. Otherwise there would be a large subspace H_1 of H such that $\operatorname{cn} A_1>0$, $A_1=A|_{H_1}$. This means that A_1 has an inverse $B\colon \operatorname{im} A_1\to H_1$, $\operatorname{im} B=H_1$, $\|B_1\|<\infty$. Hence $\dim H_1=\dim \operatorname{im} B\leqq\dim K$, a contradiction. Q.e.d.

Now we use Theorem 5 to determine $\psi(A)$ in an important special case.

(5.7) Corollary. *If A is a diagonal operator with the diagonal $\{d_j\colon j\in J\}$, then we have* $\psi(A)=\inf\operatorname{ess}|d_j|$.

Proof. Let A be diagonal in basis $\{e_j\}$, $Ae_j=d_je_j$ and denote $\inf\operatorname{ess}|d_j|=D$. We have $\operatorname{cn} A|_{H_1}\leqq D$ for every large subspace H_1 by Lemma (5.3). Therefore $\psi(A)\leqq D$ by Theorem 5.

On the other hand, considering the large subspaces $H_\varepsilon=\operatorname{span}\{e_i\colon |d_i|>D-\varepsilon\}$ we get $\psi(A)\geqq\sup\operatorname{cn} A|_{H_\varepsilon}=D$. Q.e.d.

Proof of Theorem 2b. For arbitrary $a, b, c\geqq 0$ it is easy to construct sets $\{a_j\}$ and $\{b_j\}$, $j\in J$, $|J|=\dim H$ such that $\inf\operatorname{ess}|a_j|=a$, $\inf\operatorname{ess}|b_j|=b$ and $\inf\operatorname{ess}|a_j+b_j|=c$. Hence Theorem 2b follows from Corollary (5.7). Q.e.d.

6. Proof of Theorem 3b

We begin with showing $\Pi(H)=R_+^3$ if H belongs to Class II. Let H_1 be another Hilbert space with the properties $\dim H_1<\dim H$, $\operatorname{Dim} H_1=\operatorname{Dim} H$. Given $a, b, c\geqq 0$, we shall construct transformations $A\colon H\to H\oplus H_1$, $B\colon H\oplus H_1\to H$ such that $\psi(A)=a$, $\psi(B)=b$, $\psi(BA)=c$. (Since Hilbert spaces of the same dimension are isomorphic, instead of H we may use any Hilbert space of the same dimension; sometimes this makes the description simpler.)

Let $P\colon H\to H_1$ be a vector-space isomorphism and $Q=P^{-1}$. Now let $Ax=(ax, Px)$ and $B(x, x_1)=bx+(c-ab)Qx_1$ $(x\in H,\ x_1\in H_1)$. We have $BA=cI_H$, therefore $\psi(BA)=c$ as wanted. Since $H'=\{(x,0)\colon x\in H\}$ is a large subspace of $H\oplus H_1$, we have $\psi(B)=\psi(B|_{H'})=\psi(bI_H)=b$ by Corollary (5.5).

We have

$$\|Ax\| = \sqrt{a^2\|x\|^2+\|Px\|^2}, \tag{6.1}$$

therefore $\psi(A)\geqq \operatorname{cn} A\geqq a$. On the other hand, $\psi(P)=0$ by Corollary (5.6), so that for every $\varepsilon>0$ there exists an ON basis $\{e_j\}$ satisfying $\|Pe_j\|<\varepsilon$. By (6.1) we have

$$\forall j(\|Ae_j\|\leqq a+\varepsilon),\quad \psi(A)\leqq a+\varepsilon$$

and we are ready.

Now we prove

$$\psi(BA)\geqq\psi(B)\psi(A) \tag{6.2}$$

if H belongs to Class I.

(6.3) LEMMA. *If H belongs to Class* I, *then we have* $\dim H'<\dim H\Rightarrow\operatorname{Dim} H'<\dim H$ *for every Hilbert space H'.*

PROOF. Let $\dim H=d$, $\operatorname{Dim} H=D$, $\dim H'=d'$, and $\operatorname{Dim} H'=D'$. If $d'<\aleph_0$, then $D'=d'<d$ and we are ready. If $d'\geqq\aleph_0$, then we have $D'=d'^{\aleph_0}$. Suppose $D'\geqq d$. This implies $D'^{\aleph_0}\geqq d^{\aleph_0}=D$. On the other hand, $D'^{\aleph_0}=(d'^{\aleph_0})^{\aleph_0}=d'^{\aleph_0}=D'$. This yields $D'\geqq D$, a contradiction. Q.e.d.

Call a linear manifold M in a Hilbert space H large, if its algebraic codimension is less than $\dim H$, that is, in a decomposition $H=M+M_1$ we have $\operatorname{Dim} M_1<\dim H$. Evidently if a large manifold is a subspace, it must be a large subspace. The converse, that a large subspace must be a large manifold, is wrong if H belongs to Class II, but according to Lemma (6.3) it is true in Class I.

(6.4) LEMMA. *Let H be a Hilbert space of Class* I, *$A\colon H\to K$ a transformation, $\{e_j\colon j\in J\}$ an ON system in H, $|J|=\dim H$ and M a large manifold in H. We have*

$$\operatorname{cn} A|_M\leqq s=\sup\|Ae_j\|. \tag{6.5}$$

(Strictly speaking, $\operatorname{cn} A|_M$ has not been defined yet but its meaning is evident.)

Proof. First suppose $\dim H=\aleph_0$. Then M has a codimension $m<\aleph_0$. Consider a subspace $H_1=\operatorname{span}\{e_{j_1}, \dots, e_{j_n}\}$, $n>m$. For $H_2=H_1\cap M$

(6.6) $$\dim H_2 \geqq \dim H_1 - \operatorname{codim} M = n-m$$

holds. Moreover we have $h(A|_{H_2})\leqq h(A|_{H_1})\leqq\sqrt{n}s$, where h denotes the Hilbert—Schmidt norm (see Section 4) Therefore by Theorem 4

$$\operatorname{cn} A|_M \leqq \operatorname{cn} A|_{H_2} \leqq \psi(A|_{H_2}) = \frac{h(A|_{H_2})}{\sqrt{\dim H_2}} \leqq \sqrt{\frac{n}{n-m}}\,s;$$

making $n\to\infty$ we get (6.5).

Next let $\dim H>\aleph_0$. Since H belongs to Class I, this implies $\dim H>2^{\aleph_0}$. (Otherwise a subspace of countable dimension would have the same algebraic dimension.) Let $H=M+M_1$, $\operatorname{Dim} M_1<\dim H$ and $e_j=e_j^0+e_j^1$, $e_j^0\in M$, $e_j^1\in M_1$. Since $|M_1|=2^{\aleph_0}\operatorname{Dim} M_1<\dim H=|J|$, there must be an infinity of j's with the same e_j^1, say $j_1, j_2, \dots$. Now let

$$H_1 = \operatorname{span}\{e_{j_1}, \dots, e_{j_n}\}, \quad H_2 = H_1\cap M.$$

Since $e_{j_i}-e_{j_k}\in M$, we have $\dim H_2\geqq n-1$, which is (6.6) with $m=1$ and the proof can be completed similarly. Q.e.d.

Now we prove (6.2). Let $\varepsilon>0$ and choose a large subspace H_1 for which $\operatorname{cn} B|_{H_1}\geqq\psi(B)-\varepsilon$. Since H belongs to Class I, H_1 is a large manifold and hence $M_1=A^{-1}H_1$ is a large manifold as well. Let H_2 be a large subspace for which $\operatorname{cn} A|_{H_2}\geqq\psi(A)-\varepsilon$. H_2 and so $M_2=M_1\cap H_2$ are large manifolds, therefore by Lemma (6.5) we have

(6.7) $$\operatorname{cn} BA|_{M_2} \leqq \psi(BA).$$

On the other hand, for $x\in M_2$ we have $Ax\in H_1$ and so

$$\|BAx\| \geqq (\psi(B)-\varepsilon)\|Ax\| \geqq (\psi(B)-\varepsilon)(\psi(A)-\varepsilon)\|x\|.$$

Therefore

(6.8) $$\operatorname{cn} BA|_{M_2} \geqq (\psi(A)-\varepsilon)(\psi(B)-\varepsilon).$$

(6.7) and (6.8) imply $\psi(BA)\geqq(\psi(A)-\varepsilon)(\psi(B)-\varepsilon)$; since ε was arbitrary, we are ready.

At last we have to show that there is no further connection between $\psi(A)$, $\psi(B)$ and $\psi(BA)$. Let $c\geqq ab$; we shall construct transformations A and B satisfying $\psi(A)=a$, $\psi(B)=b$ and $\psi(BA)=c$. Since the case $c=0$ is obvious, we may confine ourselves to the case $c=1$, $ab\leqq1$.

If $ab\neq0$, let A be a diagonal operator whose diagonal contains numbers a and $1/b$, each $\dim H$ times and $B=A^{-1}$. By Corollary (5.7) we have $\psi(A)=a$ and $\psi(B)=b$ as required.

If $b=0$, tt A_0 be an arbitrary invertible transformation with $\psi(A_0)=a$ and define $A\colon H\to H\oplus H$ and $B\colon H\oplus H\to H$ by $Ax=(A_0x, 0)$, $B(x,y)=A_0^{-1}x$. We have $\psi(A)=\psi(A_0)=a$ by Corollary (5.5), further $BA=I_H$ and $\psi(B)=0$ by

Lemma (5.3) (applied to a system $\{(0, e_j)\}$, where $\{e_j\}$ is an ON basis in H) and Theorem 5.

The remaining subcase $a=0$, $b\neq 0$ will be omitted, because it is of secondary importance and I managed to find only a rather complicated proof.

Acknowledgement. I am indebted to J. Bognár and A. Szép for their valuable remarks.

Reference

[1] P. R. HALMOS, *A Hilbert space problem book*, Van Nostrand (1967).

(Received January 17, 1978)

MATHEMATICAL INSTITUTE
OF THE HUNGARIAN ACADEMY OF SCIENCES
H—1053 BUDAPEST, REÁLTANODA U. 13—15.

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (3—4), (1979), 307—315.

ON AN EXTREMAL PROBLEM CONNECTED WITH THE FUNDAMENTAL POLYNOMIALS OF INTERPOLATION

By

CATHERINE BALÁZS (Budapest)

The aim of this paper is giving elementary proofs for some theorems of KARLIN and STUDDEN's book [5] and making them more complete with new results of similar nature.

The following problem will be investigated; let $a<b$ finite or infinite real numbers, and denote $\mathscr{P}_n$ the class of n-tuples $(x_1, x_2, \dots, x_n)$ with $a \leqq x_1 < \dots < x_n \leqq b$ (x_1 and x_n are finite). Let $w(x)$ be a positive function on (a, b), further

$$r_i(x) \overset{\text{def}}{=} \frac{l_i(x)}{w^{1/2}(x_i)} \quad (i = 1, 2, \dots, n)$$

where

$$l_i(x) \overset{\text{def}}{=} \frac{\omega_n(x)}{\omega_n'(x_i)(x-x_i)} \quad \text{and} \quad \omega_n(x) = \prod_{i=1}^{n} (x-x_i).$$

Consider the following extremal problem: determine the value

$$M = \inf_{\mathscr{P}_n} \sup_{a<x<b} w(x)\{r_1^2(x) + \dots + r_n^2(x)\} = \inf_{\mathscr{P}_n} \tilde{M}(x_1, \dots, x_n)$$

and the system of points $\{x_i\}_{i=1}^n$ which minimizes $\tilde{M}$ for a fixed $w(x)$.

FEJÉR's [1] well-known result is the case $w(x) \equiv 1$, $[a, b] = [-1, 1]$; $\tilde{M}$ is minimal if and only if $\{x_i\}_{i=1}^n$ are the zeros of the integral of the Legendre polynomial $\int_{-1}^{x} P_{n-1}(t)\,dt$ and $M=1$. We investigate the cases of the classical weight functions in this paper. The contents of Theorems 1, 3, 4 and 5 are given in KARLIN and STUDDEN's book [5] on p. 336 in Theorems 4.1, 4.2, 4.3 and 4.4. Theorem 2 is a result of the present author. The proofs given in [5] are based on deep general facts (such as von Neumann's minimax theorem), while our proofs are quite simple and elementary. However, the mentioned authors gave a common general setting of extremal problems which seemed to be rather different before.

THEOREM 1. *If* $w(x) = (1-x)^{\alpha+1}(1+x)^{\beta+1}$ $(\alpha, \beta > -1)$ *and* $[a, b] = [-1, 1]$, *then* $M=1$ *and this infimum is attained if and only if* $\{x_i\}_{i=1}^n$ *are the zeros of the Jacobi polynomial* $P_n^{(\alpha,\beta)}(x)$.

Now consider the case when one of α or β is equal to -1:

THEOREM 2. *If* $w(x) = (1-x)^{\alpha+1}$ $(\alpha > -1)$ *or* $w(x) = (1+x)^{\beta+1}$ $(\beta > -1)$ *and* $[a, b] = [-1, 1]$ *then* $M=1$ *and this infimum is attained uniquely when* $\{x_i\}_{i=1}^n$ *are* *the zeros of the Jacobi polynomial* $P_n^{(\alpha,-1)}(x)$ *and* $P_n^{(-1,\beta)}(x)$, *respectively, where*

(1)
$$P_n^{(\alpha,\beta)}(x) = \frac{1}{n!}\sum_{\nu=0}^{n}\binom{n}{\nu}(n+\alpha+\beta+1)\dots(n+\alpha+\beta+\nu)(\alpha+\nu+1)\dots(\alpha+n)\left(\frac{x-1}{2}\right)^{\nu}$$

(α, β *are arbitrary real numbers).*

We remark that Theorems 1 and 2 and Fejér's theorem mentioned above give a complete characterization of the zeros of Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$, $\alpha, \beta \geqq -1$. However, Theorem 2 is a marginal case $\alpha=-1$ or $\beta=-1$ of Theorem 1, its proof cannot be done simply as a special case of Theorem 1.

THEOREM 3. *If* $w(x)=e^{-x}$ *and* $[a, b)=[0, \infty)$ *then* $M=1$ *and this infimum is attained if and only if* $x_1=0$ *and* $x_2, \dots, x_n$ *are the zeros of the Laguerre polynomial* $L_{n-1}^{(1)}(x)$.

THEOREM 4. *If* $w(x)=x^{\alpha+1}e^{-x}$ $(\alpha>-1)$ *and* $[a, b)=[0, \infty)$ *then* $M=1$ *and this infimum is attained if and only if* $\{x_i\}_{i=1}^n$ *are the zeros of the Laguerre polynomial* $L_n^{(\alpha)}(x)$.

Taking into consideration the relation $xL_{n-1}^{(1)}(x)=L_n^{-1}(x)\cdot(-n)$ (SZEGŐ [8], (5.2.1)) Theorem 3 seems to be a marginal case $\alpha=-1$ of Theorem 4.

THEOREM 5. *If* $w(x)=e^{-x^2}$ *and* $(a, b)=(-\infty, \infty)$ *then* $M=1$ *and this infimum is attained if and only if* $\{x_i\}_{i=1}^n$ *are the zeros of the Hermite polynomial* $H_n(x)$.

The elementary proofs of Theorems 1, 2, 3, 4 and 5 are based on similar ideas. The proofs of Theorems 1, 4 and 5 follow the same way so we prove them by the aid of Theorem 6 stated under more general conditions. Theorems 2 and 3 need a special analysis.

THEOREM 6. *Suppose that* $w(x)>0$ *and* $(1/w(x))^{(2n)}\geqq 0$ $(a<x<b)$. *Then the system of points* $(a<)x_1<\dots<x_n(<b)$ *minimizes the expression*

(2)
$$\tilde{M}(x_1, \dots, x_n) = \sup_{a<x<b} w(x)\sum_{i=1}^{n}\frac{l_i^2(x)}{w(x_i)} \quad (n = 1, 2, \dots)$$

if and only if

(3)
$$\left(w(x)\omega_n'(x)\right)'_{x=x_i} = 0 \quad (i = 1, \dots, n)$$

and $M=\min\limits_{\mathcal{P}_n}\tilde{M}(x_1, \dots, x_n)=1$.

Finally consider an example which differs from the previous cases because the weight function $w(x)$ is "bad" at one point. Let $w(x)=e^{-x^2}|x|^{\beta}$ $(\beta>-1)$ and $(a, b)=(-\infty, \infty)$. The orthogonal polynomials $\{H_n^{(\beta)}(x)\}_{n=1}^{\infty}$ associated with the weight function $w(x)$ are called Sonin—Markov polynomials and are generalizations of the Hermite polynomials $H_n(x)=H_n^{(0)}(x)$ $(i=1, \dots, n)$ (SZEGŐ [9], p. 440). If $-1<\beta<0$ then $\tilde{M}(x_1, \dots, x_n)$ has no finite value for any system of points. A unique solution exists in the special case $\beta=0$ treated in Theorem 5. If $\beta>0$, then $1/w(x)$ is not differentiable at $x=0$, so the problem of uniqueness can not be solved in this way. I can prove the following:

THEOREM 7. *If* $w(x)=e^{-x^2}|x|^\beta$ $(\beta>0)$, $(a, b)=(-\infty, \infty)$ *and* $n=2m$, $m=1, 2, \ldots$ *then* $M=1$ *and this infimum is attained if* $\{x_i\}_{i=1}^n$ *are the zeros of the Sonin—Markov polynomial* $H_{2m}^{(\beta)}(x)$.

The condition $x_i\neq 0$ $(i=1, \ldots, n)$ must be fulfilled for any minimizing system of points in the case $\beta>0$, thus if $n=2m+1$ $(m=1, 2, \infty)$ then $M=1$ cannot be attained for the zeros of $H_{2m+1}^{(\beta)}(x)$, because $H_{2m+1}^{(0)}(0)=0$ (KIS [6], (27)).

PROOF OF THEOREM 6. $M\geqq 1$, because

$$w(x)\sum_{i=1}^n \frac{l_i^2(x)}{w(x_i)}=1 \tag{4}$$

at $x=x_i$, for arbitrary system of points.

Suppose now that (3) is fulfilled for a system of points $\{x_i\}_{i=1}^n$. We shall prove for it the inequality

$$V_n(x)=\frac{1}{w(x)}-\sum_{i=1}^n \frac{l_i^2(x)}{w(x_i)}\geqq 0 \quad (a<x<b). \tag{5}$$

(4) and (5) will show that $\tilde{M}(x_1, \ldots, x_n)=M=1$. The inequality (5) can be seen as follows: $V_n(x_i)=0$ $(i=1, \ldots, n)$. Using

$$2l_i(x_i)=\frac{\omega_n''(x_i)}{\omega_n'(x_i)} \quad (i=1, \ldots, n) \tag{6}$$

and (3), we get

$$V_n'(x_i)=-\frac{w'(x_i)}{w^2(x_i)}-\frac{2l_i'(x_i)}{w(x_i)}=-\frac{w'(x_i)\omega_n'(x_i)+w(x_i)\,\omega_n''(x_i)}{w^2(x_i)\omega_n'(x_i)}=$$

$$=\frac{-\left(w(x)\omega_n'(x_i)\right)'_{x=x_i}}{w^2(x_i)\omega_n'(x_i)}=0 \quad (i=1, \ldots, n)$$

that is the polynomial $\sum_{i=1}^n \frac{l_i^2(x)}{w(x_i)}$ of degree $2n-2$ coincides with $1/w(x)$ at $2n$ points, so it is the unique Hermite interpolating polynomial of degree at most $2n-1$ of $1/w(x)$ with nodes given above. We denote it by $H(x)$. Then the following theorem is true for $H(x)$ (NATANSON [7], p. 377).

$$V_n(x)=\frac{1}{w(x)}-H(x)=\frac{1}{(2n)!}\left[\frac{1}{w(x)}\right]^{(2n)}_{x=\xi}\prod_{i=1}^n (x-x_i)^2 \quad (a<\xi<b).$$

We obtain (5) by $(1/w(x))^{(2n)}>0$.

Conversely, suppose that $M=1$. Then the expression

$$0\leqq \frac{w(x)\,l_i^2(x)}{w(x_i)}\leqq 1 \quad (i=1, \ldots, n)$$

has a maximum equal to 1 at $x=x_i$, so

$$\left[\frac{w(x)\,l_i^2(x)}{w(x_i)}\right]'_{x=x_i}=0 \quad (i=1, 2, \ldots, n)$$

that is

$$\left(w(x)l_i^2(x)\right)'_{x=x_i} = w'(x_i)+w(x_i)2l_i'(x_i) = \frac{\left(w(x)\omega_n'(x)\right)'_{x=x_i}}{\omega_n'(x_i)} = 0 \quad (i=1,2,\dots,n)$$

which gives (2).

REMARK. Theorem 6 investigates systems of points $\{x_i\}_{i=1}^n$ with $x_1>a$ and $x_n<b$, and does not exclude the existence of such minimizing systems of points where $x_1=a$ or $x_n=b$. In this latter case we can have $\left(w(x)\omega_n'(x)\right)'_{x=x_i}\neq 0$ for $i=1$ and $i=n$, respectively. But if we suppose

$$\lim_{x\to a+0} w(x) = \lim_{x\to b-0} w(x) = 0$$

then for a minimizing system of points $\{x_i\}_{i=1}^n$ the relations $a<x_1$ and $x_n<b$ are true, otherwise $\tilde{M}(x_1,\dots,x_n)$ is not a finite value. So we can apply Theorem 6 in this case and we have that only systems of points satisfying (3) may be minimizing ones.

We shall need a

LEMMA. *The following inequalities hold:*

$$\left(\frac{1}{(1-x)^{\alpha+1}(1+x)^{\beta+1}}\right)^{(2n)} > 0 \quad (\alpha,\beta>-1,\ -1<x<1,\ n=1,2,\dots) \tag{7}$$

and

$$\left(\frac{e^x}{x^{\alpha+1}}\right)^{(2n)} > 0 \quad (\alpha>-1,\ x>0,\ n=1,2,\dots) \tag{8}$$

These inequalities were got by I. Joó. R. Askey (SZEGŐ [8], p. 391) gave a simple proof for (7) and (8). However, he used a deep theorem on the distribution of the zeros of the Jacobi and Laguerre polynomials. We remark that (7) and (8) follow trivially by differentiation under the integral in the formulae

$$\frac{1}{(1-x)^{\alpha+1}(1+x)^{\beta+1}} = \left(\frac{1}{\Gamma(\alpha+1)}\int_0^\infty v^\alpha e^{-(1-x)v}\,dv\right)\cdot\left(\frac{1}{\Gamma(\beta+1)}\int_0^\infty s^\beta e^{-(1+x)s}\,ds\right) =$$

$$= \frac{1}{\Gamma(\alpha+1)\Gamma(\beta+1)}\int_0^\infty\int_0^\infty v^\alpha s^\beta e^{-(v+s)} e^{x(v-s)}\,dv\,ds \quad (\alpha,\beta>-1,\ -1<x<1)$$

and

$$\frac{e^x}{x^{\alpha+1}} = \frac{1}{\Gamma(\alpha+1)}\int_0^\infty u^\alpha e^{-x(u-1)}\,du \quad (\alpha>-1,\ x>0)$$

which one can easily obtain from

$$\frac{1}{x^{\alpha+1}} = \frac{1}{\Gamma(\alpha+1)}\int_0^\infty u^\alpha e^{-xu}\,du \quad (\alpha>-1,\ x>0).$$

This last equality follows from the integral representation of the Gamma function

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}\,dt \quad (\operatorname{Re} z>0)$$

(DUNCAN [2], p. 264) with $z=\alpha+1$, $t=u\cdot x$.

PROOF OF THEOREM 1. By the Lemma and the Remark, Theorem 6 shows that we have to search for systems of points satisfying (3):

$$\big(w(x)\omega_n'(x)\big)'_{x=x_i} =$$
$$= (1-x_i)^\alpha(1+x_i)^\beta\{[-(\alpha+1)(1+x_i)+(\beta+1)(1-x_i)]\omega_n'(x)+(1-x_i)^2\omega_n''(x_i)\} = 0$$
$$(i=1,\dots,n)$$

that is

$$(1-x_i)^\alpha(1+x_i)^\beta\cdot\{(1-x_i^2)\omega_n''(x_i)-[\beta-\alpha-(\alpha+\beta+2)x_i]\omega_n'(x_i)\}=0 \quad (i=1,2,\dots,n).$$

Taking into consideration the fact $x_i\neq\pm 1$ and substituting x for x_i we can write

$$(1-x^2)\omega_n''(x)+[\beta-\alpha-(\alpha+\beta+2)x]\omega_n'(x) = C\omega(x),$$

since the left hand side is a polynomial of degree n having zeros $\{x_i\}_{i=1}^n$. Comparing the corresponding coefficients,

$$C=-n(n-1)-n(\alpha+\beta+2),$$

and by Theorem 4.2.2. of SZEGŐ [8]

$$\omega_n(x) = \text{const. } P_n^{(\alpha,\beta)}(x).$$

PROOF OF THEOREM 2. $M\geqq 1$, because (4) is valid.

Let us suppose e.g. that $\alpha=-1$, $\beta>-1$. We note that the important relation

(9) $$P_n^{(-1,\beta)}(x) = \text{const. }(x-1)P_{n-1}^{(1,\beta)}(x)$$

will be very useful in our proof. (9) is a trivial consequence of (1).

Let $-1<x_1<\dots<x_n=1$ be the zeros of the Jacobi polynomial $P_n^{(-1,\beta)}(x)$. Then $\tilde M(x_1,\dots,x_n)=M=1$. To see this we have to prove the inequality

(10)
$$A_n=\frac{1}{(1+x)^{\beta+1}}-\frac{1}{2^{\beta+1}}\left[\frac{P_{n-1}^{(1,\beta)}(x)}{P_{n-1}^{(1,\beta)}(1)}\right]^2-\sum_{i=1}^{n}\frac{1}{(1+x_i)^{\beta+1}}\left(\frac{1-x}{1-x_i}\right)^2\left[\frac{P_{n-1}^{(1,\beta)}(x)}{P_{n-1}^{(1,\beta)\prime}(x_i)(x-x_i)}\right]^2\geqq 0$$
$$(-1<x<1)$$

which is analogous to (5). (10) will be justified by the aid of

(11) $$V_n=A_n-B_n\geqq 0,$$

where

$$B_n=(1-x)\frac{1}{2^{\beta+2}}\left\{(\beta+1)+2^2\frac{P_{n-1}^{(1,\beta)\prime}(1)}{P_{n-1}^{(1,\beta)}(1)}\right\}\left[\frac{P_{n-1}^{(1,\beta)}(x)}{P_{n-1}^{(1,\beta)}(1)}\right]^2\geqq 0.$$

The proof of (11) is similar to that of (5). Using the relations

$$P_n^{(\alpha,\beta)}(1)=\binom{n+\alpha}{n}(\neq 0)$$

and

$$P_n^{(\alpha,\beta)\prime}(1)=\frac{1}{2}(\alpha+\beta+n+1)\binom{\alpha+n}{n-1}\quad(n\geqq 1)$$

(SZEGŐ [8], pp. 65 and 63, respectively) one can establish $V_n(x_i)=0$, $V_n'(x_i)=0$ $(i=1, \dots, n-1)$ and $V_n(1)=0$, $V_n'(1)=0$. So

$$V_n = A_n - B_n = \frac{1}{w(x)} - H(x)$$

where $H(x)$ is a Hermite interpolation polynomial of degree $2n-1$ of $1/w(x)$ based on the nodes $x_1, \dots, x_{n-1}$ and $x_n=1$ with multiplicities 2. By

$$\left(\frac{1}{w(x)}\right)^{(2n)} = \left(\frac{1}{(1+x)^{\beta+1}}\right)^{(2n)} > 0 \quad (-1 < x < 1)$$

and by the theorem mentioned earlier concerning Hermite interpolation we get

$$V_n = \frac{1}{w(x)} - H(x) = \frac{1}{(2n)!}\left[\frac{1}{w(x)}\right]^{(2n)}_{x=\xi} \prod_{i=1}^{n} (x-x_i)^2 \geqq 0 \quad (a < \xi < b)$$

which gives (10) by (11).

Conversely, let us suppose that

$$(1+x)^{\beta+1} \sum_{i=1}^{n} \frac{l_i^2(x)}{(1+x_i)^{\beta+1}} \leqq 1 \quad (-1 \leqq x \leqq 1). \tag{12}$$

$x_i \neq -1$ follows trivially from (12). The expression

$$0 \leqq (1+x)^{\beta+1} \frac{l_i^2(x)}{(1+x_i)^{\beta}} \leqq 1$$

has a maximum equal to 1 at $x=x_i$, when $-1<x_i<1$. So

$$\left[(1+x)^{\beta+1} \frac{l_i^2(x)}{(1+x_i)^{\beta+1}}\right]'_{x=x_i} = 0 \quad (-1 < x_i < 1)$$

that is

$$\frac{\beta+1}{1+x_i} + 2l_i'(x_i) = 0 \quad (-1 < x_i < 1) \tag{13}$$

which is equivalent to

$$(\beta+1)\,\omega_n'(x_i) + (1+x_i)\,\omega_n''(x_i) = 0 \quad (-1 < x < 1)$$

if we use (6) and the fact $x_i \neq -1$. Substituting x for x_i the left-hand side is a polynomial of degree at most $n-1$ with leading coefficient $(\beta+1)n+n(n-1)\neq 0$. So (13) holds only for $x_1, \dots, x_{n-1}$, and $x_n=1$ must be true. Denoting

$$\tilde{\omega}(x) = \prod_{i=1}^{n-1} (x-x_i) \quad \text{and} \quad l_i(x) = \frac{x-1}{x_i-1} \tilde{l}_i(x) \quad (i = 1, \dots, n-1),$$

(13) can be written in the form

$$(1-x_i^2)\tilde{\omega}''(x_i) + [\beta-1-(\beta+3)x_i]\tilde{\omega}'(x_i) = 0 \quad (i = 1, \dots, n-1)$$

using $2\tilde{l}_i(x_i)=\dfrac{\tilde{\omega}''(x_i)}{\tilde{\omega}'(x_i)}$. Replacing x for x_i the left-hand side is a polynomial of degree at most $n-1$ having zeros $\{x_i\}_{i=1}^{n-1}$, that is

$$(1-x^2)\tilde{\omega}''(x)+[\beta-1-(\beta+3)x]\tilde{\omega}'(x) = C\tilde{\omega}(x),$$

where $C=-(n-1)(n+\beta+1)$. By Theorem 4.2.2 of SZEGŐ [8] we have $\tilde{\omega}(x)=$ $=\text{const.}\ P_{n-1}^{(1,\beta)}(x)$, so taking into consideration (9) and the fact $x_n=1$, $x_1, \dots, x_n$ are uniquely the zeros of $P_n^{(-1,\beta)}(x)$ if $\tilde{M}(x_1, \dots, x_n)=1$.

PROOF OF THEOREM 3. We need the useful equality

$$L_n^{(-1)}(x) = -\frac{x}{n} L_{n-1}^{(1)}(x) \quad (x \geqq 0) \tag{14}$$

(SZEGŐ [8], (5.2.1)).

$M \geqq 1$, for (4) is valid.

Now let us suppose that $x_1=0$, $x_2, \dots, x_n$ are the zeros of $L_n^{(-1)}(x)$, we have to prove $\tilde{M}(x_1, \dots, x_n)=1$. This is an easy consequence of

$$e^x-\left(\frac{L_{n-1}^{(1)}(x)}{L_{n-1}^{(1)}(0)}\right)^2-\sum_{i=2}^{n} e^{x_i}\frac{x^2}{x_i^2}\left(\frac{L_{n-1}^{(1)}(x)}{L_{n-1}^{(1)\prime}(x_i)(x-x_i)}\right)^2 \geqq 0 \quad (x \geqq 0). \tag{15}$$

(15) follows trivially from Lemma 4.1 of JOÓ [3]:

$$e^x-[1+(n-1)x]\frac{(L_{n-1}^{(1)}(x))^2}{n^2}-\sum_{i=2}^{n} e^{x_i}\frac{x^2}{x_i^2}l_i^2(x)-x\frac{(L_{n-1}^{(1)}(x))^2}{n^2} \geqq 0$$

$$\left(x \geqq 0,\ n = 1, 2, \dots;\ L_{n-1}^{(1)}(x_i) = 0,\ i = 2, \dots, n\right)$$

by $L_{n-1}^{(1)}(0)=n$ and by omitting negative terms on the left hand side.

Conversely, assume that $\tilde{M}(x_1, \dots, x_n)=1$. We show that $L_n^{(-1)}(x_i)=0$ $(i=1, \dots, n)$. Namely $0\leqq e^{x_i}e^{-x}l_i^2(x)\leqq 1$ $(x_i>0)$, so by (6)

$$\left(e^{x_i}e^{-x}l_i^2(x)\right)'_{x=x_i} = -1+\frac{\omega_n''(x_i)}{\omega_n'(x_i)} = 0, \quad \text{when} \quad x_i > 0. \tag{16}$$

But (16) can be satisfied only by $n-1$ points x_i, for the degree of the polynomial $\omega_n''(x)-\omega_n'(x)$ is $n-1$. As a consequence we get $x_1=0$. Denoting $\tilde{\omega}(x)=\prod_{i=2}^{n}(x-x_i)$ and $l_i=\dfrac{x}{x_i}\tilde{l}_i$ $(i=2, \dots, n)$ and using (6) we write (16) in the form

$$x_i\tilde{\omega}''(x_i)+(2-x_i)\tilde{\omega}'(x_i) = 0 \quad (i = 2, \dots, n).$$

Substituting x for x_i, the left-hand side is a polynomial of degree at most $n-1$ having zeros $x_2, \dots, x_n$, that is

$$x\tilde{\omega}''(x)+(2-x)\tilde{\omega}'(x) = C\tilde{\omega}(x),$$

where $C=-(n-1)$. Then $\tilde{\omega}(x)=\text{const.}\ L_{n-1}^{(1)}(x)$ (SZEGŐ [8], (5.1.3)). Hence using (14) and $x_1=0$ our Theorem is proved.

Proof of Theorems 4 and 5. These proofs are completely analogous to that of Theorem 1: Theorem 6, Remark and Lemma show that we have to search for systems of points satisfying (3). We get by simple calculation and by (5.1.2) and (5.5.2) of Szegő [8], that (3) holds and $\tilde{M}(x_1, \ldots, x_n)=M=1$ if and only if $L_n^{(\alpha)}(x_i)=0$ $(\alpha>-1;\, i=1, \ldots, n;\, n=1, 2, \ldots)$ and $H_n(x)=0$ $(i=1, \ldots, n;\, n=1, 2, \ldots)$, respectively.

Proof of Theorem 7. $M\geqq 1$, because (4) holds. Now let $\{x_i\}_{i=1}^{2m}$ be the zeros of the Sonin—Markov polynomial $H_{2m}^{(\beta)}(x)$. To see $\tilde{M}(x_1, \ldots, x_{2m})=M=1$ we have to prove

$$V_n(x) = \frac{e^{x^2}}{|x|^\beta} - \sum_{i=1}^{2m} \frac{e^{x_i^2}}{|x_i|^\beta} l_i^2(x) = \frac{e^{x^2}}{|x|^\beta} - Q(x) \geqq 0 \quad (x \neq 0). \tag{17}$$

Since $H_{2m}^{(\beta)}(x)=a_m L_m^{(\alpha)}(x^2)$, where $L_m^{(\alpha)}$ is the Laguere polynomial of parameter $\alpha=\dfrac{\beta-1}{2}$ and a_m is a constant (Kis [6], (26)), the zeros of $H_{2m}^{(\beta)}(x)$ can be denoted by $\xi_{-m}, \xi_{-m+1}, \ldots, \xi_{-1}, \xi_1, \ldots, \xi_m$, where $\xi_i=x_{i+m}^2$, $\xi_i>0$, $\xi_{-i}=-\xi_i$ $(i=1, \ldots, m)$. $V_n(x)$ is an even function because $Q(x)$ is an even polynomial of x, so applying the substitution $y=x^2$ it is enough to prove

$$V_n(x) = A_n(y) = \frac{e^y}{y^{\beta/2}} - \sum_{i=1}^{m} \frac{e^{\xi_i^2}}{|\xi_i|^\beta} [\tilde{l}_i^2(\sqrt{y})+\tilde{l}_{-i}^2(\sqrt{y})] = \frac{e^y}{y^{\beta/2}} - R(y) \geqq 0 \quad (0<y<\infty)$$

where $\tilde{l}_i(x)=l_{i+m}(x)$ and $\tilde{l}_{-i}(x)=l_{m-i}(x)$ $(i=1, \ldots, m)$ and $R(y)$ is a polynomial of degree $n-1=2m-1$. Clearly

$$A_n(\xi_i^2) = V_n(\xi_i) = 0 \quad (i = 1, \ldots, m). \tag{18}$$

We show that

$$\frac{d}{dy} A_n(\xi_i^2) = 0 \quad (i = 1, \ldots, m). \tag{19}$$

We get (19) by a standard calculation if we use

$$2\tilde{l}_i(\xi_i) = \frac{\omega_n''(\xi_i)}{\omega_n'(\xi_i)} = \frac{1}{\xi_i} + 2\xi_i \frac{L_m^{(\alpha)\prime\prime}(\xi_i^2)}{L_m^{(\alpha)\prime}(\xi_i^2)} = \frac{2\xi_i^2-\beta}{\xi_i} \quad (i = 1, \ldots, m)$$

which follows from

$$\omega_n(x) = \prod_{i=-m}^{m} (x-\xi_i) = c_m L_m^{(\alpha)}(x^2), \quad \omega_n'(\xi_i) = 2c_m \xi_i L_m^{(\alpha)\prime}(\xi_i^2),$$

$$\omega_n''(\xi_i) = c_m[2L_m^{(\alpha)\prime}(\xi_i^2) + 4\xi_i^2 L_m^{(\alpha)\prime\prime}(\xi_i^2)]$$

(c_m is a constant depending only on m) and the differential equation of $L_m^{(\alpha)}(x)$ (Szegő [8], (5.1.2)). (18) and (19) mean that $A_n(y)$ has at least $n=2m$ zeros in $(0, \infty)$. Since $\lim_{y\to 0+} A_n(y)=\infty$ and $\lim_{y\to\infty} A_n(y)=\infty$, so if there were one point $\eta>0$, where $A_n(\eta)<0$, then $A_n(y)$ would have at least one more zero i.e. altogether

$2m+1$ zeros in $(0, \infty)$, and $A_n^{(2m)}(y)$ would have at least one positive zero according to Rolle's theorem. But this is impossible because we have from (8)

$$A_n^{(2m)}(y) = \left(\frac{e^y}{y^{\beta/2}}\right)^{(2m)} > 0 \quad (y > 0, \beta > 0)$$

so (17) is proved.

Finally I thank Professor J. Szabados for his help during the preparation of this paper and Professor O. Kis for calling my attention to the Sonin—Markov case.

References

[1] L. Fejér, Bestimmung derjenigen Abszissen eines Intervalles, für welche die Quadratsumme der Grundfunktionen der Lagrangeschen Interpolation im Intervalle ein möglichst kleines Maximum besitzt, *Annali della Scuola Normale Superiore di Pisa*, **1** (1932), 3—16.

[2] J. Duncan, *The Elements of Complex Analysis*, Műszaki Könyvkiadó (Budapest, 1974) (in Hungarian).

[3] J. Joó, Stable interpolation on an infinite interval, *Acta Math. Acad. Sci. Hungar.*, **25** (1974), 147—157.

[4] I. Joó, An interpolation-theoretical characterization of the classical orthogonal polynomials, *Acta Math. Acad. Sci. Hungar.*, **26** (1975), 163—169.

[5] S. Karlin and W. J. Studden, *Tchebycheff Systems: with Applications in Analysis and Statistics*, Wiley (New York, 1966).

[6] О. Киш, О Лагранжевом интерполировании с узлами в корнях многочленов Сонина-Маркова, *Acta Math. Acad. Sci. Hungar.*, **23** (1973), 389—417.

[7] I. P. Natanszon, *Konstruktiv függvénytan*, Akadémiai Kiadó (Budapest, 1952) (in Hungarian).

[8] G. Szegő, *Orthogonal Polynomials*. Amer. Math. Soc. Coll. Publ. (New York, 1975).

[9] Г. Сегё, *Ортогональные многочлены*. Физматгиз, (Москва, 1962).

(Received January 20, 1978)

KARL MARX UNIVERSITY
DEPARTMENT OF MATHEMATICS
1093 BUDAPEST
DIMITROV TÉR 8

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (3—4), (1979), 317—320.

ON CONTROLLABLY PERIODIC PERTURBATIONS OF AUTONOMOUS FUNCTIONAL DIFFERENTIAL EQUATIONS

By

B. M. GARAY (Budapest)

As usual, let $C([a, b], \mathbf{R}^n)$ be the space of continuous functions mapping the interval $[a, b]$ into an n-dimensional real vector space $\mathbf{R}^n$. Let r be a fixed positive number. For $C([-r, 0], \mathbf{R}^n)$ we write briefly C. For any φ in C, we define $|\varphi| = \sup_{-r\leqq\theta\leqq 0} |\varphi(\theta)|$, where $|x|$ is the norm of a vector in $\mathbf{R}^n$. If $x\in C([a-r, a+A], \mathbf{R}^n)$, $A>0$ and t is a given element of $[a, a+A]$, define x_t in C by $x_t(\theta)=x(t+\theta)$, $-r\leqq\theta\leqq 0$.

Let f be a function taking C into $\mathbf{R}^n$. Assume that $f(\varphi)$ has a continuous Frechet derivative, and consider the autonomous functional differential equation

$$\dot{x}(t) = f(x_t). \tag{1}$$

It is well known that the continuous differentiability of f implies the existence and uniqueness of a solution $x(\delta, \varphi)$ of (1) through (δ, φ), for any (δ, φ), i.e. we can define the operator Φ: $\mathbf{R}\times C\to C$, $(t, \varphi)\to x_t(0, \varphi)$, for $t\geqq 0$. The smoothness hypothesis on f implies that $\Phi(t, \varphi)$ has a continuous partial Frechet derivative $\frac{d}{d\varphi}\Phi(t, \varphi)$ with respect to the second variable in its domain of definition.

Let us assume that (1) has a non-constant periodic solution p with period $\tau_0>r$. The linear variational equation relative to p is

$$\dot{y}(t) = \frac{d}{d\varphi} f(p_t)\, y_t. \tag{2}$$

It is known that for a solution $y(0, \varphi)$ of (2) belonging to the initial values $(0, \varphi)$,

$$y_t(0, \varphi) = \frac{d}{d\varphi}\Phi(t, p_0)\varphi$$

holds. From $\tau_0>r$ it follows the compactness of the linear operator $\frac{d}{d\varphi}\Phi(\tau_0, p_0)$. We call its eigenvalues the characteristic multipliers of (2). Since $\dot{p}$ is a nontrivial τ_0-periodic solution of (2), one of the characteristic multipliers of (2) is 1. Let us assume that 1 is a simple characteristic multiplier, i.e. $\dot{p}_0$ is a generator for the generalised eigenspace of the multiplier 1, or, equivalently

$$C = H\oplus\{\lambda\dot{p}_0 | \lambda\in\mathbf{R}\} \tag{3}$$

where

$$\{\lambda\dot{p}_0|\lambda\in\mathbf{R}\} = \mathrm{Ker}\left(\frac{d}{d\varphi}\Phi(\tau_0, p_0)-\mathrm{Id}_C\right),$$

$$H = R\left(\frac{d}{d\varphi}\Phi(\tau_0, p_0)-\mathrm{Id}_C\right) \quad \text{and} \quad \frac{d}{d\varphi}\Phi(\tau_0, p_0)-\mathrm{Id}_C$$

restricted to H is continuously invertible on H.

For some fixed ϱ, $0<\varrho<\tau_0$, denote $\{\mu\in\mathbf{R}\,|\,|\mu|<\varrho\}$ and $\{\tau\in\mathbf{R}\,|\,|\tau-\tau_0|<\varrho\}$ by I_0 and I_{τ_0}, resp. Let g be a function taking $\mathbf{R}\times C\times I_0\times I_{\tau_0}$ into $\mathbf{R}^n$. Assume that g is continuous and has continuous partial derivatives in its second and fourth variables, and in addition, $g(s+1, \varphi, \mu, \tau)\equiv g(s, \varphi, \mu, \tau)$. Consider the perturbed system

$$\dot{x}(t) = f(x_t)+\mu g\left(\frac{t}{\tau}, x_t, \mu, \tau\right). \tag{4}$$

The attention is drawn to the fact that the period of the non-autonomous periodic perturbation $\mu g\left(\frac{t}{\tau}, x_t, \mu, \tau\right)$ occurs in g as a parameter (besides the "small parameter" μ) and can be chosen appropriately if necessary. We express this fact by saying that the period of the perturbation is controllable and therefore, we call (4) the controllably periodic perturbation of (1).

Our assumptions assure a unique solution $x(\vartheta, \varphi, \mu, \tau)$ of (4) through (ϑ, φ), for $(\vartheta, \varphi)\in\mathbf{R}\times C$. For $t\geqq\vartheta$, we define the operator Ψ taking $\mathbf{R}\times\mathbf{R}\times C\times\mathbf{R}\times\mathbf{R}$ into C by $\Psi(t, \vartheta, \varphi, \mu, \tau)=x_t(\vartheta, \varphi, \mu, \tau)$. Ψ is continuous and has continuous partial derivatives in its third and fifth variables. We recall that for $t\geqq 0$, $\tau\in I_{\tau_0}$, $\Psi(t+\tau, \tau, \varphi, 0, \tau)=\Phi(t, \varphi)$.

We are going to prove the following

THEOREM. *Under the previous conditions, to given μ and ϑ, where $|\mu|$ and $|\vartheta-\tau_0|$ are sufficiently small, there belongs a unique period $\tau(\mu, \vartheta)$ and a unique $h(\mu, \vartheta)\in H$ such that*

$$\omega(t, \vartheta, \mu) = \Psi\big(t, \vartheta, p_0+h(\mu, \vartheta), \mu, \tau(\mu, \vartheta)\big)$$

is a periodic solution of (4) *with period $\tau(\mu, \vartheta)$. The functions τ: $\mathbf{R}\times\mathbf{R}\to\mathbf{R}$, h: $\mathbf{R}\times\mathbf{R}\to H$ are continuous in some neighbourhood of* $(0, \tau_0)\in\mathbf{R}\times\mathbf{R}$, $h(0, \tau_0)=0$, $\tau(0, \tau_0)=\tau_0$ *and* $\omega(t, \vartheta, 0)=p_{t-\vartheta}$.

PROOF. For given μ and ϑ, where $|\mu|$ and $|\vartheta-\tau_0|$ are sufficiently small, we have to find a $\tau\in I_{\tau_0}$ and an $h\in H$ such that the solution $\Psi(t, \vartheta, p_0+h, \mu, \tau)$ of (4) through (ϑ, p_0+h) satisfies

$$\Psi(\vartheta+\tau, \vartheta, p_0+h, \mu, \tau) = \Psi(\vartheta, \vartheta, p_0+h, \mu, \tau) \equiv p_0+h$$

or, equivalently, with the function z defined by

$$z(\vartheta, h, \mu, \tau) = \Psi(\vartheta+\tau, \vartheta, p_0+h, \mu, \tau)-p_0-h, \quad z(\vartheta, h, \mu, \tau) = 0.$$

z, taking $\mathbf{R}\times H\times\mathbf{R}\times\mathbf{R}$ into C, is defined in a neighbourhood of $(\tau_0, 0, 0, \tau_0)$. z is continuous and continuously differentiable in its second and fourth variables. We will show that the conditions of the implicit function theorem are fulfilled. We have

$$z(\tau_0, 0, 0, \tau_0) = \Psi(2\tau_0, \tau_0, p_0, 0, \tau_0) - p_0 = \Phi(\tau_0, p_0) - p_0 = p_{\tau_0} - p_0 = 0.$$

Since the partial derivatives $\frac{d}{d\tau} z$ and $\frac{d}{dh} z$ are continuous, $\frac{d}{d(\tau, h)} z$ (the total derivative of z with respect to (τ, h)) exists and is continuous in its domain of definition. We have to show that $\frac{d}{d(\tau, h)} z(\tau_0, 0, 0, \tau_0)$, taking $\mathbf{R}\times H$ into C, is continuously invertible on C. We have

$$\frac{d}{d\tau} z(\tau_0, 0, 0, \tau_0) = \frac{d}{dt} \Psi(2\tau_0, \tau_0, p_0, 0, \tau_0) = \frac{d}{dt} \Phi(\tau_0, p_0) = \dot{p}_{\tau_0} = \dot{p}_0$$

and

$$\frac{d}{dh} z(\tau_0, 0, 0, \tau_0) = \left(\frac{d}{d\varphi} \Psi(2\tau_0, \tau_0, p_0, 0, \tau_0) - \mathrm{Id}_C\right)\Big|_H =$$

$$= \left(\frac{d}{d\varphi} \Phi(\tau_0, p_0) - \mathrm{Id}_C\right)\Big|_H$$

Therefore, for $(\lambda, h)\in\mathbf{R}\times H$

$$\frac{d}{d(\tau, h)} z(\tau_0, 0, 0, \tau_0)(\lambda, h) = \left(\frac{d}{d\varphi} \Phi(\tau_0, p_0) - \mathrm{Id}_C\right) h + \lambda \dot{p}_0. \tag{5}$$

Using the open mapping theorem, due to (3) and the remarks following it, (5) implies the desired result. Q.E.D.

In the proof of the theorem, we have used the implicit function theorem which has guaranteed the uniqueness of the implicit function. This leads to the following corollary, stating the existence of a unique periodic solution of the perturbed equation, in the autonomous case. This special case was investigated by J. K. HALE [2]. We have his theorem as a

COROLLARY. *Let us assume that the conditions of the theorem are fulfilled and, in addition, g does not depend on s and* τ. *This is the case of autonomous perturbation,* (4) *has the form*

$$\dot{x}(t) = f(x_t) + \mu g(x_t, \mu). \tag{6}$$

Then there exist positive constants μ_0, a, b *such that for any* $|\mu| < \mu_0$, (6) *has (in the orbital sense) exactly one periodic solution belonging to* $Z^p_{a,b}$, *where*

$$Z^p_{a,b} = \{\varphi\in C(\mathbf{R}, \mathbf{R}^n)|\ \varphi \text{ is } \tau\text{-periodic}, |\tau - \tau_0| < a,$$

$$\text{and there exist } \delta, t\in\mathbf{R} \text{ such that } |\varphi_\delta - p_t| < b\}.$$

PROOF. By the theorem, taking μ_0, a, b sufficiently small, (6) has at least one solution in $Z^p_{a,b}$. We have to show that every two solutions of (6), belonging to $Z^p_{a,b}$, have the same orbit, for μ_0, a, b sufficiently small.

Let x^i $(i=1, 2)$ be a solution of (6), x^i periodic with period τ_i and for suitable $\delta_i, \tau_i \in \mathbf{R}$, $(i=1, 2)$

$$x^i_{\delta_i} = p_{t_i} + \varphi^i, \quad \varphi^i \in C.$$

We may assume that $0 \leqq \delta_i < \tau_i$, $-\tau_i - \tau_0 \leqq t_i < -\tau_i$. Let $T_i = t_i + \tau_i - \delta_i$ and $y^i \in C(\mathbf{R}, \mathbf{R}^n)$ defined by $y^i(t) = x^i(t - T_i)$ $(i=1, 2)$. We have

$$y^i_{Ti} = x^i_0 = x^i_{\tau_i} = p_{T_i} + \varkappa^i, \quad \varkappa^i \in C \quad (i = 1, 2)$$

and, since $T_i < 0$, $y^i_0 = p_0 + \eta^i$, $\eta^i \in C$ $(i=1, 2)$.

According to the continuity of Ψ, taking μ_0, a, b sufficiently small, we have $\varphi^i, \varkappa^i$ and η^i sufficiently small. Applying (3) and the continuity of Ψ, there exist $\vartheta_i \in \mathbf{R}$ $(i=1, 2)$ such that $y^i_{\vartheta_i} = p_0 + h^i$ with $h^i \in H$ and $|\vartheta_i - \tau_0|$ being sufficiently small.

Define $w \in C(\mathbf{R}, \mathbf{R}^n)$ by $w(t) = y^2(t + \vartheta_2 - \vartheta_1)$. Since (6) is autonomous and y^2 is a solution, w is a solution, too. We have $w_{\vartheta_1} = y^2_{\vartheta_2}$, therefore $w_{\vartheta_1}, y^1_{\vartheta_1} \in p_0 + H$, which implies by the theorem $w_{\vartheta_1} = y^1_{\vartheta_1}$.

Acknowledgement. The author wishes to acknowledge the kind help of Professor M. Farkas in preparing the paper.

References

[1] M. FARKAS, Controllably periodic perturbations of autonomous systems, *Acta Math. Acad. Sci. Hungar.*, **22** (1971), 337—348.

[2] J. K. HALE, Solutions near simple periodic orbits of functional differential equations, *J. of Diff. Eq.*, **7** (1970), 126—138.

[3] J. K. HALE, *Functional differential equations*, Springer-Verlag (New York—Berlin—Heidelberg, 1971).

(Received February 2, 1978)

UNIVERSITY OF TECHNOLOGY
DEPARTMENT OF MATHEMATICS
1111 BUDAPEST
STOCZEK U. 2—4.

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (3—4), (1979), 321—327.

ON A PROBLEM OF G. O. H. KATONA AND T. TARJÁN

By

E. GYŐRI (Budapest) and A. V. KOSTOCHKA (Novosibirsk)

Definitions

For each real number x the lower (upper) integer of x is denoted by $[x]$ ($[x]^*$). For each graph $V(G)$ and $E(G)$ denote the vertex-set and the edge-set of G, resp. Throughout this paper graph will mean simple graph without loops and multiple edges. The complement of a graph G is denoted by $\bar{G}$. $\chi(G)$ and $\sigma(G)$ denote the chromatic number of G, and the maximal degree of vertices of G, resp.

Let G be a graph, $V_0 \subset V(G)$. Then $G(V_0)$ and $G-V_0$ is the subgraph of G spanned by V_0, and $V(G)-V_0$, resp. If $V_0=\{v\}$ then we write $G-v$ instead of $G-\{v\}$.

Let G be a graph, $v \in V(G)$. Then $d_G(v)$ or sometimes simply $d(v)$ denotes the degree of the vertex v in G.

$$N(v) = \{w \in V(G)\colon (v, w) \in E(G)\}, \quad \bar{N}(v) = V(G)-\{v\}-N(v).$$

For each natural number r, s, K_r and $K_{r,s}$ denote the complete graph of r vertices and the complete bipartite graph of $r+s$ vertices, resp.

Introduction

At the 5th Hungarian Combinatorial Colloquium G. O. H. Katona and T. Tarján set the following conjecture:

Let G be an arbitrary graph and let

$$\pi(G) = \min\left\{\sum_{i=1}^{m} |V(G_i)|\colon G_i \text{ is complete graph}, \bigcup_{i=1}^{m} E(G_i) = E(G)\right\}.$$

Then only the Turán-graph $H=K_{[n/2],[n/2]}$ has the property

$$\pi(H) = \max_{|V(G)|=n} \pi(G).$$

A similar theorem is proved by P. ERDŐS, A. W. GOODMANN and L. PÓSA in [2]: For an arbitrary graph let

$$f(G) = \min\{m\colon G_i \text{ is complete graph}, \bigcup_{i=1}^{m} E(G_i) = E(G),$$
$$E(G_i) \cap E(G_j) = \varnothing \text{ if } i \neq j\}.$$

Then only the Turán-graph $H=K_{[n/2],[n/2]}$ has the property

$$f(H) = \max_{|V(G)|=n} f(G).$$

The conjecture of Katona and Tarján was proved independently by F. R. K. Chung, by J. Kahn and by the authors.

First we give a direct proof of the conjecture then we prove a stronger theorem; namely the conjecture remains valid even if we suppose that the covering complete graphs are edge-disjoint.

Direct proof of the conjecture

LEMMA 1. (BROOKS' theorem, [1]). *For any graph* G, $\chi(G)\leq\sigma(G)+1$. *Moreover, if* $\chi(G)=\sigma(G)+1$ *and* $\sigma(G)\neq 2$ *then* $G\supset K_{\sigma(G)+1}$.

LEMMA 2. *If*

$$|V(G)|=n,\quad \sigma(G)=\left[\frac{n-1}{2}\right],\quad G\neq \overline{K}_{[n/2],[n/2]^*},$$

then there exists a vertex $v_0\in V(G)$ *such that* $d_G(v_0)=\left[\frac{n-1}{2}\right]$ *and* $G-v_0\neq$ $\neq\overline{K}_{[(n-1)/2],[(n-1)/2]}$.

PROOF. Let $v\in V(G)$ be an arbitrary vertex such that $d_G(v)=\left[\frac{n-1}{2}\right]$. If $G-v\neq\overline{K}_{[(n-1)/2],[(n-1)/2]^*}$, then we assume $v_0=v$.

Suppose that $V(G-v)=V_1\cup V_2$, $V_1\cap V_2=\emptyset$, $G(V_1)=K_{[(n-1)/2]}$, $G(V_2)=K_{[(n-1)/2]^*}$.

Case 1: n is odd. Since $G\neq\overline{K}_{(n-1)/2,(n+1)/2}$, there exist vertices $v_1\in V_1$ and $v_2\in V_2$ such that $\{(v_1,v),(v_2,v)\}\subset E(G)$. Then $G-v_1\neq K_{(n-1)/2,(n-1)/2}$ and let $v_0=v_1$.

Case 2: n is even. Since $\sigma(G)=\sigma(G(V_1))$, v is neighbouring only with vertices of V_2 in the graph G. But $|V_2|=d_G(v)$. Therefore $G=\overline{K}_{n/2,n/2}$, contradiction. Q.e.d.

The problem of Katona and Tarján can be formulated in the following way too:

Let G be an arbitrary graph and let

$$\bar{\pi}(G)=\min\Big\{\sum_{i=1}^{m}|W_i|:\ W_i\subset V(G),\ W_i \text{ is independent},$$

$$\forall(v,w)\notin E(G)\ \exists i_0:\ \{v,w\}\subset W_{i_0}\Big\}$$

(W_i is called independent if $E(G(W_i))=\emptyset$).

The only the graph $\overline{K}_{[n/2],[n/2]^*}$ has the property that

$$\bar{\pi}(K_{[n/2],[n/2]^*})=\max_{|V(G)|=n}\bar{\pi}(G)=\left[\frac{n^2}{2}\right]$$

among all graphs of n vertices.

LEMMA 3. *Let G be an arbitrary graph,* $v_0\in V(G)$. *Then*

$$\bar{\pi}(G)\leq\bar{\pi}(G-v_0)+|\overline{N}(v_0)|+\chi(G(\overline{N}(v_0))).$$

PROOF. Let $W_1, W_2, \ldots, W_k$ be an optimal system of the independent vertex-sets for $G-v_0$. Let f be a colouring of $G(\overline{N}(v_0))$ by $t=\chi(G(\overline{N}(v_0)))$ coloura and

$V_i \subset \overline{N}(v_0)$ coloured by the i-th colour $(i=1, 2, \dots, t)$. Then the system $W_1, W_2, \dots, W_k, V_1 \cup \{v_0\}, V_2 \cup \{v_0\}, \dots, V_t \cup \{v_0\}$ will be a desired system for G since

$$\sum_{i=1}^{k} |W_i| + \sum_{j=1}^{t} |V_j \cup \{v_0\}| = \bar{\pi}(G-v_0) + |\overline{N}(v_0)| + t.$$

Q.e.d.

THEOREM 1. *For an arbitrary natural number n and for an arbitrary graph G with $V(G)=n$ the following inequality holds:*

$$\bar{\pi}(G) \leqq \left[\frac{n^2}{2}\right].$$

Moreover, the equality holds if and only if $G=\overline{K}_{[n/2],[n/2]^*}$.

PROOF. For $n \leqq 3$ the theorem is trivial. Suppose that the theorem holds for every n such that $n<k$. Let G be an arbitrary graph such that $|V(G)|=k$. If $G=\overline{K}_{[n/2],[n/2]}$ then $\bar{\pi}(G)=\left[\frac{k^2}{2}\right]$. Let $G \neq \overline{K}_{[n/2],[n/2]}$ and $\sigma(G)=r$. Choose a vertex v_0 of degree r in G. It is clear that $|\overline{N}(v_0)|=k-r-1$.

Case 1: $r \geqq \frac{k}{2}$ According to Lemma 3

$$\bar{\pi}(G) \leqq \bar{\pi}(G-v_0)+2(k-r-1) \leqq \bar{\pi}(G-v_0)+k-2.$$

But by the induction hypothesis $\bar{\pi}(G-v_0) \leqq \left[\frac{(k-1)^2}{2}\right]$, consequently $\bar{\pi}(G)<\frac{k^2}{2}$.

Case 2: $r=\left[\frac{k-1}{2}\right]$. Applying Lemma 2 we may suppose that $G-v_0 \neq K_{[(k-1)/2],[(k-1)/2]^*}$. Therefore according to Lemma 3

$$\bar{\pi}(G) \leqq \bar{\pi}(G-v_0)+2\left(k-1-\left[\frac{k-1}{2}\right]\right) < \left[\frac{(k-1)^2}{2}\right]+2k-2-2\left[\frac{k-1}{2}\right] \leqq \left[\frac{k^2}{2}\right].$$

Case 3: $r<\left[\frac{k-1}{2}\right]$, k is even. Since $r<\left[\frac{k-1}{2}\right]$, thus $G-v_0 \neq K_{[(k-1)/2],[(k-1)/2]^*}$.

Consequently applying Lemma 3 and Lemma 1 we obtain that

$$\bar{\pi}(G) \leqq \bar{\pi}(G-v_0)+(k-r-1)+\chi\big(G(\overline{N}(v_0))\big) <$$
$$< \left[\frac{(k-1)^2}{2}\right]+(k-r-1)+r+1 = \left[\frac{k^2}{2}\right].$$

Case 4: Either k is odd, $r<\left[\frac{k-1}{2}\right]$, $r \neq 2$ or $k=7$, $r=2$. If $\chi\big(G(\overline{N}(v_0))\big) \leqq r$, then $\bar{\pi}(G) \leqq \bar{\pi}(G-v_0)+(k-r-1)+r<\left[\frac{(k-1)^2}{2}\right]+k-1=\left[\frac{k^2}{2}\right]$ by Lemma 3. Assume that $\chi\big(G(\overline{N}(v_0))\big)=r+1$. Then by Lemma 1 $G\big(\overline{N}(v_0)\big) \supset K_{r+1}$ (or if $k=7$, $r=2$ then $G\big(\overline{N}(v_0)\big)$ contains an odd circuit). Let v_1 be a vertex belonging to this complete graph (or if $k=7$, $r=2$ then to this odd circuit). Let us consider the vertex v_1 as v_0. If $\chi(G(\bar{N}(v_1)))\leqq r$ then the theorem is valid for G. But otherwise we obtain that there are two vertex-disjoint complete subgraphs of $r+1$ vertices in G. Let $\{v_1, v_2, \dots, v_{r+1}\}$ and $\{w_1, w_2, \dots, w_{r+1}\}$ be the vertex-sets of these complete subgraphs. According to the induction hypothesis

$$\bar{\pi}(G-\{v_1, v_2, \dots, v_{r+1}, w_1, w_2, \dots, w_{r+1}\}) \leqq \left[\frac{(k-2r-2)^2}{2}\right].$$

By Lemma 1

$$\chi(G-\{v_1, v_2, \dots, v_{r+1}, w_1, w_2, \dots, w_{r+1}\}) \leqq r+1.$$

Let $V_1, V_2, \dots, V_{r+1}$ be the disjoint independent vertex-sets covering $V(G)-$ $-\{v_1, v_2, \dots, v_{r+1}, w_1, w_2, \dots, w_{r+1}\}$. Let $D_{ij}=\{v_i, w_j\}\cup V_{i+j}$ for $i, j=1, 2, \dots,$ $r+1$ where the sum is considered mod $(r+1)$.
It is clear that

$$\bar{\pi}(G) \leqq \bar{\pi}(G-\{v_1, v_2, \dots, v_{r+1}, w_1, w_2, \dots, w_{r+1}\})+$$

$$+\sum_{i=1}^{r+1}\sum_{j=1}^{r+1}|D_{ij}| \leqq \left[\frac{(k-2r-2)^2}{2}\right]+(r+1)\sum_{j=1}^{r+1}D_{1j} =$$

$$= \left[\frac{k^2}{2}\right]-2(r+1)k+2(r+1)^2+(r+1)\big(k-2(r+1)+2(r+1)\big) =$$

$$= \left[\frac{k^2}{2}\right]-(r+1)(k-2r-2) < \left[\frac{k^2}{2}\right]$$

since $r<\left[\frac{k-1}{2}\right]$ then $r+1\leqq\left[\frac{k-1}{2}\right]$ and $(k-2r-2)>0$.

Case 5: k is odd, $k\geqq 9$, $r=2$. If G does not contain odd circuits of length greater than *3* then we may now finish the proof as in Case 4. Otherwise let $(v_1, v_2, \dots,$ $v_{2p+1}, v_1)$ be an odd circuit in G $(p\geqq 2)$. Let us consider $G-\{v_1, v_3\}$. According to the induction hypothesis $\bar{\pi}(G-\{v_1, v_3\})\leqq\left[\frac{(k-2)^2}{2}\right]$. It is obvious that

$$\chi(G-\{v_1, v_2, v_3, v_4, v_{2p+1}\}) \leqq 3.$$

Let V_1, V_2, V_3 be disjoint independent sets covering $V(G)-\{v_1, v_2, v_3, v_4, v_{2p+1}\}$. Then

$$\bar{\pi}(G) \leqq \bar{\pi}(G-\{v_1, v_3\})+\sum_{i=1}^{3}|\{v_1, v_3\}\cup V_i|+|\{v_1, v_4\}|+$$

$$+|\{v_3, v_{2p+1}\}| \leqq \left[\frac{(k-2)^2}{2}\right]+6+(k-5)+2+2 < \left[\frac{k^2}{2}\right].$$

Q.e.d.

A generalization of the conjecture

If G is a graph then let

$p(G) = \min\{\sum |V(G_i)|: E(G) = \cup E(G_i),\ G_i$'s are edge-disjoint complete graphs$\}$.

We prove the following theorem:

THEOREM 2. *Let G be an arbitrary graph such that $|V(G)|=n$. Then* $p(G) \leqq \left[\frac{n^2}{2}\right]$, *moreover if G is not the Turán-graph $K_{[n/2],[n/2]}$ then* $p(G) < \left[\frac{n^2}{2}\right]$.

The following trivial lemma will be used:

LEMMA 4. *Let G be an arbitrary graph such that $|V(G)|=n$ and such that $d(x) \geqq d$ for every vertex x of G. Then there exists a partition of $V(G)$ such that the number of the classes of the partition is $n-d$ and such that every class of the partition spans a complete graph.*

PROOF OF THE THEOREM. We proceed by induction. If $n=1, 2, 3$ then the statement is obvious. Suppose that the statement holds for every graph such that $|V(G)| \leqq n$. Let G be an arbitrary graph such that $|V(G)|=n+1$. Let x be a vertex of G such that $d(x) = \min_{y \in V(G)} d(y)$.

Case 1: The degree of every element of $N(x)$ is at least $d(x)+1$.

Case 1.1: $d(x) \leqq \frac{n}{2}$. Consider the edges incident to x as $d(x)$ classes of a partition and consider an optimal partition of $E(G-x)$. Then combining these two partitions we obtain that

$$p(G) \leqq p(G-x)+2d(x) \leqq \left[\frac{n^2}{2}\right]+n \leqq \left[\frac{(n+1)^2}{2}\right]$$

by the induction hypothesis. Equality holds if $d(x)=\frac{n}{2}$ (then n is even) and $G-x=K_{n/2,n/2}$. If there is an edge in $G(N(x))$ then we can improve the covering of the edges incident to x and $p(H) \leqq \left[\frac{n^2}{2}\right]$ also for the remaining graph H. But if there is no edge in $G(N(x))$ and $G-x=K_{n/2,n/2}$ then $G=K_{n/2+1,n/2}=K_{[(n+1)/2]^*,[(n+1)/2]}$.

Case 1.2: $d(x) > \frac{n}{2}$. The degree of every element of $N(x)$ is at least $d(x)$ in $G-x$. Every element of $N(x)$ is adjacent to at most $n-d(x)$ elements of $V(G)-N(x)-\{x\}$ because $|V(G)-N(x)-\{x\}|=n-d(x)$. That is $d_{G(N(x))}(v) \geqq 2d(x)-n$ for every vertex $v \in N(x)$. Then according to Lemma 4 there is a partition $\mathcal{P}=\{P_1\dots, P_{n-d(x)}\}$ of $N(x)$ such that $G(P_i)$ is a complete graph for $i=1, \dots, n-d(x)$. But then also $G(P_i \cup \{x\})$ is a complete graph for $i=1, \dots, n-d(x)$. Consider the edge-sets of these complete graphs as the classes of a partition and consider an optimal partition of the remaining graph H. Combining these partitions and using the induction hypothesis we obtain that

$$p(G) \leqq p(H)+\sum_{i=1}^{n-d(x)} (|P_i|+1) \leqq \left[\frac{n^2}{2}\right]+|N(x)|+n-d(x) = \left[\frac{n^2}{2}\right]+n \leqq \left[\frac{(n+1)^2}{2}\right].$$

If n is odd then $\left[\frac{n^2}{2}\right]+n<\left[\frac{(n+1)^2}{2}\right]$. Hence we may suppose that n is even and that G is not a complete graph. $\left(d(x)<n, V(G)-N(x)-\{x\}\neq\emptyset.\right)$ Then omitting $\bigcup_{i=1}^{n-d(x)} E(G(P_i\cup\{x\}))$ from $E(G)$ we obtain that the degrees of the elements of $V(G)-N(x)-\{x\}$ are at least $\frac{n}{2}+1$ in the remaining graph. Then $H\neq K_{n/2,n/2}$ and $p(H)<\left[\frac{n^2}{2}\right]$ by the induction hypothesis, i.e. $p(G)<\left[\frac{(n+1)^2}{2}\right]$ in this case.

Case: 2 There exists a vertex $y\in N(x)$ that $d(y)=d(x)$. Let y be fixed.

Case 2.1: $d(x)\leqq\frac{n+1}{2}$. Consider the $2d(x)-1$ edges incident either to x or to y as $2d(x)-1$ classes of one element. Adding these classes to an optimal partition of $E(G-\{x,y\})$ we obtain that

$$p(G)\leqq p(G-\{x,y\})+4d(x)-2\leqq\left[\frac{(n-1)^2}{2}\right]+2n=\left[\frac{(n+1)^2}{2}\right]$$

by the induction hypothesis. But if the equality is valid then $d(x)=\frac{n+1}{2}$ (i.e. n is odd) and $G-\{x,y\}=K_{(n-1)/2,(n-1)/2}$. If either $E\big(G(N(x))\big)\neq\emptyset$ or $E\big(G(N(y))\big)\neq\emptyset$ then we can improve the covering of the edges incident either to x or to y. But if $E\big(G(N(x))\big)=E\big(G(N(y))\big)=\emptyset$, $d(x)=d(y)=\frac{n+1}{2}$, $G-\{x,y\}=K_{n-1/2,n-1/2}$ then $G=K_{n+1/2,n+1/2}$.

Case 2.2: $d(x)>\frac{n+1}{2}$. Then $N(x)\cap N(y)\neq\emptyset$. Let $z_1, z_2, \dots, z_t\in N(x)\cap N(y)$ such that $G(\{z_1, z_2, \dots, z_t\})$ is a maximal complete graph in $G\big(N(x)\cap N(y)\big)$. Then the elements of $\big(N(x)-\{z_1, \dots, z_t\}\big)\cap\big(N(x)-N(y)\big)$ are not adjacent to y and the elements of $\big(N(x)-\{z_1\dots, z_t\}\big)\cap\big(N(x)\cap N(y)\big)$ are not adjacent to each elements of $\{z_1, \dots, z_t\}$. Hence the degrees of the vertices are at least $d(x)-\big(n-d(x)+t+1\big)=$ $=2d(x)-n-t-1$ in the spanned subgraph $G\big(N(x)-\{z_1, \dots, z_t\}\big)$. Similarly the degrees of the vertices are at least $2d(x)-n-t-1$ in the spanned subgraph $G\big(N(y)-\{z_1, \dots, z_t\}\big)$. Then there exists a partition $\mathcal{R}=\{R_1, \dots, R_{n-d(x)}\}$ of $N(x)-$ $-\{y, z_1, \dots, z_t\}$ such that the classes R_i of the partition span complete graphs in the graph G $\big(d(x)-t-1-(2d(x)-n-t-1)=n-d(x)\big)$. Let D be the edge-set of the complete graph spanned by $\{x, y, z_1, \dots, z_t\}$, let E_i be the edge-set of the complete graph spanned by $R_i\cup\{x\}$ for $i=1, 2, \dots, n-d(x)$. Omitting $D\cup\left(\bigcup_{i=1}^{n-d(x)} E_i\right)$, the degrees of the vertices remain at least $2d(x)-n-2t$ in the graph spanned by $N(y)-\{x, z_1, \dots, z_t\}$ by the maximality of the set $\{z_1, \dots, z_t\}$. There exists a partition $\mathcal{S}=\{S_1, \dots, S_{n+t-d(x)-1}\}$ of $N(y)-$

$-\{x, z_1, \ldots, z_t\}$ such that the classes S_j of the partition span complete graphs not intersecting

$$D \cup \left(\bigcup_{i=1}^{n-d(x)} E_i \right), \quad \big(d(x)-t-1-(2d(x)-n-2t) = n+t-d(x)-1\big).$$

Let D_j be the edge-set of the complete graph spanned by $S_j \cup \{y\}$. Let H be the graph with

$$V(H) = V(G)-\{x, y\},\ E(H) = E(G)-D-\bigcup_{i=1}^{n-d(x)} E_i - \bigcup_{j=1}^{n+t-d(x)-1} D_j.$$

Then $p(H) \leqq \left[\frac{(n-1)^2}{2}\right]$ by the induction hypothesis. Adding D, E_i's, D_j's to this partition as further classes then we obtain that

$$p(G) \leqq t+2+\sum_{i=1}^{n-d(x)} (|R_i|+1)+\sum_{j=1}^{n+t-d(x)-1} (|S_j|+1)+$$

$$+\left[\frac{(n-1)^2}{2}\right] = t+2+\big(d(x)-t-1\big)+\big(n-d(x)\big)+$$

$$+\big(d(x)-t-1\big)+\big(n+t-d(x)-1\big)+\left[\frac{(n-1)^2}{2}\right] = \left[\frac{(n-1)^2}{2}\right]+2n-1 < \left[\frac{(n+1)^2}{2}\right].$$

Q.e.d.

References

[1] R. L. Brooks, On colouring the nodes of a network, *Cambridge Philos. Soc.*, **37** (1941), 194—197.

[2] P. Erdős, A. W. Goodman and L. Pósa, The representation of a graph by set intersection, *Can. J. Math.*, **XVIII** (1966), 106—112.

[3] Fan R. K. Chung, *On the decomposition of graphs* (to appear).

[4] Jeff Kahn, *Proof of a conjecture of Katona and Tarján*, (to appear).

(Received February 3, 1978)

MATHEMATICAL INSTITUTE
OF THE HUNGARIAN ACADEMY OF SCIENCES
1053 BUDAPEST, REÁLTANODA U. 13—15.

MATHEMATICAL INSTITUTE
OF THE SOVIET ACADEMY OF SCIENCES
SIBERIAN BRANCH
NOVOSIBIRSK, USSR

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (3—4), (1979), 329—336.

WEAKLY HOMOMORPHICALLY CLOSED SEMISIMPLE CLASSES*

By

T. ANDERSON (Vancouver) and R. WIEGANDT (Budapest)

A class $\mathbf{M}$ of rings is said to be *weakly homomorphically closed*, if $\mathbf{M}$ satisfies condition

$$(*) \qquad \text{If} \quad I \lhd A \in \mathbf{M}, \quad \text{and} \quad I^2 = 0, \quad \text{then} \quad A/I \in \mathbf{M}.$$

In this note we shall deal with weakly homomorphically closed semisimple classes. As long as homomorphically closed semisimple classes are very rare (see e.g. [3]), weakly homomorphically closed semisimple classes are appropriate to characterize supernilpotent radical classes: among the hereditary radicals the supernilpotent radicals are precisely those which have weakly homomorphically closed semisimple classes (Corollary 3). The semisimple class belonging to the class of all idempotent rings is also weakly homomorphically closed (Theorem 3). At the end of this note we shall construct the smallest weakly homomorphically closed class containing a given class in two different ways.

1. Preliminaries

In what follows under a ring we always mean a not necessarily associative one. We shall work in a universal class $\mathbf{A}$ of rings, that is, in a class $\mathbf{A}$ which is homomorphically closed and hereditary ($I \lhd A \in \mathbf{A}$ implies $I \in \mathbf{A}$). A radical class and a semisimple class will always mean a *radical and a semisimple class in the sense of Kurosh and Amitsur*, respectively. For the basic notions and results of the radical theory we refer to [5] and [9]. A radical class $\mathbf{R}$ in a universal class $\mathbf{A}$ will be called *hypersolvable*, if $\mathbf{R}$ contains all zero-rings of $\mathbf{A}$. According to the usual terminology, the hereditary hypersolvable radical classes are referred as to *supernilpotent* radicals. A not necessarily hereditary radical class consisting of idempotent rings, will be called a *subidempotent radical class*.

We shall use the *upper radical* operator $\mathcal{U}$ and *semisimple operator* $\mathcal{S}$ acting on classes as

$$\mathcal{U}\mathbf{S} = \{A \in \mathbf{A} \mid A \text{ has no nonzero homomorphic image in } \mathbf{S}\}$$

and

$$\mathcal{S}\mathbf{R} = \{A \in \mathbf{A} \mid A \text{ has no nonzero ideal in } \mathbf{R}\},$$

respectively.

* The first author gratefully acknowledges the support given by the Hungarian Government and the National Research Council of Canada.

Let us recall that a semisimple class $\mathbf{S}$ may be defined by the following two conditions.

(S 1) If $A \in \mathbf{S}$, then every nonzero ideal of A has a nonzero homomorphic image in $\mathbf{S}$;

(S 2) If every nonzero ideal of a ring $A \in \mathbf{A}$ has a nonzero homomorphic image in $\mathbf{S}$, then $A \in \mathbf{S}$.

Further, if $\mathbf{R}$ is an radical class, then $\mathcal{S}\mathbf{R}$ is a semisimple class; for every semisimple class $\mathbf{S}$ the class $\mathcal{U}\mathbf{S}$ is always a radical class and $\mathbf{S}=\mathcal{S}\mathcal{U}\mathbf{S}$ holds.

In proving our results we make use of some known statements listed below.

PROPOSITION 1 (cf. [1] Corollary 24.4 or [8] Theorem 5.1). *The subdirectly irreducible abelian groups are the cyclic groups $C(p^k)$ for all prime powers p^k and the quasi-cyclic groups $C(p^\infty)$ for all primes.*

A special case of the assertion of Lemma 31.1 of [1] is

PROPOSITION 2. *The discrete direct sum H of the cyclic groups $C(q^k)$, $k=1, 2, \dots$, has a subgroup B such that $H/B \cong C(q^\infty)$.*

PROPOSITION 3 ([5] Theorem 11 or [9] Proposition 8.5). *Every semisimple class $\mathbf{S}$ is closed under extensions, that is $B \in \mathbf{S}$ and $A/B \in \mathbf{S}$ imply $A \in \mathbf{S}$.*

PROPOSITION 4 ([5] Theorem 12 or [9] Theorem 22.8). *Every semisimple class is closed under subdirect sums, that is if A is a subdirect sum of rings taken from a semisimple class $\mathbf{S}$, then also $A \in \mathbf{S}$ holds.*

PROPOSITION 5 ([7] Theorems 1 and 2). *In a universal class of associative or alternative rings a class $\mathbf{S}$ is a semisimple class if and only if $\mathbf{S}$ is hereditary, closed under subdirect sums and under extensions.*

An ideal L of a ring A is said to be *large* in A, if $L \cap I \neq 0$ holds for every ideal $I \neq 0$ of A. In view of Proposition 5, Theorem 8 of [6] can be sharpened immediately as

PROPOSITION 6. *Let $\mathbf{A}$ be a universal class of associative or alternative rings. A subclass $\mathbf{S}$ of $\mathbf{A}$ is a semisimple class and the upper radical $\mathcal{U}\mathbf{S}$ is hereditary if and only if $\mathbf{S}$ satisfies the following conditions:*

(1) *$\mathbf{S}$ is hereditary,*

(2) *$\mathbf{S}$ is closed under subdirect sums,*

(3) *If L is a large ideal in a ring $A \in \mathbf{A}$ and $L \in \mathbf{S}$, then also $A \in \mathbf{S}$ holds.*

2. Hypersolvable and subidempotent radicals

We shall restrict our attention in this section to varieties of rings which satisfy certain mild conditions on the defining identities. However, one important result remains true without all these assumptions, as will be pointed out at the appropriate place.

Throughout this section $\mathbf{A}=\mathbf{A}(\Omega)$ will denote the variety of rings satisfying a given set Ω of identities, and it is assumed that $\mathbf{A}(\Omega)$ satisfies the following two conditions

(I) **A** contains all zero-rings,
(II) If $f\in\Omega$ and g is a linearization of f, then each ring in **A** satisfies g.

We refer the reader to JACOBSON [4] for technical details about the defining identities of a variety, and shall note here that (II) is equivalent to Jacobson's condition (L), while (I) is somewhat weaker than his condition (H) (see [4] pp. 27 and 29). At any rate, *the conditions* (I) *and* (II) *are valid in the important cases of Lie, associative or alternative rings, and all rings.*

PROPOSITION 7. *If* **R** *is a hypersolvable radical class, then* $\mathcal{S}\mathbf{R}$ *is weakly homomorphically closed.*

The proof is obvious.

For a ring $A\in\mathbf{A}$ let A^n denote all sums of all products of n-elements of A. We call *A nilpotent,* if $A^n=0$ for some $n>1$. For a ring $A\in\mathbf{A}$ let us define $A^{(0)}=A$ and $A^{(n)}=(A^{(n-1)})(A^{(n-1)})$ for $n=1, 2, \ldots$. The ring A is said to be *solvable*, if there exists an $n\geqq 1$ such that $A^{(n)}=0$. The smallest such natural number n will be called the *degree of solvability*. Note that a nilpotent ring is always solvable and for associative rings the two notions coincide, for alternative rings, however, a solvable ring need not be nilpotent (see [2]).

THEOREM 1. *If* **S** *is a weakly homomorphically closed semisimple class containing a nonzero solvable ring, then* **S** *contains all zero-rings.*

PROOF. Let $A\neq 0$ be a solvable ring in **S**. Since **S** is a semisimple class, either $A^2=0$ or the ideal A^2 of A has a homomorphic image $B\neq 0$ in **S** and the degree of solvability of B is less than that of A. By induction on the degree of solvability we get that **S** contains a zero-ring $C\neq 0$. Again, since **S** is a semisimple class, the ideal (a) (that is the additive group) of C generated by a nonzero element $a\in C$, has a nonzero homomorphic image in **S**. Hence a zero-ring $Z(n)$ over the cyclic group of n elements is in **S** where n may be infinite. We claim that $Z(\infty)\in\mathbf{S}$. If n is finite, then by a similar reasoning we get $Z(p)\in\mathbf{S}$ for a prime p. In virtue of Proposition 3 it follows by induction that $Z(p^k)\in\mathbf{S}$ for every $k=1, 2, \ldots$. Moreover by Proposition 4 we get

$$Z(\infty)\cong\sum_{\text{subdirect}} Z(p^k)\in\mathbf{S}$$

in view of $Z(p^k)\cong Z(\infty)/(p^k)$. Thus $Z(\infty)\in\mathbf{S}$. For any prime q and prime power q^r the ideal (q^r) of $Z(\infty)$ satisfies $(q^r)^2\subseteqq(Z(\infty))^2=0$, so by the assumption it follows

$$Z(q^r)\cong Z(\infty)/(q^r)\in\mathbf{S}.$$

We show also that the zero ring $Z(q^\infty)$ over the quasi-cyclic group $C(q^\infty)$ is in **S**. Consider the zero-rings built over the groups H, B and $C(q^\infty)$ occurring in Proposition 2. Denoting these zero-rings again by H and B, Proposition 2 yields $H/B\cong Z(q^\infty)$. Since $H\in\mathbf{S}$ (by Proposition 4) and $B^2=0$, we conclude $Z(q^\infty)\in\mathbf{S}$. Thus $Z(p^k)\in\mathbf{S}$ holds for every prime p and $k=1, 2, \ldots, \infty$. But by Proposition 1 these zero rings are exactly the subdirectly irreducible ones. Take any zero-ring $X\in\mathbf{A}$. As is well known, X is a subdirect sum of subdirectly irreducible rings and in addition the subdirectly irreducible components are zero-rings. Since all the subdirectly irreducible zero-rings are in **S**, Proposition 4 implies $X\in\mathbf{S}$, and the statement is proved.

COROLLARY 1. *If* $\mathbf{S}$ *is a weakly homomorphically closed semisimple class containing a nonzero solvable ring, then* $\mathcal{U}\mathbf{S}$ *is a subidempotent radical class.*

REMARK. The assertions of Theorem 1 and Corollary 1 hold for any universal class $\mathbf{A}$ satisfying condition (I).

Let $A \in \mathbf{A}$ be a ring and M an A-bimodule. Define the ring

$$A * M = \{(a, m) \mid a \in A,\ m \in M\}$$

with the operation

$$(a_1, m_1) + (a_2, m_2) = (a_1 + a_2, m_1 + m_2)$$

and

$$(a_1, m_1)(a_2, m_2) = (a_1 a_2, a_1 m_2 + m_1 a_2).$$

This construction is the so-called *split-null extension* and $A * M$ is in the variety $\mathbf{A}$ (cf. [4] pp. 79—82). In particular, if A is associative (or alternative or Lie), then so is $A * M$. Obviously $A * M$ has the following properties;

(a) $M_0 = \{(0, m) \mid m \in M\} \lhd A * M$;

(b) $A' = \{(a, 0) | a \in A\}$ is a subring of $A * M$ and $A' \cong A$;

(c) $M_0^2 = 0$;

(d) $A * M / M_0 \cong A$.

LEMMA. *If* $\mathbf{R}$ *is a hereditary radical class and* $\mathcal{S}\mathbf{R}$ *contains all zero-rings, then*

$$\mathbf{R}(A * M) M_0 + M_0 \mathbf{R}(A * M) = 0 \tag{1}$$

and

$$\mathbf{R}(A * M) \subseteq A'. \tag{2}$$

PROOF. Since $\mathbf{R}$ is hereditary, we have

$$\mathbf{R}(A * M) \cap M_0 \subseteq \mathbf{R}(M_0).$$

But $M_0^2 = 0$ implies $M_0 \in \mathcal{S}\mathbf{R}$, that is $\mathbf{R}(M_0) = 0$. Hence we get

$$\mathbf{R}(A * M) M_0 + M_0 \mathbf{R}(A * M) \subseteq \mathbf{R}(A * M) \cap M_0 = 0,$$

proving the validity of (1).

To prove (2), define the map $\varphi\colon \mathbf{R}(A * M) \to M_0$ by the rule $\varphi\big((a, m)\big) = (0, m)$. Trivially φ preserves addition. Further, we have

$$\varphi\big((a_1, m_1)\big)\varphi\big((a_2, m_2)\big) = (0, m_1)(0, m_2) = (0, 0),$$

while

$$\varphi\big((a_1, m_1)(a_2, m_2)\big) = \big(\varphi(a_1 a_2, a_1 m_2 + m_1 a_2)\big) = (0, a_1 m_2 + m_1 a_2) =$$

$$= (a_1, m_1)(0, m_2) + (0, m_1)(a_2, m_2) \in \mathbf{R}(A * M) M_0 + M_0 \mathbf{R}(A * M) = 0$$

by (1). Hence φ is a ring-homomorphism of $\mathbf{R}(A * M)$ into M_0. Since $\mathbf{R}(M_0) = 0$, we have $\varphi(\mathbf{R}(A * M)) \subseteq \mathbf{R}(M_0) = 0$ which means just $\mathbf{R}(A * M) \subseteq A'$ and also (2) is proved.

THEOREM 2. *Let* **S** *be a semisimple class in a variety* **A** *such that*
(i) **S** *is weakly homomorphically closed;*
(ii) **S** *contains all zero-rings;*
(iii) $\mathbf{R}=\mathscr{U}\mathbf{S}$ *is hereditary.*
Then $\mathbf{S}=\mathbf{A}$ *and* $\mathbf{R}=\mathbf{0}$ *where* **0** *denotes the class of one-element rings.*

PROOF. Let A be any ring of **A**. Note that every ring $A\in\mathbf{A}$ is an A-bimodule (cf. [4] pp. 80—81). Apply the Lemma for the extension $A*A$. Since by (2) $\mathbf{R}(A*A)\subseteq A'$, every element of $\mathbf{R}(A*A)$ has the form $(a, 0)$. On the other hand, by (1) $\mathbf{R}(A*A)$ annihilates A_0 and so $(a, 0)(0, b)=(0, ab)=(0, 0)$ holds for every $(a, 0)\in\mathbf{R}(A*A)$ and $(0, b)\in A_0$. Thus $ab=0$ holds for every $b\in A$. In particular, by (b) $\mathbf{R}(A*A)$ is a zero-ring and by the assumption we get $\mathbf{R}(A*A)\in\mathbf{R}\cap\mathscr{S}\mathbf{R}=\mathbf{0}$. Consequently $A*A$ is in **S**. Using (d), (c) and the assumption (i), it follows $A\in\mathbf{S}$ for every $A\in\mathbf{A}$.

From Theorems 1 and 2 it follows immediately

COROLLARY 2. *Let* **A** *be a variety of rings. If* **S** *is a weakly homomorphically closed semisimple class, then the upper radical* $\mathscr{U}\mathbf{S}$ *is either hypersolvable, or a non-hereditary subidempotent radical class.*

The next corollary characterizes the supernilpotent radicals by means of weakly homomorphically closed semisimple classes.

COROLLARY 3. *Let* **A** *be a variety. A hereditary radical class* $\mathbf{R}(\neq\mathbf{0})$ *is hypersolvable if and only if the semisimple class* $\mathscr{S}\mathbf{R}$ *is weakly homomorphically closed.*

PROOF. Proposition 7 yields the necessity. If **R** is not supernilpotent, then there exists a zero-ring $A\neq 0$ which is not in **R**. Hence $0\neq A/\mathbf{R}(A)\in\mathscr{S}\mathbf{R}$ and so $\mathscr{S}\mathbf{R}$ contains a nonzero zero-ring. Now Theorems 1 and 2 establish the sufficiency.

Proposition 6 and Corollary 3 yield immediately a characterization of supernilpotent radicals by conditions imposed on their semisimple classes.

COROLLARY 4. *Let* **A** *be a variety of associative or alternative rings. A subclass* **S** *of* **A** *is weakly homomorphically closed and satisfies conditions* (1), (2) *and* (3) *of Proposition* 6 *if and only if* $\mathbf{R}=\mathscr{U}\mathbf{S}$ *is either supernilpotent or consists of one-element rings.*

COROLLARY 5. *Let* **A** *be as in Corollary* 4. *If* **Q** *is a subclass of* **A** *such that* **Q** *is weakly homomorphically closed, contains all zero-rings and satisfies conditions* (1), (2) *and* (3) *of Proposition* 6, *then* $\mathbf{Q}=\mathbf{A}$ *holds.*

We know that the upper radical of a weakly homomorphically closed semisimple class is either hypersolvable or subidempotent. The semisimple classes of hypersolvable radicals are in fact weakly homomorphically closed, and a non-hereditary subidempotent radical may have a weakly homomorphically closed semisimple class. Thus we pose the following

PROBLEM. *Describe the subidempotent radicals having weakly homomorphically closed semisimple classes (in the universal class of all associative rings).*

Next we give an example of a non-hereditary subidempotent radical with weakly homomorphically closed semisimple class. Let **A** denote the variety of all associative rings and let **Z** denote that of all zero-rings. Let us consider the upper radical $\mathbf{B}=\mathcal{U}\mathbf{Z}$. If A is an idempotent ring, then A has no nonzero homomorphic image in **Z**, hence $A\in\mathbf{B}$. If $0\neq B\in\mathbf{B}$, then we have $B/B^2\in\mathbf{Z}\cap\mathbf{B}=\mathbf{0}$, implying that B is idempotent. Hence **B** is the class of all idempotent rings of **A**. Note that **B** is not hereditary.

THEOREM 3. *The semisimple class $\mathcal{S}\mathbf{B}$ is weakly homomorphically closed and contains all zero-rings.*

PROOF. Let A be a ring from $\mathcal{S}\mathbf{B}$, and $I\lhd A$ such that $I^2=0$. We claim that $A/I\in\mathcal{S}\mathbf{B}$. Take a nonzero ideal J/I of A/I and suppose $J/I\in\mathbf{B}$, that is $(J/I)^2=J/I$. This means

$$j+I\in(J+I)(J+I)\subseteq J^2+I$$

for every $j\in J$. Hence $J\subseteq J^2+I$ holds, and so we get

$$J^2\subseteq(J^2+I)(J^2+I)\subseteq J^4+(J^2\cap I).$$

Iterating this latter relation and taking into account $I^2=0$, we get

$$J^4\subseteq\big(J^4+(J^2\cap I)\big)\big(J^4+(J^2\cap I)\big)\subseteq J^8+J^4(J^2\cap I)+(J^2\cap I)J^4\subseteq J^6.$$

Hence $J^4\subseteq J^6\subseteq J^4$ holds, and we get

$$J^8\subseteq J^4=J^6=J^2J^4=J^2J^6=J^8.$$

Thus J^4 is an idempotent ring and so $J^4\in\mathbf{B}$, for **B** contains all idempotent rings. But J^4 is an ideal of the ring $A\in\mathcal{S}\mathbf{B}$, so J^4 has a nonzero homomorphic image in $\mathcal{S}\mathbf{B}$, provided $J^4\neq0$. This is, however, impossible, so necessarily $J^4=0$ holds. Thus from $J^2\subseteq J^4+(J^2\cap I)$ we get $J^2\subseteq J^2\cap I$ and by $J\subseteq J^2+I$ we have $J\subseteq I$. Thus $J/I=0$, contradicting the choice of J/I. Hence $A/I\in\mathcal{S}\mathbf{B}$ and so $\mathcal{S}\mathbf{B}$ is weakly homomorphically closed.

The second assertion is trivial.

COROLLARY 6. *Let* **Q** *be a subclass of the class of all associative rings with the following properties:*

(i) **Q** *is hereditary;*
(ii) **Q** *is subdirectly closed;*
(iii) **Q** *is closed under extensions;*
(iv) **Q** *is weakly homomorphically closed;*
(v) **Q** *contains a nonzero nilpotent ring;*
(vi) **Q** *does not contain nonzero idempotent rings.*

Then $\mathcal{U}\mathbf{Q}$ is the class of all idempotent rings. Conversely, if **B** *is the class of all idempotent rings, then the class $\mathbf{Q}=\mathcal{S}\mathbf{B}$ satisfies conditions* (i)—(vi).

PROOF. By Proposition 5 **Q** is a semisimple class. Hence by (iv) and (v) Theorem 1 implies that **Q** contains all zero-rings, so $\mathcal{U}\mathbf{Q}$ consists of idempotent rings. But (vi) yields that every idempotent ring is in $\mathcal{U}\mathbf{Q}$. Thus $\mathcal{U}\mathbf{Q}$ is the class of all idempotent rings.

Conversely, for the class **B** of all idempotent rings $\mathbf{Q}=\mathcal{S}\mathbf{B}$ is a semisimple class, so (i), (ii) and (iii) are satisfied. In view of Theorem 3 **Q** has property (iv). Conditions (v) and (vi) are fulfilled by the definition of **Q**.

3. Weak homomorphic closures

In this section we are going to give two procedures for constructing the smallest weakly homomorphically closed class containing a given one.

Let $\mathbf{M}$ be any abstract class of rings in an arbitrary universal class $\mathbf{A}$. Define for each ordinal α a class $\mathbf{M}_\alpha$ as follows: $\mathbf{M}_1=\mathbf{M}$,

$$\mathbf{M}_\alpha = \left\{ A\in\mathbf{A} \,\middle|\, \begin{array}{l}\text{there is a ring } B\in\mathbf{M}_\beta \text{ for an ordinal } \beta<\alpha \\ \text{and an ideal } I \text{ of } B \text{ such that } B/I\cong A \text{ and } I^2=0\end{array} \right\}$$

for every $\alpha>1$. Obviously $\mathbf{M}_1\subseteq\ldots\subseteq\mathbf{M}_\alpha\subseteq\ldots$.

THEOREM 4. *Let ω denote the first limit ordinal. Then $\mathbf{M}_\omega=\mathbf{M}_{\omega+1}$ holds and $\mathbf{M}_\omega$ is the smallest weakly homomorphically closed class containing* $\mathbf{M}$.

PROOF. Let A be a ring of $\mathbf{M}_{\omega+1}$. By definition there exists a ring $B\in\mathbf{M}_\omega$ and an ideal $J\lhd B$ such that $J^2=0$ and $B/J\cong A$. Since $B\in\mathbf{M}_\omega$, there is a ring $C\in\mathbf{M}_\gamma$, $\gamma<\omega$, and an ideal $K\lhd C$ such that $K^2=0$ and $C/K\cong B$. Hence we have the isomorphism

$$A\cong B/J\cong\frac{C/K}{L/K}\cong C/L$$

with a suitable ideal $L\lhd C$, moreover, by $J\cong L/K$ we have $L^2\subseteq K$. Let P be the ideal of C generated by L^2. Now $P\subseteq K$ and $K^2=0$ imply $P^2=0$. For the ideal P we have the isomorphism

$$A\cong C/L\cong\frac{C/P}{L/P}.$$

Since $C\in\mathbf{M}_\gamma$ and $P^2=0$, we have $C/P\in\mathbf{M}_{\gamma+1}$, further $(L/P)^2=0$ and the above isomorphism shows that $C/L\in\mathbf{M}_{\gamma+2}$. Thus every ring $A\in\mathbf{M}_{\omega+1}$ is contained in a class $\mathbf{M}_\beta$ where β is a finite ordinal depending on A. Hence $\mathbf{M}_{\omega+1}\subseteq\mathbf{M}_\omega$ is proved. The opposite inclusion is trivial.

If $B\in\mathbf{M}_\omega$ and $I\lhd B$ such that $I^2=0$, then by the construction $A=B/I\in\mathbf{M}_{\omega+1}=$ $=\mathbf{M}_\omega$ holds, proving that $\mathbf{M}_\omega$ is weakly homomorphically closed.

Finally, let $\mathbf{L}$ be any weakly homomorphically closed class containing $\mathbf{M}$. Then clearly $\mathbf{M}_2\subseteq\mathbf{L}$ holds and so by induction we get $\mathbf{M}_\beta\subseteq\mathbf{L}$ for every finite β. Thus also $\mathbf{M}_\omega\subseteq\mathbf{L}$ is valid.

COROLLARY 7. *A class* $\mathbf{M}$ *is weakly homomorphically closed if and only if* $\mathbf{M}_2=\mathbf{M}$.

In what follows we assume that the universal class $\mathbf{A}$ satisfies the following condition:

If $I\lhd A\in\mathbf{A}$, then $I^2\lhd A$.

Note that the classes of associative, alternative and Lie rings do satisfy this condition, but that of Jordan rings do not.

An ideal I of a ring is called solvable, if I is a solvable ring. Let $\mathbf{M}$ be a subclass of the universal class $\mathbf{A}$ and let us define the class $\overline{\mathbf{M}}$ by

$$\overline{\mathbf{M}} = \left\{ A\in\mathbf{A} \,\middle|\, \begin{array}{l}\text{there is a ring } B\in\mathbf{M} \text{ and a solvable} \\ \text{ideal } I \text{ of } B \text{ such that } B/I\cong A\end{array} \right\}$$

THEOREM 5. *$\overline{\mathbf{M}}$ is the smallest weakly homomorphically closed class containing* $\mathbf{M}$.

PROOF. We shall exhibit $\overline{\mathbf{M}}=\mathbf{M}_\omega$. Certainly $\mathbf{M}\subseteq\overline{\mathbf{M}}$. Take a ring A from $\overline{\mathbf{M}}$ and an ideal I of A such that $I^2=0$. Now there is a ring $B\in\mathbf{M}$ and a solvable ideal K of B such that $B/K\cong A$. Hence $I\cong J/K$ with a suitable ideal J of B. Since $I^2=0$ and K is solvable, also J is solvable. Further,

$$A/I\cong\frac{B/K}{J/K}\cong B/J$$

holds. Since $B\in\mathbf{M}$ and J is solvable, we get $A/I\cong B/J\in\overline{\mathbf{M}}$. Thus $\overline{\mathbf{M}}$ is weakly homomorphically closed and by Theorem 4 $\mathbf{M}_\omega\subseteq\overline{\mathbf{M}}$ follows.

Consider a ring $A\in\overline{\mathbf{M}}$. Now there exists a ring $C\in\mathbf{M}$ and a solvable ideal L of C such that $A\cong C/L$. If $C/L^{(n)}\in\mathbf{M}_\omega$ for some $n\geqq 1, 2, \dots$, then by definition we have

$$C/L^{(n-1)}\cong\frac{C/L^{(n)}}{L^{(n-1)}/L^{(n)}}\in\mathbf{M}_\omega.$$

So by induction we get $A\cong C/L=C/L^{(0)}\in\mathbf{M}_\omega$, provided that there is an n such that $C/L^{(n)}\in\mathbf{M}_\omega$. But L is solvable, so for an appropriate n $L^{(n)}=0$ holds and therefore $C/L^{(n)}\cong C\in\mathbf{M}\subseteq\mathbf{M}_\omega$ proves the relation $\overline{\mathbf{M}}\subseteq\mathbf{M}_\omega$.

COROLLARY 8. *In the variety* $\mathbf{A}$ *of all associative rings the equality*

$$\mathbf{M}_\omega=\left\{A\in\mathbf{A}\,\middle|\,\begin{array}{l}\text{there is a ring } B\in\mathbf{A} \text{ and a nilpotent}\\ \text{ideal } K \text{ of } B \text{ such that } A\cong B/K\end{array}\right\}$$

holds.

References

[1] L. FUCHS, *Abelian groups* (Budapest, 1958).

[2] G. V. DOROFEEV, An instance of a solvable, though non-nilpotent alternative ring, *Usp. Mat. Nauk*, **15 (93)** (1960) 147—150 (Russian).

[3] B. J. GARDNER and P. N. STEWART, On semi-simple radical classes, *Bull. Austral. Math. Soc.*, **13** (1975), 349—353.

[4] N. JACOBSON, *Structure and representation of Jordan algebras*, AMS Coll. Publ. 39 (Providence, 1968).

[5] W. G. LEAVITT, *The general theory of radicals*, Univ. of Nebraska (Lincoln, 1970).

[6] L. C. A. VAN LEEUWEN, Properties of semisimple classes, *J. Nat. Sci. & Math., Lahore*, **15** (1975), 59—67.

[7] L. C. A. VAN LEEWUEN, C. ROOS and R. WIEGANDT, Characterization of semisimple classes, *J. Austral. Math. Soc.*, **23** (1977), 172—182.

[8] B. M. SCHEIN, Homomorphisms and subdirect decompositions of semigroups, *Pacific J. Math.*, **17** (966), 529—547.

[9] R. WIEGANDT, *Radical and semisimple classes of rings*, Queen's papers in pure and appl. Math., *37* Kingston, 1974).

(Received February 6, 1978)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BRITISH COLUMBIA
VANCOUVER, B.C., CANADA

AND

MATHEMATICAL INSTITUTE
OF THE HUNGARIAN ACADEMY OF SCIENCES
H—1053 BUDAPEST, REÁLTANODA U. 13—15.

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (3—4), (1979), 337—339.

REMARK ON GENERALIZED FUNCTION LATTICES

By
E. T. SCHMIDT (Budapest)

1. Introduction

G. Birkhoff has introduced the exponentation of partially ordered sets; if X, Y are partially ordered sets then Y^X denote the set of all order-preserving maps of X to Y partially ordered by $f \leqq g$ if and only if $f(x) \leqq g(x)$ for each $x \in X$. Let L be a lattice and P a partially ordered set; then L^P is a lattice, the so called function lattice. Studying the structure and decomposition of function lattices, D. DUFFUS and I. RIVAL [2] have proved the following theorem:

Let L be a finite lattice and P a finite partially ordered set with $|P|=n$. Then

$$\mathrm{Con}\,(L^P) \cong (\mathrm{Con}\,(L))^n.$$

This theorem asserts that the congruence lattice of L^P is a direct power of Con (L). For infinite P this theorem does not remain valid, e.g. if $L \cong \underset{\sim}{2}$ (where $\underset{\sim}{2}$ denotes the two element chain) and P is the chain of rationals, then Con $(\underset{\sim}{2}^P)$ is not a direct power of $\underset{\sim}{2}$.

The purpose of this paper is to give a generalization of this theorem for arbitrary partially ordered sets P. For this generalization we need the notion of the extension of a (finite) lattice by a bounded distributive lattice (see [5]) which generalizes the notion of the function lattice.

2. Totally order disconnected spaces

A subset E of a partially ordered set X is increasing if $x \in E$, $y \geqq x$ imply $y \in E$. Analogously we get the notion of a decreasing set. Let $(X, \mathcal{T}, \leqq)$ be an ordered space, i.e. a set X with a topology $\mathcal{T}$ endowed with the relation $\leqq$. Each set $\mathcal{U}$ consisting of the increasing sets in $\mathcal{T}$ and the set $\mathcal{L}$ consisting of the decreasing sets in $\mathcal{T}$ defines a topology on X. The triple $(X, \mathcal{T}, \leqq)$ is called totally order disconnected if given $x, y \in X$, $x \not\geqq y$, there exist disjoint $\mathcal{T}$-clopen sets $U \in \mathcal{U}$, $L \in \mathcal{L}$ such that $y \in U$, $x \in L$. (See CANFELL [1] or PRIESTLEY [3].)

Let D be a bounded distributive lattice and X the poset of all ultrafilters of D, i.e. $\leqq$ is the set-theoretical inclusion. $\mathcal{T}$ is the product topology induced from Hom $(D, \underset{\sim}{2})$ which is the set of all homomorphisms of D onto $\underset{\sim}{2}$ (i.e. $\mathcal{T}$ is the weak topology induced by Hom $(D, \underset{\sim}{2})$). Then $(X, \mathcal{T}, \leqq)$ is totally order disconnected. The main theorem of [3] assert that D is isomorphic to the dual lattice of $(X, \mathcal{T}, \leqq)$, i.e. to the lattice of all clopen increasing subsets.

Let L be an arbitrary lattice. $L[D]$ is the lattice of all continuous monotone maps of the totally order disconnected space X into the discrete space L. The constant mappings form a sublattice of $L[D]$ isomorphic to L. We identify L with this sublattice. If $a \in L$ then denote the corresponding diagonal element by $\bar{a}$. The reformulation of Priestley's theorem is the following; for every bounded distributive lattice D, $\underset{\sim}{2}[D] \cong D$ holds.

If D is finite, then it is easy to show that $L[D]$ is isomorphic to L^X, i.e. $L^X \cong L[\underset{\sim}{2}^X]$.

Let L be finite. If a/b is a prime quotient of L then the corresponding quotient $\bar{a}/\bar{b}$ of $L[D]$ is isomorphic to D.

We call $L[D]$ a generalized function lattice.

3. The congruence lattice of $L[D]$

The following theorem generalizes the result of Duffus and Rival.

THEOREM. *Let L be a finite lattice and D a bounded distributive lattice. Then*

$$\mathrm{Con}\,(L[D]) \cong (\mathrm{Con}\,(L))[\mathrm{Con}\,(D)].$$

Using this result, we can prove the theorem of Duffus and Rival as follows. Let L be a finite lattice and D a finite distributive lattice. P denotes the dual of the poset of all join-irreducible elements of D. Then $L[D]$ is the function lattice L^P. On the other hand, $\mathrm{Con}\,(D)$ is a finite Boolean algebra isomorphic to $\underset{\sim}{2}^n$, where $n=|P|$. Then $(\mathrm{Con}\,(L))[\mathrm{Con}\,(D)] \cong (\mathrm{Con}\,(L))^n$, hence $\mathrm{Con}\,(L^P) \cong (\mathrm{Con}\,(L))^n$.

PROOF. $L[D]$ is a subdirect power of L having the following two properties:
(i) $L[D]$ contains the constant mappings, i.e. the diagonal elements.
(ii) if a covers b in L then the quotient $\bar{a}/\bar{b}$ of $L[D]$ is isomorphic to D; we have a natural isomorphisms $\varepsilon_{ab}\colon \bar{a}/\bar{b} \to D$ which is the extension of the mappings $a \to 1$, $b \to 0$ $(0, 1 \in \underset{\sim}{2})$.

We will prove slightly more: if S is an arbitrary subdirect power of L satisfying (i) and (ii) then $\mathrm{Con}\,(S) \cong (\mathrm{Con}\,L)\,[\mathrm{Con}\,(D)]$.

Let θ be a congruence relation of S. Then θ_{ab} denotes the restriction of θ to the quotient $\bar{a}/\bar{b}$, where $a \succ b$ in L. $\bar{\theta}_{ab}$ denotes the extension of θ_{ab} to S, then $\bar{\theta}_{ab}$ is the smallest congruence relation of S which, restricted to $\bar{a}/\bar{b}$, is θ_{ab}.

If a/b runs over all prime quotients we get the family $\{\theta_{ab}\}$. We shall show that θ is uniquely determined by this family (i.e. $\theta \neq \Phi$ implies the existence of $a, b \in L$, $a \succ b$ such that $\theta_{ab} \neq \Phi_{ab}$). Let $u \equiv v(\theta)$, $u > v$, $u, v \in S$, i.e. $u=(u(i))$, $v=(v(i))$ where $u(i)$ resp. $v(i)$ are the i-th components ($i \in X$ and X is the set of all ultrafilters of D). Then $u(i) \geqq v(i)$ for all i. If $u(i) > v(i)$ for some i we choose the elements a, $b \in L$ such that $u(i) \geqq a \succ b \geqq v(i)$ (L is finite). Then $u \equiv v(\theta)$ implies $(u \wedge \bar{a}) \vee \bar{b} \equiv (v \wedge \bar{a}) \vee \bar{b}(\theta)$, i.e. $(u \wedge \bar{a}) \vee b \equiv (v \wedge \bar{a}) \vee b(\theta_{ab})$. The i-th components of these elements are a and b, hence the join of all $\bar{\theta}_{ab}$ is the congruence relation θ. We have therefore that θ is determined by the family $\{\Theta_{ab}\}$ where each θ_{ab} is a congruence relation on the suitable $\bar{a}/\bar{b} \cong D$.

Conversely, let $\{\theta^*_{ab}\}$ be a family of congruence relations $(\theta^*_{ab} \in \mathrm{Con}\,(\bar{a}/\bar{b}),\ a \succ b)$ such that $\theta(a, b) \leqq \theta(c, d)$ $(a \succ b,\ c \succ d)$ implies $\varepsilon_{ab}\theta^*_{ab} \geqq \varepsilon_{cd}\theta^*_{cd}$ in $\mathrm{Con}\,(D)$. Then it is easy to see that there exists an "extension" $\theta \in \mathrm{Con}\,(S)$ such that the restriction of θ to $\bar{a}/\bar{b}$ is θ^*_{ab}.

$\mathrm{Con}(L)$ and $\mathrm{Con}(D)$ are distributive lattices, hence by a theorem of R. QUACKENBUSH [4] $(\mathrm{Con}(L))[\mathrm{Con}(D)]$ is isomorphic to the free product $\mathrm{Con}(L) * \mathrm{Con}(D)$ in the variety of distributive lattices. The free product is commutative, therefore we get

$$(\mathrm{Con}(L))[\mathrm{Con}(D)] \cong (\mathrm{Con}(D))[\mathrm{Con}(L)].$$

But L is a finite lattice, i.e. $\mathrm{Con}(L)$ is a finite distributive lattice. Thus if Y denotes the dual of the partially ordered set of all join irreducible elements of $\mathrm{Con}(L)$ then $(\mathrm{Con}(D))[\mathrm{Con}(L)]$ is nothing else than the function lattice $\mathrm{Con}(D)^Y$.

A join-irreducible congruence relation of L has the form $\theta(a, b)$, where a covers b. This implies that we have a one-to-one correspondence between $\mathrm{Con}(S)$ and $(\mathrm{Con}(D))^Y$ which proves our theorem.

Let L be a finite simple lattice, i.e. $\mathrm{Con}(L) \cong \underset{\sim}{2}$. Then $(\mathrm{Con})(L)[\mathrm{Con}(D)] \cong$ $\cong \underset{\sim}{2}[\mathrm{Con}(D)] \cong \mathrm{Con}(D)$, thus we have

COROLLARY 1. *If L is a finite simple lattice then* $\mathrm{Con}(L[D])$ *is isomorphic to* $\mathrm{Con}(D)$.

If L is a finite modular lattice then $\mathrm{Con}(L) \cong \underset{\sim}{2}^n$. Hence we get

COROLLARY 2. *If L is a finite modular lattice then* $\mathrm{Con}(L[D]) \cong (\mathrm{Con}(D))^n$ *where n is the number of irreducible congruences of L.*

PROBLEM. Does the theorem remain valid for an infinite L?

References

[1] M. J. CANFELL, *Semi-algebras and rings of continuous functions*, Thesis, University of Edinburgh, 1968.

[2] D. DUFFUS and J. RIVAL, *Structure and decomposition results for function lattices*, Preprint of the University of Calgary.

[3] H. A. PRIESTLEY, Representation of distributive lattices by means of ordered stone spaces, *Bull. London Math. Soc.*, **2** (1970), 186—190.

[4] R. QUACKENBUSH, Free products of bounded distributive lattices, *Algebra Universalis*, **2** (1972), 393—394.

[5] E. T. SCHMIDT, Lattices generated by partial lattices, *Colloqia Math. Soc. János Bolyai*. Lattice Theory (1974), 343—353.

(Received February 10, 1978)

MATHEMATICAL INSTITUTE
OF THE HUNGARIAN ACADEMY OF SCIENCES
1053 BUDAPEST, REÁLTANODA U. 13—15.

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (3—4), (1979), 341—346.

THE 3-CENTRE AND COMMUTATIVITY THEOREMS

By

H. E. BELL* (St. Catharines)

KOVÁCS [9] defined the n-centre of the ring R to be the set

$$Z_n(R) = \{a \in R \mid P_n(a, b_1, \dots, b_{n-1}) = 0 \text{ for all } b_i \in R\},$$

where

$$P_n(X_1, X_2, \dots, X_n) = \sum_{\sigma \in S_n} (-1)^\sigma X_{\sigma(1)} X_{\sigma(2)} \dots X_{\sigma(n)}$$

is the standard polynomial of degree n. In general, constraining some subset of R to lie in $Z_n(R)$ does not imply much commutativity — indeed the Amitsur—Levitski Theorem states that if R is the ring of $k \times k$ matrices over the field F, then $Z_{2k}(R)=R$; however, $Z_3(R)$ can play a role analogous to that of the centre $Z(R)=Z_2(R)$ in the study of commutativity conditions for rings.

In this note we begin by defining a generalized 3-centre $H_3(R)$, which contains $Z_3(R)$ and is in general larger than $Z_3(R)$. We then give a very elementary treatment of some of the properties of $Z_3(R)$ and $H_3(R)$, and illustrate how they can be employed in the formulation of commutativity theorems.

To make our arguments more transparent, we display $P_3(X, Y, W)$ in terms of commutators:

$$(\dagger) \qquad P_3(X, Y, W) = X(YW-WY) - Y(XW-WX) + W(XY-YX).$$

Throughout the paper, Z will denote the centre of the ring R, N the set of nilpotent elements, and $A(S)$ the two-sided annihilator of the subset S. For ring elements x, y, the commutator $xy-yx$ will be written $[x, y]$; and $C(R)$ will stand for the commutator ideal of R. An additive subgroup H of R will be called a *Lie ideal* if $[r, x] \in H$ for all $x \in H$ and $r \in R$.

1. The polynomials K_n and the definition of $H_3(R)$

Begin by defining a certain sequence $K_1, K_2, \dots, K_n, \dots$ of integral polynomials in the countable family $\{X_1, X_2, \dots, X_n, \dots\}$ of non-commuting indeterminates, each K_i involving the first $2i+1$ indeterminates. Take $K_1(X_1, X_2, X_3)$ to be the standard polynomial $P_3(X_1, X_2, X_3)$; and having defined $K_1, \dots, K_{n-1}$, define

$$(1) \qquad K_n(X_1, \dots, X_{2n+1}) = P_3\big(K_{n-1}(X_1, \dots, X_{2n-1}), X_{2n}, X_{2n+1}\big).$$

* Supported by the National Research Council of Canada, Grant No. A3961.

For $n=1, 2, \ldots,$ define

(2) $$H_3^{(n)}(R) = \{x \in R \mid K_n(x, x_2, \ldots, x_{2n+1}) = 0 \text{ for all } x_2, \ldots, x_{2n+1} \in R\};$$

and define

(3) $$H_3(R) = \bigcup_{n=1}^{\infty} H_3^{(n)}(R).$$

Clearly $H_3^{(1)}(R) = Z_3(R)$, so $Z_3(R) \subseteq H_3(R)$.

Kovacs has introduced a family of generalized Jacobi identities. Specifically, if $Q(X_1, \ldots, X_n)$ is an integral polynomial in n non-commuting indeterminates, and X_{n+1} is an additional indeterminate, Q will be said to satisfy the n-th Jacobi identity if

(4) $$Q([X_{n+1}, X_1], X_2, \ldots, X_n) + Q(X_1, [X_{n+1}, X_2], \ldots, X_n) + \ldots$$
$$\ldots + Q(X_1, \ldots, [X_{n+1}, X_n]) + [Q(X_1, \ldots, X_n), X_{n+1}] \equiv 0,$$

the notation $\equiv$ denoting identity of polynomials.

It is proved in [9], by a careful consideration of permutations, that the standard polynomial S_n satisfies the n-th Jacobi identity for $n=2, 3, \ldots$. The case of interest to us is that of $n=3$, which is easy to verify directly; using this as a first step in a straightforward but tedious induction, we can show that for each $n=1, 2, \ldots,$ the polynomial K_n defined above satisfies the $2n+1$-st Jacobi identity.

2. Properties of $Z_3(R)$ and $H_3(R)$

Our first theorem, listed first because of the elementary nature of its proof, indicates a property that $H_3(R)$ has in common with Z.

THEOREM 1. *If R is any ring with $N \subseteq H_3(R)$, then N is an ideal of R.*

PROOF. An easy inductive argument shows that for each positive integer n and each $x, r_1, r_2, \ldots, r_{2n-1}, r_{2n} \in R$, we have

$$K_n(x, xr_1, r_2x, xr_3, r_4x, \ldots, xr_{2n-1}, r_{2n}x) = \pm xr_{2n-1}xr_{2n-3} \ldots xr_1xr_2xr_4 \ldots xr_{2n}x + p,$$

where p is a sum of products of ring elements, each product having x^2 as a factor. Thus, if $x \in H_3(R)$, there exists an integer $n=n(x)$ such that for each $r_1, r_2, \ldots, r_{2n} \in R$,

(5) $$xr_1xr_2 \ldots xr_{2n}x = sx^2 + x^2t + u,$$

where u is a sum of terms of form vx^2w.

We now assume that $N \subseteq H_3(R)$, and argue by induction on the degree of nilpotence that for each $x \in N$, there exists $j=j(x)$ such that any product of ring elements having at least j factors equal to x is trivial. Suppose first that $x^2=0$; clearly any product with two adjacent factors equal to x is 0, and from (5) any other product with at least $2n(x)+1$ factors equal to x is 0. Now assume our result obtains for any nilpotent element in $H_3(R)$ having index of nilpotence less than

$t, t \geqq 3$, and consider $x \in H_3(R)$ with $x^t = 0$. From (5), any product having at least $2n(x)+1$ factors equal to x may be written as a sum of products, each with x^2 as a factor; and our inductive hypothesis, together with the observation that x^2 has index less than t, can be invoked to show that any product with at least $j(x^2)(2n(x)+1)$ factors equal to x is trivial. This completes our induction.

It is immediate that for arbitrary $x \in N$ and $r \in R$, both rx and xr are in N. Furthermore, if $x, y \in N$ and $k = j(x)+j(y)-1$, $(x-y)^k$ is a sum of products each of which has at least $j(x)$ factors equal to x or at least $j(y)$ factors equal to y; hence $x-y \in N$ and we have completed the proof that N is an ideal.

The next theorem is in part an extension to $H_3(R)$ of some results in [9].

THEOREM 2. (i) *For any ring R, $H_3(R)$ is a Lie ideal.*

(ii) *If R is a prime ring with non-zero centre, $Z_3(R) = Z \cap A(C(R))$.*

(iii) *If R is an arbitrary ring with 1, $Z_3(R) = Z \cap A(C(R))$.*

(iv) *If R is any prime ring, either R is commutative or $H_3(R) = \{0\}$.*

PROOF. (i) Let $x \in H_3(R)$ and $r \in R$. Then for some n, $x \in H_3^{(n)}(R)$ — i.e. $K_n(x, y_1, \dots, y_{2n}) = 0$ for all $y_1, y_2, \dots, y_{2n} \in R$. Since K_n satisfies the $2n+1$-st Jacobi identity, it follows from (4) that $K_n([r, x], y_1, \dots, y_{2n}) = 0$ for all $y_1, \dots, y_{2n} \in R$ and hence that $[r, x] \in H_3(R)$.

(ii) Substitute $a \in Z_3(R)$, $0 \neq z \in Z$, and $r \in R$ for X, Y, W in (†), thereby obtaining $0 = -z(ar - ra)$. Since the centre of R contains no nontrivial zero divisors of R, we get $a \in Z$. Using this fact and substituting in (†) the elements $a \in Z_3(R)$ and $r, s \in R$, we get $a(rs - sr) = 0$. Thus $Z_3(R) \subseteqq Z \cap A(C(R))$. The reverse inclusion is handled similarly.

(iii) The proof is omitted; it is similar to that of (ii).

(iv) Let R be prime and suppose that $x^2 = 0$ for $x \in H_3(R)$. By (5), the right ideal xR is nilpotent, hence $x = 0$; thus $H_3(R)$ contains no non-zero elements squaring to 0. We can show more — namely, that $H_3(R)$ contains no non-trivial zero divisors of R. For if $x \in H_3(R)$, and if y is an element such that $xy = 0$, (i) shows that for each $r \in R$, $[x, yr]$ is an element of $H_3(R)$ squaring to zero, hence is trivial. Thus, $yrx = 0$ and the primeness of R forces $y = 0$.

Now consider $Z_3(R) = H_3^{(1)}$, and suppose it contains a non-zero element x. Then $P_3(x, y, x^2) = 0$ for all $y \in R$, and by (†), $xyx^2 - x^2yx = 0$ for all $y \in R$. Since x is not a zero divisor, cancelling it from this equation shows that $x \in Z$. It follows from (ii) that $Z_3(R)C(R) = 0$ and therefore R must be commutative. We have now shown that if $H_3^{(1)}(R) \neq 0$, R is commutative.

Using this result to launch an inductive argument, assume that $H_3^{(n-1)}(R) \neq 0$ implies R commutative. Suppose $0 \neq x \in H_3^{(n)}(R)$. Then $P_3(K_{n-1}(x, y_1, \dots, y_{2n-2}), y_{2n-1}, y_{2n}) = 0$ for all $y_1, y_2, \dots, y_{2n} \in R$, so that either $K_{n-1}(x, y_1, \dots, y_{2n-2}) = 0$ for all $y_1, \dots, y_{2n-2} \in R$ (and hence $x \in H_3^{(n-1)}(R)$), or there exist $y_1, \dots, y_{2n-2} \in R$ for which $K_{n-1}(x, y_1, \dots, y_{2n-2})$ is a non-zero element of $H_3^{(1)}(R)$. In either event R is commutative and the proof of (iv) is complete.

We intend to prove commutativity theorems for rings with some subset required to lie in $H_3(R)$. To use $Z_3(R)$ instead of $H_3(R)$ would not yield many new results for rings with 1, since $Z_3(R)$ is contained in the ordinary centre. (Kovacs gives an example, however, showing that in the absence of 1, $Z_3(R)$ need not lie in Z.) The following example shows that even for rings with 1, $H_3(R)$ need not lie in Z, and incidentally, establishes our claim that $H_3(R)$ is in general larger than $Z_3(R)$.

EXAMPLE. Let R be any non-commutative ring with 1 having $C(R)\subseteq Z$. (For example, such a ring may be obtained by taking a non-commutative ring S with $S^3=0$ and adjoining an identity by the standard Dorroh method.) Now for $x, y, z, w\in R$ we have $(xy-yx)zw=w(xy-yx)z=wz(xy-yx)$, so $[x, y][z, w]=0$ and hence $C(R)^2=0$. Thus, in view of (iii), $C(R)\subseteq Z\cap A(C(R))=Z_3(R)$. By (†), $P_3(x, y, z)\in C(R)$ for all $x, y, z\in R$; so we get $P_3(P_3(x, y, z), u, v)=0$ for all $x, y, z, u, v\in R$. Thus $H_3^{(2)}(R)=H_3(R)=R$.

3. Some commutativity theorems

A long-standing result of HERSTEIN [7] is that periodic rings with central nilpotent elements are commutative. Our Theorem 1 yields the following extension.

THEOREM 3. *Let R be any periodic ring with $N\subseteq H_3(R)$. Then $C(R)$ is nil, and N is an ideal.*

PROOF. By Theorem 1, N is an ideal. The factor ring $\bar{R}=R/N$ is then periodic with no non-zero nilpotent elements, hence has the property that for each $x\in\bar{R}$, $x^{n(x)}=x$ for some $n(x)>1$. (For proof, see [3].) Thus, $\bar{R}$ is commutative by Jacobson's "$x^n=x$ theorem", and $C(R)\subseteq N$.

There are many theorems in the literature asserting either that a ring R is commutative or that $C(R)$ is nil, under hypotheses of the following form: there exists an integer $n\geqq 1$ such that for each ordered n-tuple $(x_1, \dots, x_n)$ of ring elements, there is a polynomial p of a certain prescribed form, in n non-commuting indeterminates, for which $p(x_1, \dots, x_n)=0$. Often such a theorem has an extension, which we shall call the *classical extension*, asserting that if $p(x_1, \dots, x_n)\in Z$, then $C(R)$ is nil; this extension can on occasion be rather difficult to establish. In view of result (iv), there is always an immediate extension of the form "if $p(x_1, \dots, x_n)\in$ $\in H_3(R)$, then $C(R)$ is nil". As examples, we give such extensions for three recent commutativity theorems. In the statements, the parts in parentheses, as well as the bibliographic references following the theorem numbers, describe the original theorems, which we refer to as Theorems 4—0, 5—0, and 6—0.

THEOREM 4. [2] *Let R be a ring such that for each $x, y\in R$, there exist positive integers $n=n(x, y)$ and $m=m(x, y)$ for which $xy-y^m x^n\in H_3(R)$ $(xy-y^m x^n=0)$. Then $C(R)$ is nil (R is commutative).*

THEOREM 5. [8] *Let R have the property that for each $x, y\in R$, there exist positive integers $n=n(x, y)$ and $m=m(x, y)$ such that $[x^n, y^m]\in H_3(R)$ $([x^n, y^m]=0)$. Then $C(R)$ is nil.*

THEOREM 6. [1] *Suppose $q(X)$ is a polynomial in n non-commuting indeterminates, its coefficients being integers with highest common factor 1; and let R be a ring with the property that for each $x_1, \dots, x_n\in R$, $q(x_1, \dots, x_n)\in H_3(R)$ $(q(x_1, \dots, x_n)=0)$. If there exists no prime p for which the ring of 2×2 matrices over $GF(p)$ satisfies the identity $q(X)=0$, then R has nil commutator ideal.*

PROOF OF THEOREM 4. Since the intersection of the prime ideals of R is a nil ideal, it suffices to show that for each prime ideal P of R, the prime ring $\bar{R}=R/P$

is commutative. By (iv), either $\bar{R}$ is commutative or satisfies the condition $xy=y^m x^n$; and in the latter case, $\bar{R}$ is commutative by Theorem 4—0.

PROOF OF THEOREMS 5 AND 6. We require a more sophisticated use of the prime ideal structure theory. It is clearly sufficient to establish commutativity of rings R satisfying our hypotheses and having no non-trivial nil ideals. Such rings are subdirect sums of prime rings without nil ideals (see [5, Theorem 65]), and these are commutative by (iv), together with Theorems 5—0 and 6—0.

These examples were chosen in part because the classical extensions (those in which Z replaces $H_3(R)$) are problematical. It is not known whether the classical extensions of Theorems 4—0 and 5—0 are true; if they are, their proofs can be expected to be difficult. The classical extension of Theorem 6—0 is definitely false, for the polynomial $(XY-YX)^2$ is a nonvanishing central polynomial for the ring of 2×2 matrices over every field $GF(p)$ [6].

As the example at the end of Section 2 shows, constraining certain subsets of R to lie in $H_3(R)$ or $Z_3(R)$ cannot be expected to yield full commutativity; however, in special cases, it may yield a slightly stronger conclusion than the one that $C(R)$ is nil. Our concluding theorem illustrates this fact.

THEOREM 7. *Let R be a finite ring such that*
(a) $xy=0$ *implies* $yx=0$, *and*
(b) $N\subseteqq Z_3(R)$.
Then R is a direct sum of a commutative ring and a nil ring.

PROOF. Use induction on the number $|R|$ of elements of R. There is no trouble beginning — all rings of order 2 are commutative; thus, suppose $|R|=n$ and that all rings S satisfying (a) and (b) and having $|S|<n$ may be written as direct sums of commutative rings and nil rings. If R has 1, the nilpotent elements are central by (iii) of Theorem 2, and R is commutative by the theorem of [7]; and if R is nil, there is nothing to prove. Thus, we can assume that R contains a non-zero idempotent e which is not a multiplicative identity element. Now for arbitrary $x\in R$, we have $(ex-exe)e=0$, so $e(ex-exe)=ex-exe=0$; similarly, $xe-exe=0$ and e is therefore central. It follows that R has a direct sum decomposition $R=eR\oplus A(e)$; and we get our result by applying the inductive hypothesis to eR and $A(e)$.

Condition (a) by itself will not yield the conclusion of Theorem 7, for there exist finite non-commutative rings in which the product of any two zero divisors is zero [4]; similarly, condition (b) alone is not sufficient, as we see by noting that for any field F, the ring

$$R=\left\{\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix} \middle| \, a,b\in F\right\}$$

has $Z_3(R)=R$ and $Z=\{0\}$.

Finally, we remark that it is not clear whether $H_3(R)$ is a subring in general. However, if $H_3(R)$ is replaced by the subring H which it generates, conclusion (iv) of Theorem 2 still holds and we can use H in applications of the kind in Theorems 4—6.

References

[1] H. E. BELL, On some commutativity theorems of Herstein, *Arch. Math.*, **24** (1973), 34—38.
[2] H. E. BELL, A commutativity condition for rings, *Canad. J. Math.*, **28** (1976), 986—991.
[3] H. E. BELL, Some commutativity results for periodic rings, *Acta Math. Acad. Sci. Hungar.*, **28** (1976), 279—283.
[4] B. CORBAS, Finite rings in which the product of any two zero divisors is zero, *Arch. Math.*, **21** (1970), 466—469.
[5] N. DIVINSKY, *Rings and radicals*, (University of Toronto Press, 1965).
[6] E. FORMANEK, Central polynomials for matrix rings, *J. Algebra*, **23** (1972), 129—132.
[7] I. N. HERSTEIN, A note on rings with central nilpotent elements, *Proc. Amer. Math. Soc.*, **5** (1954), 620.
[8] I. N. HERSTEIN, A commutativity theorem, *J. Algebra*, **38** (1976), 112—118.
[9] A. KOVÁCS, The n-center of a ring, *J. Algebra*, **40** (1976), 107—124.

(Received February 13, 1978)

DEPARTMENT OF MATHEMATICS
BROCK UNIVERSITY
ST. CATHARINES, ONTARIO
CANADA L2S 3A1

Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 34 (3—4), (1979), 347—350.

ON A COMMON GENERALIZATION OF BORSUK'S AND RADON'S THEOREM

By

E. G. BAJMÓCZY and I. BÁRÁNY (Budapest)

1. The well-known theorem of RADON [3] says that if $A \subset R^d$ and $|A| \geqq d+2$, then there exist $B, C \subset A$, $B \cap C = \emptyset$ such that $\operatorname{conv} B \cap \operatorname{conv} C$ is not empty. It is clear that for each finite set $A = \{a_1, \dots, a_n\}$ in R^d with $n \geqq d+2$ one can find a linear map $f\colon R^{d+1} \to R^d$ and a set $A' = \{a_1', \dots, a_n'\} \subset R^{d+1}$ such that $f(a_i') = a_i$, $i = 1, 2, \dots, n$ and $\operatorname{int}\operatorname{conv} A'$ is not empty and $\operatorname{vert}\operatorname{conv} A' = A'$. In view of this fact, Radon's theorem can be stated in the following way.

RADON'S THEOREM. *Let $P \subset R^{d+1}$ be a convex polytope with non-empty interior. Put $A = \operatorname{vert} P$. If $f\colon R^{d+1} \to R^d$ is a linear map, then there exist two disjoint sets $B, C \subset A$ such that $f(\operatorname{conv} B) \cap f(\operatorname{conv} C)$ is non-empty.*

The surprising fact here is that the word "linear" can be replaced by "continuous", namely, a continuous analogue of Radon's theorem is true;

THEOREM 1. *Let $P \subset R^{d+1}$ be a convex polytope with non-empty interior. Given an $f\colon \partial P \to R^d$ continuous map, there exist two disjoint faces, B and C, of P such that $f(B) \cap f(C) \neq \emptyset$.*

COROLLARY. *Let T be a $(d+1)$-dimensional simplex. Denote its d-faces by $L_1, L_2, \dots, L_{d+2}$. If $f\colon \partial T \to R^d$ is a continuous map, then $\bigcap_{i=1}^{d+2} f(L_i)$ is non-empty.*

If f is a linear map, then this statement is an easy consequence (in fact, equivalent) of Helly's theorem (see [3]). The interesting fact here is that in this particular case a continuous version of Helly's theorem holds true.

Let us now introduce some notions. Given a convex compact set $C \subset R^{d+1}$ with non-empty interior and a vector $a \in R^{d+1}$, $a \neq 0$, we write

$$C(a) = \{x \in C\colon \langle a, x \rangle = \max_{t \in C} \langle a, t \rangle\}.$$

Two points, x and y, of C are said to be opposite if for some $a \in R^{d+1}$, $x \in C(a)$ and $y \in C(-a)$. If C happens to be a polytope, then $C(a)$ is a proper face of C. In this case we say that the two faces $C(a)$ and $C(-a)$ are opposite.

THEOREM 2. *Given a polytope $P \subset R^{d+1}$ with non-empty interior and a continuous map $f\colon \partial P \to R^d$, there exist two opposite faces, B and C, of P such that $f(B) \cap f(C)$ is non-empty.*

It is evident that opposite faces of P are disjoint. Thus Theorem 2 implies Theorem 1.

Speaking about points instead of faces Theorem 2 can be formulated as follows.

THEOREM 2′. *Given a polytope* $P \subset R^{d+1}$ *with non-empty interior and a continuous map* $f: \partial P \to R^d$, *there exist two opposite points, x and y, of P with* $f(x)=f(y)$.

We shall prove this Theorem 2′ which yields a generalization of Borsuk's theorem [1]. In order to state Borsuk's theorem put $S^d = \{x \in R^{d+1}: \|x\|=1\}$.

BORSUK'S THEOREM. *If* $f: S^d \to R^d$ *is a continuous map, then there is a point* $x \in S^d$ *with* $f(x)=f(-x)$.

THEOREM 3. *Let* $C \subset R^{d+1}$ *be a convex compact set with nonempty interior. If* $f: \partial C \to R^d$ *is a continuous map, then there exist two opposite points, x and y, of C with* $f(x)=f(y)$.

Again, Theorem 3 implies Theorem 2′. However, we shall get Theorem 3 from Theorem 2′ by a simple continuity argument.

Further, our Theorem 3 contains Borsuk's theorem (put simply $C = \text{conv } S^d$). On the other hand, Theorem 2′ is proved using Borsuk's theorem.

2. We need a simple proposition.

PROPOSITION. *If P is a polytope in* R^d *and* $x, y, x_n \in P$ $n=1, 2, \ldots$ *and* $\lim x_n = x$, *then there is an* $\varepsilon > 0$ *and N such that* $x_n + \varepsilon \cdot (y-x) \in P$ *for* $n > N$.

PROOF. This proposition is true for any cone C (instead of P) whose vertex is x (with arbitrary $\varepsilon > 0$ and n), so it is true for $C \cap B(x, \delta)$ where $B(x, \delta)$ is the ball with center x and radius δ. But $P \cap B(x, \delta) = C \cap B(x, \delta)$ for a sufficiently small $\delta > 0$ where

$$C = \{z \in R^d: z = x + \lambda(w-x), \ \lambda > 0, \ w \in P\}$$

is a cone with vertex x.

PROOF OF THEOREM 2′. Put $Q = P - P$. Q is a polytope with non-empty interior. It is centrally symmetric with respect to the origin. For $x \in Q$ write

$$h(x) = \max \{z: \ x = z - w, \ z, w \in P\}$$

where max is meant in the lexicographic ordering of R^{d+1}. Clearly $h: Q \to P$ is well-defined. An easy computation shows that the vector w corresponding to $z = h(x)$ equals $h(-x)$.

We claim that h is continuous. Indeed, let $x, x_n \in Q$, $x = \lim x_n$ and $x_n = z_n - w_n$ where $z_n = h(x_n)$. We can choose a subsequence n_i so that z_{n_i} and, consequently w_{n_i} converge. Put $z = \lim z_{n_i}$ and $w = \lim w_{n_i}$; clearly $x = z - w$. We claim that $z = h(x)$. If not, then $z < h(x)$ in the lexicographic ordering. By the Proposition, for a sufficiently small positive ε and large i

$$z' = z_{n_i} + \varepsilon\big(h(x) - z\big) \in P \quad \text{and} \quad w' = w_{n_i} + \varepsilon\big(h(-x) - w\big) \in P.$$

Now $z' - w' = x_{n_i}$ and $z' > z_{n_i}$ contradicting $z_{n_i} = h(x_{n_i})$. This means that $z = h(x)$. Thus, every convergent subsequence of z_n tends to $h(x)$. Now by compactness $\lim z_n = h(x)$, i.e., h is continuous.

Next we claim that $x\in Q(a)$ implies $h(x)\in P(a)$ and $h(-x)\in P(-a)$. Indeed, if $x\in Q(a)$ then $\max_{t\in Q}\langle a, t\rangle=\langle a, x\rangle$. Of course, $x=h(x)-h(-x)$ and $h(x), h(-x)\in P$. Whence

$$\langle a, h(x)\rangle+\langle -a, h(-x)\rangle = \langle a, x\rangle = \max_{t\in Q}\langle a, t\rangle =$$

$$= \max_{u,v\in P}\langle a, u-v\rangle = \max_{u\in P}\langle a, u\rangle+\max_{v\in P}\langle -a, v\rangle$$

and so $h(x)\in P(a)$ and $h(-x)\in P(-a)$. This further implies that for $x\in\partial Q$ $h(x)$ and $h(-x)$ belong to ∂P.

Now we define a map $g\colon \partial Q\to R^d$ in the following way: for $x\in\partial Q$ let $g(x)=f(h(x))$. This map is welldefined and continuous. Let us observe now that the conditions of Borsuk's theorem are fulfilled for the map g (instead of S^d we have ∂Q here but this is indifferent). In this case Borsuk's theorem says that there is a point $x\in\partial Q$ with $g(x)=g(-x)$. Now there exists $a\in R^{d+1}$, $a\neq 0$ such that $x\in Q(a)$. Then $z=h(x)\in P(a)$ and $w=h(-x)\in P(-a)$, i.e., z and w are opposite points of P and $f(z)=f(h(x))=g(x)=g(-x)=f(h(-x))=f(w)$. And this is what we wanted to prove.

Proof of the Corollary. It is easy to check that if B and C are disjoint faces of the simplex T, then for any $i=1, 2, \ldots, d+2$ either $B\subset L_i$ or $C\subset L_i$ (or both). This fact proves the Corollary.

Proof of Theorem 3. Without loss of generality we may suppose that $0\in \operatorname{int} C$.

Now let P be a polytope inscribed in C, i.e., vert $P\subset\partial C$ and suppose further that $0\in\operatorname{int} P$. Then a continuous map $f_P\colon \partial P\to R^d$ can be defined as $f_P(x)=f(\lambda x)$, where λ is the unique positive number with $\lambda x\in\partial C$. By Theorem 2′, there are opposite points of P, z_P and w_P with $f_P(z_P)=f_P(w_P)$.

Now choose a sequence of inscribed polytopes $P_1, P_2, \ldots$ with $0\in\operatorname{int} P_n$. Suppose further that vert $P_n\subset$ vert P_{n+1} and $\partial C\cap\bigcup_1^\infty P_n$ is dense in ∂C. Again, for each n there exist opposite (for P_n) points z_n and w_n with $f_n(z_n)=f_n(w_n)$ where $f_n=f_{P_n}$. Since z_n and w_n are opposite points in P_n there is a vector $a_n\in S^d$ such that $z_n\in P_n(a_n)$ and $w_n\in P_n(-a_n)$.

By the compactness of C and S^d we may suppose that z_n, w_n and a_n converge, their limits are $z, w\in\partial C$ and $a\in S^d$ respectively. It is easy to see that z and w are opposite points of C (with normal a) and $f(z)=f(w)$.

3. Remarks. 1. Theorem 1 can be interpreted in the following way. Let $P\subset R^{d+1}$ be a convex polytope with non-empty interior. Then it is not possible to make a drawing of ∂P in R^d so that disjoint faces of P be disjoint in the drawing.

2. We can give a second proof of Theorem 2 which is more involved than the above one but does not make use of Borsuk's theorem. It relies on a suitably modified version of the main lemma of [2].

3. The following generalization of Theorem 3 holds true.

Theorem 4. *Let* $C\subset R^{d+1}$ *be a convex compact set with non-empty interior. Let* f *be a point to set map from* ∂C *to the family of all compact convex subsets of a compact set of* R^d. *If* f *is upper semi-continuous (i.e.,* $x_n\to x$, $y_n\in f(x_n)$, *and* $y_n\to y$

implies $y \in f(x)$), *then there exist two opposite points, z and w, of C with* $f(z) \cap f(w) \neq \emptyset$.

This theorem follows from Theorem 3 nearly the same way as Kakutani's fixed-point theorem follows from Brouwer's one.

4. We conclude with a conjecture. There is a generalization of Radon's theorem which is due to H. TVERBERG [5]. In the spirit of our formulation of Radon's theorem this generalization runs as follows:

THEOREM. *Let* $P \subset R^n$ *be a convex polytope with non-empty interior. Here* $n=(r-1)(d+1)$. *Given an* $f\colon R^n \to R^d$ *linear map there are disjoint proper faces* $B_1, B_2, \dots, B_r$ *of P, such that* $\bigcap_{i=1}^{r} f(B_i)$ *is non-empty.*

We think (but can neither prove nor disprove) that in this theorem it is enough to assume that $f\colon \partial P \to R^d$ is a continuous map.

ACKNOWLEDGEMENT. We are indebted to Prof. M. Bognár for his valuable comments on an earlier version of this paper and to L. Lovász for the question that led us from Theorem 2′ to Theorem 3.

References

[1] P. ALEXANDROFF, H. HOPF, *Topologie*, Springer (Berlin, 1935).
[2] I. BÁRÁNY, *An algorithm to compute fixed points* (unpublished).
[3] B. GRÜNBAUM, *Convex Polytopes*, Interscience (London, 1967).
[4] S. KAKUTANI, A generalization of Brouwer's fixed-point theorem, *Duke Math. J.*, **8** (1941), 457—458.
[5] H. TVERBERG, A generalization of Radon's theorem, *J. London Math. Soc.*, **41** (1966), 123—128.

(Received February 17, 1978)

OTSZK, COORDINATION AND SCIENTIFIC SECRETARIAT
1132 BUDAPEST, VICTOR HUGO U. 18—22.

INDEX

Printed in Hungary

A kiadásért felel az Akadémiai kiadó igazgatója Műszaki szerkesztő: Zacsik Annamária
A kézirat nyomdába érkezett: 1979. I. 16. — Terjedelem: 11,55 (A/5) ív 13 ábra

79-245—Szegedi Nyomda — F.v.: Dobó József igazgató

The Acta Mathematica publishes papers on mathematics in English, German, Frencn and Russian. It appears in parts of various size, making up volumes. Manuscripts should be sent to:

Acta Mathematica, H-1053 Budapest, Reáltanoda u. 13—15.

Correspondence with the publishers should be sent to:

Acta Mathematica, H-1363 Budapest, Pf. 24.

The rate of subscription is $36.00 a volume. Orders may be placed with "*Kultura*" Foreign Trade Company (1011 Budapest, Fő u. 32. Account No. 218-10990) or with representatives abroad.

Instructions for authors. Manuscripts should be typed on standard size papers (25 rows; 50 characters in each row). Mathematical symbols (except sin, log, etc.) will be set in italics, so *use handwritten letters, or indicate them by simple underlining.* The text of theorems, lemmas, corollaries are to be printed also in italics. The words THEOREM, LEMMA, COROLLARY, PROOF, REMARK, PROBLEM, etc. (or their German and French equivalents) should be typed in capital letters. Sub-headings will be set in boldface lower case; indicate them by red underlining. Other special types of symbols (such as German, Greek, script, etc.) should also be clearly identified. Mark footnotes by consecutive superscript numbers. When listing references, please follow the following pattern:

[1] G. Szegő, *Orthogonal polynomials,* AMS Coll. Publ. Vol. XXIII (Providence, 1939).
[2] A. Zygmund, Smooth functions, *Duke Math. J.,* **12** (1945), 47—76.

For abbreviation of names of journals follow the Mathematical Reviews. After the references give the author's affiliation.

Do not send abstract and second copy of the manuscript. Authors will receive only galley-proofs (one copy). Manuscripts will not be sent back to authors (neither for the purpose of proof-reading nor when rejecting a paper).

Authors obtain 100 reprints free of charge. Additional copies may be ordered from the publishers.

All the reviews of the Hungarian Academy of Sciences may be obtained among others from the following bookshops:

ALBANIA

Ndermarja Shtetnore e Botimeve
Tirana

AUSTRALIA

A. Keesing
Box 4886, GPO
Sidney

AUSTRIA

Globus Buchvertrieb
Salzgries 16
Wien I.

BELGIUM

Office International de Librairie
30, Avenue Marnix
Bruxelles 5
Du Monde Entier
5, Place St. Jean
Bruxelles

BULGARIA

Raznoiznos
1 Tzar Assen
Sofia

CANADA

Pannonia Books
2 Spadina Road
Toronto 4, Ont.

CHINA

Waiwen Shudian
Peking
P.O.B. Nr. 88.

CZECHOSLOVAKIA

Artia A. G.
Ve Smeckách 30
Praha II.
Postova Novinova Sluzba
Dovoz tisku
Vinohradska 46
Praha 2
Postova Novinova Sluzba
Dovoz tlace
Leningradska 14
Bratislava

DENMARK

Ejnar Munksgaard
Nörregade 6
Kopenhagen

FINLAND

Akateeminen Kirjakauppa
Keskuskatu 2
Helsinki

FRANCE

Office International de Documentation et Libraire
48, rue Gay Lussac
Paris 5

GERMAN DEMOCRATIC REPUBLIC

Deutscher Buchexport und Import
Leninstraße 16.
Leipzig C. I.
Zeitungvertriebsamt
Clara Zetkin Straße 62.
Berlin N. W.

GERMAN FEDERAL REPUBLIC

Kunst und Wissen
Eich Bieber
Postfach 46.
7 Stuttgart S.

GREAT BRITAIN

Collet's' Subscription Dept.
44—45 Museum Street
London W. C. I.
Robert Maxwell and Co. Ltd.
Waynflete Bldg. The Plain
Oxford

HOLLAND

Swetz and Zeitlinger
Keizersgracht 471—487
Amsterdam C.
Martinus Nijhof
Lange Voorhout 9
The Hague

INDIA

Current Technical Literature
Co. Private Ltd.
Head Office:
India House OPP.
GPO Post Box 1374
Bombay I.

ITALY

Santo Vanasia
71 Via M. Macchi
Milano
Libreria Commissionaria Sanson
Via La Marmora 45
Firenze

JAPAN

Nauka Ltd.
2 Kanada-Zimbocho 2-ehome
Chiyoda-ku
Tokyo
Maruzen and Co. Ltd.
P.O. Box 605
Tokyo
Far Eastern Booksellers
Kanada P.O. Box 72
Tokyo

KOREA

Chulpanmul
Korejskoje Obschestvo po Exportu
Importu Proizvedenij Pechati
Phenjan

NORWAY

Johan Grund Tanum
Karl Johansgatan 43
Oslo

POLAND

Export- und Import-Unternehmen
RUCH
ul. Wilcza 46.
Warszawa

ROUMANIA

Cartimex
Str. Aristide Briand 14—18.
Bucuresti

SOVIET UNION

Mezhdunarodnaja Kniga
Moscow
G—200

SWEDEN

Almquist and Wiksell
Gamla Brogatan 26
Stockholm

USA

Stechert Hafner Inc.
31 East 10th Street
New York 3 N. Y.
Walter J. Johnson
111 Fifth Avenue
New York 3 N. Y.

VIETNAM

Xunhasaba
Service d'Export et d'Import des Livres et Périodiques
19. Tran Quoc Toan
Hanoi

YUGOSLAVIA

Forum
Vojvode Misiva broj 1.
Novi Sad
Jugoslovenska Kniga
Terazije 27.
Beograd

1980. IV. 30. HU ISSN 0001—5954 **Index: 26014**