Studia Scientiarum Mathematicarum Hungarica

AUXILIO
CONSILII INSTITUTI MATHEMATICI
ACADEMIAE SCIENTIARUM HUNGARICAE

REDIGIT
L. FEJES TÓTH

ADIUVANTIBUS
M. ARATÓ, I. CSISZÁR, T. FREY, G. FREUD,
B. GYIRES, A. HEPPES, L. KALMÁR, E. MAKAI,
A. PRÉKOPA, P. RÉVÉSZ, K. TANDORI

TOMUS VIII.
FASC. 1–2.
1973

AKADÉMIAI KIADÓ, BUDAPEST

Studia Scientiarum Mathematicarum Hungarica

A Magyar Tudományos Akadémia matematikai folyóirata

Szerkesztőség: 1061 Budapest V., Reáltanoda u. 13—15.

Technikai szerkesztő: Petruska Gy.

Kiadja az Akadémiai Kiadó, 1054 Budapest V., Alkotmány u. 21.

A *Studia Scientiarum Mathematicarum Hungarica* angol, német, francia vagy orosz nyelven közöl eredeti értekezéseket a matematika tárgyköréből. Félévenként jelenik meg, évi egy kötetben.

Előfizetési ára belföldre 120,—Ft, külföldre 165,— Ft. Megrendelhető a belföld számára az Akadémiai Kiadónál, a külföld számára pedig a Kultúra Könyv és Hirlap Külkereskedelmi Vállalatnál (1011 Budapest II., Fő u. 32.).

Cserekapcsolatok felvétele ügyében kérjük a MTA Matematikai Kutató Intézete Könyvtárához (1061 Budapest V., Reáltanoda u. 13—15.) fordulni.

A közlésre szánt dolgozatokat kérjük két példányban a szerkesztőség címére küldeni.

Studia Scientiarum Mathematicarum Hungarica is a journal of the Hungarian Academy of Sciences publishing original papers on mathematics, in English, German, French or Russian.

It is published semiannually, making up one volume per year.

Editorial Office: 1061 Budapest V., Reáltanoda u. 13—15, Hungary.

Technical Editor: Gy. Petruska

Subscription rate: $ 16.00 per volume. Orders may be placed with *Kultúra* Trading Co. for Books an Newspapers, Budapest 62, P. O. B. 149 or with its representatives abroad.

For establishing exchange relations please write to the Library of the Mathematical Institute (1061 Budapest V., Reáltanoda u. 13—15.)

Papers intended for publication should be sent to the Editor in 2 copies.

315.930

Studia Scientiarum Mathematicarum Hungarica

Auxilio
Consilii Instituti Mathematici
Academiae Scientiarum Hungaricae

Redigit

L. Fejes Tóth

Adiuvantibus

M. Arató, I. Csiszár, T. Frey, G. Freud, B. Gyires, A. Heppes, L. Kalmár, E. Makai, A. Prékopa, P. Révész, K. Tandori

Tomus VIII

Akadémiai Kiadó, Budapest
1973

Studia Scientiarum
Mathematicarum Hungarica

STUDIA SCIENTIARUM MATHEMATICARUM HUNGARICA

Tomus VIII

INDEX

Huq, S. A.: Some remarks on multicategories 1
Варшамов, Р. Р.: Об одном линейном операторе в поле Галуа и его приложении 5
Mogyoródi, J.: On the rarefaction of renewal processes, III. 21
Mogyoródi, J.: On the rarefaction of renewal processes, IV. 29
Scozzafava, R.: The radical of a locally compact semialgebra 39
Chartrand, G. and *Wall, C. E.:* On the hamiltonian index of a graph 43
Rejtő, L.: A more general Farkas-type theorem 49
Jucovic, E.: On the face vector of a 4-valent 3-polytope 53
Крамли, А.: Геодезические потоки на компактнах римановых поверхностях без фокалных точек 59
Steck, G. P. and *Simmons, G. I.:* On the distributions of R_{mn}^{+} and (D_{mn}^{+}, R_{mn}^{+}) 79
Chakerian, G. D.: On a certain affine invariant functional for convex bodies 91
Major, P.: A law of the iterated logarithm for the Robbins—Monro method 95
Чан Суань Хиен: О некоторых задачах относящихся к одному семейство уравнений в частных поизводных составного типа 103
Невai, Г. П.: Замечания о тригонометрическом интерполировании и суммах Фурье 113
Székely, G. J.: On the coefficients of polynomials with negative zeros 123
Győri, I.: Asymptotic behaviour of solutions of unstable-type first order differential equations with delay 125
Makai, E.: Variatonal problems related to those which lead to the Schrödinger equation of heliumlike atoms 133
Babai, L. and *Lovász, L.:* Permutation groups and almost regular graphs 143
Lovász, L. and *Major, P.:* A note to a paper of Dudley 151
Stoyan, D.: Bounds for the extrema of the expected value of a convex function of independent random walk 153
Steinfeld, O.: Notes on regular duo elements, rings and semigroups 161
Lovász, L.: Independent sets in critical chromatic graphs 165
Elbert, A.: On the consecutive zeros of the soultions of linear second order differential equations 169
Neumann, F.: Distribution of zeros of solutions of $y'' = q(t)\,y$ in relation to their behaviour in large 177
Máté, L.: On a representation of module-homomorphisms (multipliers) 187
Mogyoródi, J.: On the rarefaction of renewal processes, V. 193
Mogyoródi, J.: On the rarefaction of renewal processes, VI. 199
Sujan Srivastava: Joint distributions based on runs and the other number of intersections 211
Simonovits, A.: Direct comparison of different priority queuing disciplines 225
Bihari, I.: Perturbations of n dimensional linear systems of differential equations 245
Scherer, K.: On extensions of a theorem of S. N. Bernstein 269
Gerencsér, L.: On a theorem of Lyapunoff 273
Freud, G.: Investigations on weighted approximation by polynomials 285
Kapoor, S. F. and *Stewart, M. J.:* A note on the Hamiltonian index of graphs 307
Fritz, J.: Stochastic approximation for finding local maxima of probability measures 309
Невai, Г. П. (Névai, G. P.) Две теоремы об отклонении сумм Фурье 323
Komlós, J. and *Révész, P.:* A modification of the Robbins—Monro process 329
Nemetz, T.: Bounds on error probability 341
Major, P.: On a non-parametric estimation of the regression function 347
Brass, H. und *Günttner, R.:* Eine Fehlerabschätzung zur Interpolation stetiger Funktionen ... 363

Halász, G.: On a result of Salem and Zygmund concerning random polynomials 369
Katona, G.: A three part Sperner theorem 379
Révész, P.: Robbins—Monro procedure in a Hilbert space and its application in the theory of learning processes I. 391
Freud, G. and *Németh, G.:* On the L_p-norms of orthonormal Hermite functions 399
Körner, J.: An extension of the class of perfect graphs 405
Berkes, I.: Functional limit theorems for lacunary trigonometric and Walsh series 411
Berkes, I.: On Strassen's version of the loglog law for multiplicative systems 425
Cvetković, D. M.: Inequalities obtained on the basis of the spectrum of the graph 433
Saxena, R. B.: Approximation of continuous functions by polynomials involving Freud's interpolatory polynomials 437
Joshi, V. N.: On the class of matrices which are convergent for a given matrix 447
Joshi, V. N.: Remarks on iterative methods for computing the generalized inverse 457
Andréka, H. and *Németi, I.:* On the equivalence of sets definable by satisfaction and ultrafilters 463
Révész, P.: Robbins—Monro procedure in a Hilbert space, II. 469
Kallenberg, O.: Characterization of continuous random processes and signed measures 473
Günttner, R.: Asymptotisch optimale Abschätzungen zwischen Halbnormen 479
Szabados, J.: Generalization of a problem of P. Turán 485

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 1—3.

SOME REMARKS ON MULTICATEGORIES

by

S. A. HUQ

Let $\mathscr{A}$ be a class $A, B, \ldots$ of objects; we also postulate that for every ordered pair of objects A, B of $\mathscr{A}$, we have a set $[A, B]_{\mathscr{A}}$ and we consider the power set $\mathscr{P}[A, B]_{\mathscr{A}}$. Let S_{AB} be the collection of elements C_{AB} of $\mathscr{P}[A, B]_{\mathscr{A}}$, which we call morphisms from $A \to B$ and denote by $A \xrightarrow{C_{AB}} B$ and which are hereditary with respect to being morphisms (i.e. every subset of C_{AB} is also a morphism from $A \to B$.)

Definition. A multicategory $\mathscr{A}$ is a class of objects as above, and with every ordered pair of objects A, B, we have the set S_{AB} of morphisms, subject to the composition law

$$\text{(H)} \qquad S_{BC} \times S_{AB} \to S_{AC}; \quad C_{BC} \circ C_{AB} = C_{AC}$$

defined by

$$\text{(H)}' \qquad \Big(\bigcup_{\gamma_i \in C_{BC}} \gamma_i\Big)\Big(\bigcup_{\beta_j \in C_{AB}} \beta_j\Big) = \Big\{\bigcup_{i,j} \gamma_i \beta_j\Big\} \in C_{AC}$$

(where the union on the left hand side has been taken over the elements of the respective sets) which is

$$\text{(1)} \qquad \text{associative:} \quad C_{CD} \circ (C_{BC} \circ C_{AB}) = (C_{CD} \circ C_{BC}) \circ C_{AB}$$

and such that

(2) for every object B, we have an element $C_{BB} \in S_{BB}$ called the identity of B, such that

$$C_{BB} \circ C_{AB} \supseteq C_{AB} \quad \text{and} \quad C_{AB} \circ C_{AA} \supseteq C_{AB},$$

whenever the compositions make sense.

Thus a multicategory is nearly a category with slightly wider identity laws and whose law of composition could be multivalued. (For categories see B. MITCHELL [3].)

We notice that if

$$C'_{BB} \in S_{BB}$$

is another identity for B, then

$$C_{BB} \circ C'_{BB} \supseteq \text{both} \quad C_{BB} \quad \text{and} \quad C'_{BB}.$$

Examples

(i) If the sets C_{AB} are singletons and the composition (H) is single valued, we are back to the definition of a category; in this case for any two identities C_{BB} and C'_{BB} of B, we must have $C_{BB} = C'_{BB}$.

MAGYAR TUDOMÁNYOS AKADÉMIA KÖNYVTÁRA

(ii) A multigroup or hypergroup [1, 2], is a multicategory with a single object.

A semigroup S with identity 1, can be made into a multicategory with single object S, in the following ways.

(iiia) the elements of the semigroup are morphisms from $S \to S$ and the composition for two morphisms a, b from $S \to S$ is given by

$$a \circ b = \{a, b, ab\}.$$

We note $a \circ (b \circ c) = (a \circ b) \circ c$ both being the set

$$\{a, b, c, ab, ac, bc, abc\}$$

and $a \circ 1 = \{a, 1\}$

(iiib) the morphisms are the elements of the power set $\mathscr{P}(S)$ having domain and codomain S, and the composition being defined as

$$S^* \circ S'^* = \{S^*, S'^*, S^* S'^*\}$$

where

$$S^* S'^* = \{ab | a \in S^*, b \in S'^* \subset S\}.$$

In view of the composition law (H) being defined by (H)′ it is quite sufficient to check the associativity for the composition in Example (iiia) for three elements $a, b, c \in S$; identity law is easy.

Now we raise the following questions

(A) Given any category $\mathscr{A}$, can we construct a multicategory having the same set of objects and morphisms such that the composition in the category $\alpha\beta$ of two morphisms α and $\beta \in \alpha \circ \beta$ in the multicategory; in this case we say that the category has been lifted to a multicategory.

(B) Given a multicategory $\mathscr{A}$, can we construct a category in which the objects and morphisms are the same but the multicategory composition

$$C_{BC} \circ C_{AB} \supseteq C_{BC} C_{AB}$$

in the category thus constructed.

We answer (A) in the affirmative and (B) in the negative. Precisely

THEOREM C. *Every category $\mathscr{C}$ can be lifted into a multicategory.*

PROOF. For two morphisms $\alpha: A \to B$ and $\beta: B \to C$, we define the multicategory composition as

$$\beta \circ \alpha = \mathrm{Hom}_{\mathscr{C}}(A, C) = \text{set of all morphisms from } A \to C.$$

Examples (iiia) and (iiib) and theorem C show that the lifting is not necessarily unique; in fact the lifting defined in theorem C is the maximal one, with the identical lifting as minimal.

Counter example to (B)

We shall use a multicategory with single object. We choose a group $G \neq \{1\}$ with a conjugacy relation on it. We consider this as a multicategory with single object G, with elements of the group as morphisms and composition

$$a \circ b = [a] \cup [b],$$

where $[a]$ denotes the equivalence class of a.

It is easy to check the associativity and the identity law. Suppose there exists a category with same objects and certain set of morphisms and in which the multicategory operation "$\circ$" contains the category composition; then for any morphism $a: G \to G$, $\exists a^{-1}: G \to G$ such that $a \cdot a^{-1} = 1$ (the identity map); so by hypothesis

$$1 \in [a] \cup [a^{-1}] \quad \text{i.e.} \quad G = [1],$$

contrary to our initial assumption that G is non trivial.

It would be interesting to study subdirect embeddings (see A. SULINSKI [4]) and their dualizations (see F. SZÁSZ—R. WIEGANDT [5], [6]) in some multicategories.

REFERENCES

[1] DRESHER, M. and ORE, O.: Theory of multigroups, *Amer. Journ. of Maths* **60** (1938), 705—733.
[2] WALL, H. S.: Hypergroups, *ibid* **59** (1937), 77—98.
[3] MITCHELL, B.: *Theory of Categories*, Academic Press, New York (1965).
[4] SULINSKI, A.: The Brown—McCoy radical in categories, *Fund. Math.* **59** (1966), 23—41.
[5] SZÁSZ, F.—WIEGANDT, R.: On the dualization of subdirect embeddings, *Acta Math. Acad. Sci. Hungar.* **20** (1969), 289—302.
[6] SZÁSZ, F.—WIEGANDT, R.: On the duality of radical and semi-simple objects in categories, *Acta Math. Acad. Sci. Hungar.* **21** (1970) 175—182.

Institute of Advanced Studies in Mathematics,
Australian National University;

at present: Department of Mathematics, University of Khartoum,
P. O. Box 321, Khartoum, Sudan.

(Received January 18, 1971.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 5—19.

ОБ ОДНОМ ЛИНЕЙНОМ ОПЕРАТОРЕ В ПОЛЕ ГАЛУА И ЕГО ПРИЛОЖЕНИИ

Р. Р. ВАРШАМОВ

Как известно, для всякого простого p и натурального числа s существует (и с точностью до изоморфизма), единственное поле Галуа $GF(q)$ порядка $q=p^s$, то есть множество из элементов с двумя операциями (сложения и умножения). Для того чтобы построить поле $GF(q^n)$ достаточно иметь полином степени n неприводимый в $GF(q)$.

Имеется ряд основополагающих работ, вскрывающих элементарную структуру полей Галуа. Особенно полная теория содержится в монографии Альберта [1] „Фундаментальные вопросы современной алгебры". Однако единственная проблема, оставленная структурными теоремами и являющаяся одновременно наиболее важной и трудной проблемой теории полей Галуа, это проблема синтеза неприводимых в поле $GF(q)$ полиномов заданной степени в явном виде. Проблемы приводимости полиномов над конечными полями, имеющие достаточно самостоятельный интерес, занимают особое место в современной технике. Можно назвать ряд областей, где теория приводимости играет решающую роль. Так, например, в теории кодирования (теории циклических кодов), в теории линейных рекурсивных последовательностей, в теории кольцевых счетчиков и др. С точки зрения методов существуют два принципиально различных подходов к решению указанной проблемы. Это взаимно дополняющие друг друга, конструктивный метод синтеза неприводимых полиномов, использующий алгебраическую структуру и внутреннее строение полей Галуа и метод анализа, основанный на исследовании мультипликативных свойств разложимых в конечных полях полиномов. Принцип на котором основан второй метод, чрезвычайно прост: из множества всевозможных полиномов степени n выбирается наугад один из них. Затем, по специально разработанному алгоритму проверяется разложимость его в поле $GF(q)$. Если исследуемый полином оказывается разложимым, то его отбраковывают, и поиск продолжается до благоприятного исхода.

Что же касается конструктивного метода, то он является наиболее интересным и притом плохо изучаемым в литературе.

Изложению некоторых результатов конструктивной теории синтеза неприводимых полиномов над конечными полями, является основным содержанием настоящей работы.

§ 1. Оператор L^0.

В этом параграфе исследуется класс операторов, областью определения которых являются полиномы с коэффициентами из поля $GF(q)$, обладающий рядом интересных алгебраических свойств.

Ниже приводятся некоторые важные в приложении результаты, устанавливающие тесную связь между степенями неприводимых функций и показателями связанных с ними полиномов. Это позволяет строить в явном виде неприводимые полиномы из некоторого специального класса полиномов.

Обозначим через $L^\theta f(x)$ выражение

$$L^\theta f(x) = \frac{1}{\theta(x)} \sum_{u=0}^{\theta} \sum_{v=0}^{n} \theta_u a_v x^{uq^v}, \tag{1}$$

где

$$f(x) = \sum_{v=0}^{n} a_v x^v, \quad \theta(x) = \sum_{u=0}^{\theta} \theta_u x^u \big(\theta(x) \not\equiv \text{const.}\big).$$

Лемма 1. *В поле $GF(q)$ полином $L^\theta r(x)$ делит без остатка $L^\theta f(x)$, где $r(x)$ — любой делитель $f(x)$*[1].

Доказательство. Подставляя в (1) $f(x) = r(x)h(x)$, где

$$r(x) = \sum_{i=0}^{r} r_i x^i \quad \text{и} \quad h(x) = \sum_{j=0}^{h} h_j x^j,$$

получим

$$L^\theta f(x) = \frac{1}{\theta(x)} \sum_{u=0}^{\theta} \sum_{i=0}^{r} \sum_{j=0}^{h} \theta_u r_i h_j x^{uq^j+i} =$$

$$= \frac{1}{\theta(x)} \sum_{j=0}^{h} h_j \theta(x)^{q^j} \big(L^\theta r(x)\big)^{q^j} = \sum_{j=0}^{h} h_j \theta(x)^{q^j-1} \big(L^\theta r(x)\big)^{q^j},$$

т. е. $L^\theta r(x) | L^\theta f(x)$. Лемма доказана.

Лемма 2. *Функция $F(x) = L^x f(x)$ является сепарабельной при всяком полиноме $f(x)$ с ненулевым свободным членом.*

Доказательство. В самом деле, в силу (1) имеем

$$\big(xF(x)\big)' = F(x) + xF(x)' = f(0).$$

Откуда, ввиду $f(0) \neq 0$, находим $\big(F(x), F(x)'\big) = \text{const}$. Лемма доказана.

Основная теорема. *Пусть $f(0) \neq 0, s_1(x), s_2(x), \ldots, s_k(x)$ — множество всех собственных делителей $f(x)$; $K_f\{L^\theta f(x)\}$ — наименьшее общее кратное $L^\theta s_1(x)$, $L^\theta s_2(x), \ldots, L^\theta s_k(x)$; N — показатель полинома $f(x)$*[2], *$\varrho(x)$ — неприводимый множитель* (1), *не делящий $K_f\{L^\theta f(x)\}$.*

Тогда c — степень полинома $\varrho(x)$ кратна N, т. е. $(c, N) = N$.

---

[1] На самом деле имеем место более сильное утверждение:

$$\big(L^\theta f_1(x), L^\theta f_2(x)\big) = L^\theta\big(f_1(x), f_2(x)\big).$$

[2] То есть минимальное натуральное число с условием

$$x^N \equiv 1 \pmod{f(x)}.$$

Доказательство. Допустим противное, т. е. $(c, N) \neq N$. Тогда, очевидно, выражение $S(x) = (x^c - 1, f(x))$ явится собственным делителем $f(x)$ и, в силу леммы 1,

$$L^\theta S(x) = \Lambda_1(x) L^\theta(x^c - 1) + \Lambda_2(x) L^\theta f(x), \tag{3}$$

так как

$$S(x) = \lambda_1(x)(x^c - 1) + \lambda_2(x) f(x).$$

По определению, $\theta(x) L^\theta(x^c - 1) = \sum_{u=0}^{\theta} \theta_u (x^{uq^c} - x^u)$, а это показывает, что $x^{q^c - 1} - 1 | \theta(x) L^\theta(x^c - 1)$ и одновременно $\varrho(x) | \theta(x) L^\theta(x^c - 1)$, так как степень неприводимого полинома $\varrho(x)$ равна c. Заметим, однако, что $(\varrho(x), \theta(x)) = \text{const}$. ввиду $(\theta(x) L^\theta f(x))' = a_0 \theta(x)'$ $(a_0 \neq 0)$. Поэтому ясно, что $\varrho(x) | L^\theta(x^c - 1)$, а также, в силу (3), $\varrho(x) | L^\theta S(x)$. Это противоречит условию теоремы. Таким образом, наше допущение неверно; следовательно, $(c, N) = N$. Теорема доказана.

Основная теорема имеет приложение в конструктивной теории синтеза неприводимых по простому модулю полиномов. Изложим некоторые полученные с её помощью результаты.

Теорема 1. *Пусть* $\theta(x) = x$. *Тогда* c — *степень неприводимого полинома* $\varrho(x)$, *удовлетворяющего условиям основной теоремы, совпадает с показателем* $f(x)$, *т. е.* $c = N$.

Доказательство. Поскольку $f(x) | x^N - 1$, то, согласно лемме 1, $L^x f(x) | L^x(x^N - 1)$ и $\varrho(x) | x^{q^N - 1} - 1$, так как $L^x(x^N - 1) = x^{q^N - 1} - 1$. Полином $\varrho(x)$ неприводим, поэтому $c | N$, а это совместно с основной теоремой дает $c = N$. Теорема доказана.

Кроме того, для $q = 2$[1] будем иметь:

Теорема 2. *Пусть* $f(x) = \prod_{i=1}^{\sigma} f_i(x)$, $f_i(x)$ — *примитивные полиномы с попарно простыми степенями* n_i *соответственно,* $|\alpha| = \sum_{i=1}^{\sigma} \alpha_i$, *где* α_i *целые* $(0 \leqq \alpha_i \leqq 1)$ $f^{(\alpha)}(x) = \prod_{i=1}^{\sigma} f^{\alpha_i}(x)$. *Тогда полином*

$$F(x) = \prod_{j=0}^{[\sigma/2]} \Big(\prod_{|\alpha| = \sigma - 2j - 1} L^x f^{(\alpha)}(x) \Big)^{-1} \prod_{|\alpha| = \sigma - 2j} L^x f^{(\alpha)}(x),$$

где $[x]$ — *наибольшее целое, не превосходящее* x — *неприводим.*

Доказательство. Легко показать, что полином $F(x)$, являясь делителем $L^x f(x)$, взаимно прост с $K_f\{L^x f(x)\}$. Кроме того, несложным подсчетом устанавливается, что степень его равна $M = \prod_{i=1}^{\sigma} (2^{n_i} - 1)$. Как известно, показатель полинома $f(x)$ также совпадает с M. Отсюда следует, согласно теореме 1, что полином $F(x)$ — неприводим. Теорема доказана.

Теорема 2 позволяет сравнительно легко находить неприводимые полиномы, степени которых представимы в виде произведения попарно взаимных

[1] Случай наиболее интересный в практическом приложении.

простых чисел Мерсенна. Так, например, в случае, когда $f(x) = (x^2+x+1)(x^3+x^2+1) = x^5+x+1$, будем иметь неприводимый полином 21-й степени

$$F(x) = \frac{L^x(x^5+x+1)}{(L^x(x^2+x+1))L^x(x^3+x^2+1)} =$$

$$= x^{21}+x^{19}+x^{18}+x^{15}+x^{14}+x^{11}+x^8+x^7+x^5+x^4+1.$$

Аналогичным образом в случае $f(x) = (x^2+x+1)(x^5+x^2+1)$ получим неприводимый полином 93-й степени

$$F(x) = (x^7+1)^3(x^4+x^2+x+1)(x^{68}+x^{37}+x^{12}+x^9+x^6+x^4)+1.$$

Теорема 3. *Пусть* $f(x) = \sum_{v=0}^{n} a_v x^v$, $f(1) \neq 0$, $\omega(x) = \sum_{u=0}^{\omega} \omega_u x^u$, $\sigma(\omega) = xL^x\omega(x)+1$, $\Lambda(x) = (x+1)\omega(x)f(x)$, $\Lambda(0) \neq 0$, K — *наименьшее общее кратное* $\{L^{\sigma(\omega^i)}f^i(x)\}_i$ $(\omega^i(x)|\omega(x),\ f^i(x)|f(x),\ \omega^i(x)f^i(x) \neq \omega(x)f(x))$ *и* $\varrho(x)$ — *неприводимый множитель* $L^{\sigma(\omega)}f(x)$, *не делящий* K. *Тогда степень полинома* $\varrho(x)$ *совпадает с показателем* $\Lambda(x)$.

Доказательство. Имеем

$$\sigma(\omega)L^{\sigma(\omega)}f(x) = \sum_{u=0}^{\omega}\sum_{v=0}^{n} \omega_u a_v x^{2^u\cdot 2^v} + \sum_{v=0}^{n} a_v =$$

$$= xL^x \sum_{u=0}^{\omega}\sum_{v=0}^{n} \omega_u a_v x^{u+v} + 1 = xL^x\omega(x)f(x)+1.$$

Однако, согласно (2),

$$xL^x\omega(x)f(x)+1 = (L^x\omega(x)f(x))^{-1}L^x(x+1)\omega(x)f(x)$$

и поэтому

$$\sigma(\omega)(L^{\sigma(\omega)}f(x))L^x\omega(x)f(x) = L^x\Lambda(x). \tag{4}$$

Отсюда, так как выражение $L^x\Lambda(x)$ не имеет кратных корней $(\Lambda(0) \neq 0)$, следует, что

$$(\sigma(\omega)L^x\omega(x)f(x),\ L^{\sigma(\omega)}f(x)) = 1. \tag{5}$$

Покажем теперь, что $\varrho(x) \nmid K_\Lambda\{L^x\Lambda(x)\}$. В самом деле, в случае $\varrho(x)|K_\Lambda\{L^x\Lambda(x)\}$ мы имели бы $\varrho(x)|L^x\Lambda^1(x)$, где $\Lambda^1(x)$ — собственный делитель $\Lambda(x)$, причем, в силу (5), $x+1|\Lambda^1(x)$. Пусть $\Lambda^1(x) = (x+1)\omega^1(x)f^1(x)$ $(\omega^1(x)|\omega(x),\ f^1(x)|f(x),\ \omega^1(x)f^1(x) \neq \omega(x)f(x))$. Согласно (4) $\sigma(\omega^1)(L^{\sigma(\omega^1)}f^1(x))L^x\omega^1(x)f^1(x) = L^x\Lambda^1(x)$, а поэтому $\varrho(x)|\sigma(\omega^1)(L^{\sigma(\omega^1)}f^1(x))L^x\omega^1(x)f^1(x)$. Но, в силу леммы 1, $L^x\omega^1(x)f^1(x)|L^x\omega(x)f(x)$ и ввиду очевидного тождества $\sigma(\omega)L^x\omega(x) \equiv L^x(x+1)\omega(x)$, также $\sigma(\omega^1)L^x\omega^1(x)|\sigma(\omega)L^x\omega(x)$. Кроме того, в силу (5), $\varrho(x) \nmid L^x\omega(x)f(x)$ и $\varrho(x) \nmid \sigma(\omega)L^x\omega(x)$. Поэтому ясно, что $\varrho(x)|L^{\sigma(\omega^1)}f^1(x)$, или, что то же самое $\varrho(x)|K$. Это противоречит условию теоремы, следовательно, $\varrho(x) \nmid K_\Lambda\{L^x\Lambda(x)\}$. Таким образом, мы убеждаемся в том, что полином $\varrho(x)$, являясь делителем $L^x\Lambda(x)$, взаимно прост с $K_\Lambda\{L^x\Lambda(x)\}$. Отсюда, согласно

теореме 1, следует, что степень его совпадает с показателем полинома $\Lambda(x)$. Теорема доказана.

Опираясь на теорему 3, можно доказать следующее.

Теорема 4. *Пусть δ означает одно из чисел 0 или 1; $\omega(x) = x^{1+\delta}+1$, $f_i(x)$ $(i=1, \sigma)$ — примитивные полиномы с попарно взаимно простыми степенями n_i соответственно. Тогда полином*

$$F(x) = \prod_{j=0}^{[\sigma/2]} \Big(\prod_{|\alpha|=\sigma-2j-1} L^{\sigma(\omega)} f^{(\alpha)}(x) \Big)^{-1} \prod_{|\alpha|=\sigma-2j} L^{\sigma(\omega)} f^{(\alpha)}(x) - \tag{6}$$

— неприводим.

Доказательство. Покажем прежде всего, что $\varrho(x) \nmid K_\Lambda\{L^x \Lambda(x)\}$, где $\varrho(x) \neq 1$ — любой неприводимый делитель $F(x)$. В самом деле, допустим противное, а именно, $\varrho(x) | K_\Lambda\{L^x \Lambda(x)\}$. Тогда очевидно, $\varrho(x) | L^x \Lambda^1(x)$, где $\Lambda^1(x)$ — собственный делитель $\Lambda(x)$ н, в силу (5), а также согласно условию теоремы $(\omega(x) = x^{1+\delta}+1)$, $\Lambda^1(x) = (x+1)\omega(x) f^1(x) (f^1(x) | f(x), f^1(x) \neq f(x))$. Отсюда, в силу (4), (5) и ввиду $\varrho(x) | L^{\sigma(\omega)} f(x)$, следует $\varrho(x) | L^{\sigma(\omega)} f^1(x)$. Это противоречит условию теоремы, как как из формулы (6) видно, что $\varrho(x) \nmid K_f\{L^{\sigma(\omega)} f(x)\}$. Таким образом, наше допущение неверно, следовательно, $\varrho(x) \nmid K_\Lambda\{L^x \Lambda(x)\}$. Заметим, теперь, что степень полинома $F(x)$ равная $\prod_{i=1}^{\sigma} (2^{n_i}-1)$ совпадает с показателем полинома $\Lambda(x)$. Поэтому, согласно теореме 3, полином $F(x) = \varrho(x)$ — неприводим. Теорема доказана.

Теорема 4 расширяет возможности построения неприводимых полиномов. С её помощью можно, например, находить неприводимые полиномы, степени которых представляются в виде $2^{1+\delta} \prod (2^{n_i}-1)$. Так, например, в случае $\delta=0$ и $f(x) = x^5+x+1$ получим неприводимый полином 42-й степени:

$$F(x) = \frac{L^{\sigma(x+1)}(x^5+x+1)}{(L^{\sigma(x+1)}(x^2+x+1)) L^{\sigma(x+1)}(x^3+x^2+1)} = x^{42}+x^{41}+x^{40}+x^{35}+x^{32}+$$

$$+x^{30}+x^{29}+x^{28}+x^{27}+x^{26}+x^{25}+x^{23}+x^{18}+x^{15}+x^{14}+x^{13}+x^8+x^6+1.$$

Аналогичным образом при $\delta=1$ и $f(x) = x^3+x+1$ будем иметь неприводимый полином 28-й степени:

$$F(x) = L^{\sigma(x^2+1)}(x^3+x+1) = (x^{17}+x^2)(x^{11}+x^8+x^7+x^5+x^3+x^2+x+1)+1$$

и т. д.

§ 2. Операторная подготовка $x \to \sigma_g(x, \delta)$

В этом параграфе делается попытка несколько расширить возможности построения в явном виде неприводимых полиномов.

В современной технике чаще всего используется двоичная система счисления, поэтому коды с основанием 2 в практическом отношении наиболее важны. Кроме того, большинство результатов, полученных для них без особого труда могут быть обобщены и на случай более высокого основания $q>2$. Поэтому ради простоты изложения и большей наглядности в дальнейшем почти

везде, если это не будет специально оговорено, ограничимся рассмотрением случая $q=2$. Обозначим через $I_n^u(x)$ $(I_n^{(0)}(x)=1, I_n^{(t(n))}(x)=I_n(x))$ произведение u различных неприводимых над полем $GF(2)$ полиномов степени n, общее число которых $t(n)=\frac{1}{n}\sum_{u|n}\mu(u)2^{n/u}$, где $\mu(u)$ — функция Мёбиуса. Кроме того введем в рассмотрение функцию $\sigma_g(x,\delta)=xL^xg(x)+\delta$, где $g(x)\neq x$ — неприводимый полином m-й степени, а δ — произвольный элемент поля $GF(2)$.

Предположим теперь, что полином n-й степени $f(x)$ неразложим и $(n, T)=1$, где T — показатель функции $(x+1)g(x)$. Тогда,

Теорема 5. Имеет место соотношение

$$f(\sigma_g(x,\delta))=I_n^{(l_2)}(x)I_{nT}^{(T^{-1}(2^m-1)l_2)}(x), \tag{7}$$

где $l_1=l_2=1$, *если* $g(x)\neq x+1$ *и* $l_1=2, 0$, *соответственно* $l_2=0, 2$, *если* $g(x)=x+1$.

Доказательство. Рассмотрим сначала случай, когда n является простым числом. Тогда очевидно

$$x^{2^n}+x=x(x+1)I_n(x)$$

и после замены x через $\sigma_g(x,\delta)$ получим

$$x\left[x^{2^n-1}(L^xg(x))^{2^n}+L^xg(x)\right]=(xL^xg(x))(xL^xg(x)+1)I_n(\sigma_g(x,\delta)),$$

или, в силу леммы 1,

$$L^x(x^n+1)g(x)=(L^x(x+1)g(x))I_n(\sigma_g(x,\delta))$$

Отсюда, ввиду

$$L^x(x^n+1)|L^x(x^n+1)g(x),\ L^x(x^n+1)g(x)|L^x(x^{nT}+1)$$

(так как $(x^n+1)g(x)|x^{nT}+1$ и $(n, T)=1$) ясно, что

$$L^x(x^n+1)g(x)=(L^x(x+1)g(x))I_n(\sigma_g(x,\delta))=\Phi(x)I_n(x)L^x(x+1)g(x), \tag{8}$$

где $\Phi(x)=\prod_{u|nT}I_u^{(\varepsilon_u)}(x)$, ε_u — целые неотрицательные чисца.

В дальнейшем, задача будет заключаться в том, чтобы показать, что $(\Phi(x), I_{nT}(x)^{-1}(x^{2^{nT}-1}+1))=1$, т. е., что $\varepsilon_u=0$ для всякого $u\neq nT$.

В самом деле, допустим противное, а именно, предположим, что среди простых делителей $\Phi(x)$ найдется хотя бы один (обозначим его через $\lambda(x)$), степень которого N является собственным делителем количества nT. Тогда, как нетрудно понять, ввиду $(n, T)=1$, будет место равенства

$$((x^n+1)g(x), x^N+1)=(x^{(n,N)}+1)g(x)^H\quad\left(H=\left[\frac{(N,T)}{T}\right]\right).$$

Последнее равенство позволяет написать соотношение:

$$b_1(x)(x^n+1)g(x)+b_2(x)(x^N+1)=(x^{(n,N)}+1)g(x)^H,$$

или, согласно лемме 1,

$$B_1(x)L^x(x^n+1)g(x)+B_2(x)L^x(x^N+1) = L^x(x^{(n,N)}+1)g(x)^H. \tag{9}$$

Здесь $b_1(x)$, $b_2(x)$, $B_1(x)$, $B_2(x)$ — полиномы с коэффициентами из поля GF (2).

Из (9) непосредственно вытекает выражение

$$\lambda(x)|L^x(x^{(n,N)}+1)g(x)^H$$

каторое ввиду $L^x(x^{(n,N)}+1)g(x)^H|I_n(x)L^x(x+1)g(x)$, а также (8) и леммы 2, дает противоречащее принятому допущению соотношение

$$\lambda(x)\nmid\Phi(x).$$

Таким образом, наше предположение неверно; следовательно $N=nT$. Итак, установлено, что степень любого неприводимого делителя $\varphi(x)$ совпадает с nT. Исходя из этого, в силу (8) и учитывая, что ст. $I_n(\sigma_g(x,\delta)) = nt(n)2^m$ и ст. $I_n(x)=nt(n)$ можно написать выражение

$$I_n(\sigma_g(x,\delta)) = I_n(x)I_{nT}^{(T^{-1}(2^m-1)t(n))}(x). \tag{10}$$

Здесь ст. $f(x)$ обозначает степень полинома $f(x)$. Последнее равенство позволяет получить формулу

$$f(\sigma_g(x,\delta)) = I_n^{(\alpha)}(x)I_{nT}^{(\beta)}(x), \tag{11}$$

справедливую для любого неприводимого полинома $f(x)$, n-й степени. Из соотношения (11), ввиду ст. $f(\sigma_g(x,\delta)) = n2^m$, ст. $I_n^{(\alpha)}(x) = n\alpha$ и $I_{(nT)}^{(\beta)}(x)=nT\beta$ находим $n2^m = n\alpha+nT\beta$ или

$$2^m = \alpha+T\beta. \tag{12}$$

Если $g(x)\neq x+1$, то $T|2^m-1$ ($T\neq 1$) и, согласно (12), $\alpha\neq 0$, а это, в свою очередь, в силу $(I_n(x), I_{nT}(x)) = 1$, и (10) дает соотношение $\alpha\leqq 1$, позволяющее заключить, что $\alpha=1$ и следовательно $\beta=T^{-1}(2^m-1)$.

Если $g(x) = x+1$, то $m=1$, $T=2$ и, согласно (12), получим

$$2 = \alpha+2\beta. \tag{13}$$

Исследуя уравнение (13) относительно переменных α и β, находим два его возможные решения в целых неотрицательных числах: $\alpha=2$, $\beta=0$ и $\alpha=0$, $\beta=1$. В итоге после соответствующей замены получим формулу (7). Аналогично доказывается справедливость (7) и при $n=1$.

Пусть теперь — целое n представимо в виде произведения h ($h<1$) любых равных или неравных сомножителей. При $h=1$ справедливость формулы (10) доказана. Примем, что она имеет место для чисел, меньших h, т. е. для всякого целого неотрицательного числа, представимого в виде произведения самое большее $h-1$ равным или неравным сомножителей. Тогда, имея

$$x^{2^n}+x = x(x+1)I_n(x)\prod_{1<v|n,\,v\neq n} I_v(x) \tag{14}$$

и заменяя в нем x через $\sigma_g(x, \delta)$, получим

$$x\left[x^{2^n-1}(L^x g(x))^{2^n} + L^x g(x)\right] =$$
$$= (xL^x g(x))(xL^x g(x)+1) I_n(\sigma_g(x, \delta)) \prod_{1<v|n,\, v\neq n} I_v(\sigma_g(x, \delta)).$$

Отсюда, согласно допущению и в силу леммы 1, находим

$$(15)\quad L^x(x^n+1)g(x) = (L^x(x+1)g(x)) I_n(\sigma_g(x, \delta)) \prod_{1<v|n,\, v\neq n} I_v(x) I_{Tn}^{(T^{-1}(2^m-1)t(n))}(x).$$

Ясно, ввиду

$$L^x(x^n+1) | L^x(x^n+1)g(x), \quad (n, T)=1 \quad \text{и} \quad L^x(x^n+1)g(x) | x^{2^{nT}-1}+1,$$

что

$$(16)\quad L^x(x^n+1)g(x) = \Phi_1(x)(L^x(x+1)g(x)) I_n(x) \prod_{1<v|n,\, v\neq n} I_v(x) I_{vT}^{(T^{-1}(2^m-1)t(n))}(x),$$

где $\Phi_1(x) = \prod_{v|nT} I_v^{(\varepsilon_v^1)}(x)$, ε_v^1 — целые неотрицательные числа.

В дальнейшем, рассуждая аналогично предыдущему, можно показать, что $\varepsilon_v^1 = 0$ $(v \neq nT)$. А это, в свою очередь, согласно (15) и (16), даст соотношение

$$I_n(\sigma_g(x, \delta)) = I_n(x) I_{nT}^{(T^{-1}(2^m-1)t(n))}(x),$$

позволяющее заключить, что

$$f(\sigma_g(x, \delta)) = I_n^{(l_1)}(x) I_{nT}^{(T^{-1}(2^m-1)l_2)}(x).$$

Этим полностью завершено доказательство теоремы. Опираясь на теорему 5 можно получить следующий результат.

Пусть $f(x)$ и $g(x) \neq x+1$-два примитивных полинома, удовлетворяющих условию теоремы 5. Тогда,

Теорема 6. 1) *Сравнение*

$$(17)\qquad \sigma_g(x^\tau, \delta) \equiv x \pmod{f(x)},$$

всегда разрешимо относительно τ;

2) *Полином степени* $n(2^m-1)$

$\psi(x)^{-1} f(\sigma_g(x, \delta))$, *где* τ — *решение* (17) *и* $\psi(x)$ — *минимальная функция для элемента* $\zeta^\tau (f(\zeta)=0)$, *неприводим.*

Доказательство. В самом деле, пусть $\zeta(f(\zeta)=0)$ примитивный элемент в расширении поля $GF(2)$ и пусть, кроме того, ζ_1 является корнем $f_1(x)$ — неприводимого делителя n-й степени выражения $f(\sigma_g(x, \delta))$. Согласно теореме 5, для любой простой функции $f(x)$ такой полином всегда существует. Далее, в силу цикличности мультипликативной группы поля Галуа, находим $\zeta_1 = \zeta^v$, где v-некоторое целое число. Отсюда автоматически следует существование решения (17). Таким образом, первое утверждение теоремы доказано.

Далее, поскольку неприводимый полином $f_1(x)$ степени n имеет в качестве своего корня элемент ζ_1, то, следовательно, он совпадает с минимальной функцией для ζ_1. При этом, так как полином $g(x) \neq x+1$ примитивный, то

$T = 2^m - 1$ и, согласно теореме 5, выражение $f(\sigma_g(x, \delta))$ разлагается на два простых сомножителя. Отсюда, в частности, так как $f(\sigma_g(x, \delta))$ делится на $\psi(x)$, следует, что полином $\psi(x)^{-1} f(\sigma_g(x, \delta))$ степень которого равна nT, неприводим. Итак, теорема доказана полностью.

Рассмотрим следующий пример.

Пусть даны два примитивных полинома $g(x) = x^2 + x + 1$ и $f(x) = x^4 + x + 1$. Находим, что $\tau = 2$ является корнем сравнения $\sigma_g(x^\tau, 0) = x^{4\tau} + x^{2\tau} + x^\tau \equiv x \pmod{f(x)}$. Поэтому $\psi(x) = f(x) = x^4 + x + 1$ и, согласно теореме 6, полином

$$\frac{f(\sigma_g(x, 0))}{x^4 + x + 1} = x^{12} + x^9 + x^8 + x^6 + x^3 + x^2 + 1 \text{ —}$$

— неприводим.

Пусть $g(x)$ — примитивный полином m-й степени и $\varepsilon > 0$ — любое целое число. Тогда,

Теорема 7. *Всякий простой делитель функции* $\sigma_{x^\varepsilon g+1}(x, \delta)$, $f(x)$, *удовлетворяющий условию* $(\text{ст. } f(x), 2^m - 1) = 1$, *делит выражение* $f(\sigma_g(x, \delta))$; *при этом частное* $f(x)^{-1} f(\sigma_g(x, \delta))$ *неприводимо.*

Доказательство. Подставляя в $\sigma_g(x^\tau, \delta)$ вместо τ выражение 2^ε, будем иметь $\sigma_g(x^{2^\varepsilon}, \delta) = \sigma_{x^\varepsilon g+1}(x, \delta) + x$. Между тем, согласно условию, $(\text{ст. } f(x), 2^m - 1) = 1$, и, кроме того полином $f(x)$ делит $\sigma_{x^\varepsilon g+1}(x, \delta)$, т. е. $\sigma_g(x^{2^\varepsilon}, \delta) \equiv x \pmod{f(x)}$. А это значит, что 2^ε является корнем (17). Исходя из этого, в силу теоремы 6, заключем, что функция $\psi(x) = f(x)$ является делителем $f(\sigma_g(x, \delta))$ и при этом выражение $f(x)^{-1} f(\sigma_g(x, \delta))$ неприводимо. Теорема доказана.

Замечание 1. Нетрудно показать, что условия, фигурирующие в теореме 7, являются также и необходимыми, т. е. можно утверждать следующее:

Для того, чтобы полином $f(x)^{-1} f(\sigma_g(x, \delta))$ *был неприводим, необходимо и достаточно, чтобы* $(\text{ст. } f(x), 2^m - 1) = 1$, *а также существовало натуральное число* ε, *удовлетворяющее условию* $f(x) | x(L^x x^\varepsilon g(x) + 1) + \delta$.

Таким образом, например, полагая $g(x) = x^2 + x + 1$, $\varepsilon = 1$ находим, что $x^2 + x + 1 | L^x x(x^2 + x + 1) + 1$ $(f(x) = x^2 + x + 1)$ в то время как $(3, 2) = 1$. Следовательно, полином шестой степени $(x^2 + x + 1)^{-1}((x^4 + x^2 + x + 1)^2 + (x^4 + x^2 + x + 1) + 1) = x^6 + x^5 + x^3 + x^2 + 1$ — неприводим. Между тем, в рассмотренном примере, если в качестве $f(x)$ взять функцию $x^4 + x^3 + 1$, то, как легко убедиться, ни для одного значения ε полином $f(x)$ не будет делить выражение $L^x x^\varepsilon g(x) + 1$.

Поэтому $f(x) \nmid f(\sigma_g(x, 0))$. Однако, корнем сравнения (17) при $g(x) = x^2 + x + 1$ и $f(x) = x^4 + x^3 + 1$ является $\tau = 9$, значит $\psi(x) = x^4 + x^3 + x^2 + x + 1$ и, следовательно, так как $(3, 4) = 1$, полином 12-й степени —

$$\frac{(x^4 + x^2 + x)^4 + (x^4 + x^2 + x)^3 + 1}{x^4 + x^3 + x^2 + x + 1} = x^{12} + x^{11} + x^8 + x^4 + x^3 + x + 1$$

неприводим.

Пусть $g(x)$ — примитивный трехчлен $g(x) = x^m + x^k + 1$ и $(m, 2^m - 1) = 1$. Тогда,

Теорема 8. *Всякий неприводимый полином* $f(x)$ *степени* m *является делителем выражения* $f(\sigma_g(x, 0))$, *при этом частное* $f(x)^{-1} f(\sigma_g(x, 0))$ *неразложимо.*

Доказательство. В самом деле, полагая $\varepsilon = m-k$, можно написать

$$\sigma_{x^\varepsilon g+1}(x, 0) = x\big(L^x x^\varepsilon g(x)+1\big) = xL^x(x^m+1)(x^{m-k}+1).$$

Поэтому ясно, что для любой неразложимой функции $f(x)$ степени m

$$f(x)|\sigma_{x^\varepsilon g+1}(x, 0).$$

Кроме того $(m, 2^m-1) = 1$. Поэтому согласно теореме 7, функция $f(x)$ делит $f\big(\sigma_g(x, 0)\big)$ и частное $f(x)^{-1}f\big(\sigma_g(x, 0)\big)$ неприводимо. Теорема доказана.

Таким образом, например, полином степени 60

$$\frac{(x^{16}+x^8+x)^4+(x^{16}+x^8+x)+1}{x^4+x+1} = (x^{11}+x^8+x^7+x^5+x^3+x^2+x+1)\times$$

$$\times(x^{49}+x^{34}+x^{19}+x^{17}+x^4+x^2+x)+x^4+x+1,$$

неприводим, так как $g(x) = x^4+x+1$ — примитивный трехчлен и полином $f(x) = x^4+x+1$ неприводим.

Полученные результаты находят интересную интерпретацию для случая производного поля. Приведем некоторые из них.

Пусть r — произвольное натуральное число, δ — любой элемент поля $GF(q^r)$; $g(x)\neq x-1$ неприводимый полином степени m над полем $GF(q)$, $f(x)$ — неприводимый полином степени n над полем $GF(q^r)$, и пусть кроме того $(nr, T)=1$, где T — показатель $g(x)$ и $\psi(x)$ — минимальная функция элемента ζ^τ, где ζ корень $f(x)=0$, а τ решение (17). Тогда

Теорема 9. *Полином*

$$\psi(x)^{-1}f\big(\sigma_g(x, \delta)\big),$$

является произведением $T^{-1}(q^m-1)$ *различных неприводимых в поле* $GF(q^r)$ *делителей степени* nT.

Теорема 9 имеет важные в приложении следствия.

Следствие 1. Для того, чтобы выражение $\psi(x)^{-1}f\big(\sigma_g(x, \delta)\big)$ не разлагалось в поле $GF(q^r)$ необходимо и достаточно, чтобы $g(\chi)$ был примитивным полиномом над полем $GF(q)$.

Как тривиальный частный случай следствия 1 при $f(x)=x$ и $\delta=0$ автоматически вытекает известная теорема Уре ([2], теорема 10) и Глисона—Марша [1], или же, при $f(x)=x$ и $g(x) = x-\theta$, где θ — первообразный элемент поля $GF(q)$, теорема Диксона [1], теорема 5—24).

Следствие 2. Пусть $g(x) = x^m-x^k-1$ примитивный трехчлен над полем $GF(q)$, $r = (m-k)N^{-1}$ — целое и $(r, q^m-1) = 1$. Тогда выражение $f(x)^{-1}f\big(\sigma_g(x,0)\big)$ является неразложимым полиномом над полем $GF(q^r)$ при всякой непроводимой в $GF(q^r)$ функции $f(x)$ степень n которой удовлетворяет условиям $n|N$ и $(n, q^m-1) = 1$.

Следствие 3. Пусть $g(x) = x^m+x^k-1$ примитивный трехчлен над полем $GF(q)$ и r — любой, взаимно простой с q^m-1 делитель (m, k). Тогда выражение $f(x)^{-1}f\big(\sigma_g(x, 0)\big)$ является неразложимым полиномом над полем $GF(q^r)$, при

всякой непроводимой в $GF(q^r)$ функции $f(x)$ степень n, которой удовлетворяет условиям $n|mr^{-1}$ и $(n, q^m-1) = 1$.

В близком отношении к теореме 9 стоит также и следующая

Теорема 10. *Пусть* $q>2$, $(n, q-1) = 1$, $f(x)$ — *неприводимый в* $GF(q^r)$ *полином степени* n, θ — *первообразный элемент поля* $GF(q^r)$ *и* δ — *произвольный элемент* $GF(q^r)$. *Тогда полином*

$$\psi(x)^{-1} f(\sigma_{x-\theta}(x, \delta))$$

степени $n(q-1)$ *неприводим в поле,* $GF(q^r)$.

Этот результат является прямым обобщением указанной выше теоремы Диксона [1], в которой приводится аналогичное утверждение, лишь для частного случая $f(x)=x$.

§ 3. Квадратичные отображения

Преобразование

$$f(x) \to f(x^2+x+\delta),$$

где δ — любой элемент поля $GF(2)$, условимся называть квадратичным отображением полинома $f(x)$[1].

В этом параграфе ставится задача исследовать алгебраическую структуру функции $f(x^2+x+\delta)$ и, тем самым расширить возможности построения неприводимых полиномов.

Если $f(x)=f_1(x)f_2(x)$ то после соответствующей замены в нем x через $x^2+x+\delta$, как легко убедиться, получаем тождество:

$$f(x^2+x+\delta) = f_1(x^2+x+\delta)f_2(x^2+x+\delta).$$

Отсюда ясно, что структура полинома $f(x^2+x+\delta)$ целиком зависит от структуры соответствующих квадратичных отображений неприводимых делителей функции $f(x)$. Поэтому в дальнейшем всюду будем предполагать, что полином $f(x)$ степени n неприводим.

Имеет место следующее альтернативное утверждение.

Теорема 11. *Функция* $f(x^2+x+\delta)$ *либо неразложима, либо является произведением двух различных простых полиномов степени* n; *при этом вероятность того, что выражение* $f(x^2+x+\delta)$ *неприводимо, равна*

$$p(n) = \frac{1}{2nt(n)} \sum_{u|n,\, 2\nmid u} \mu(u) 2^{\frac{n}{u}}.$$

Доказательство. Предварительно рассмотрим случай, когда n — натуральное нечетное число. Но тогда будут удовлетворяться условуя теоремы 5, так как $x^2+x+\delta = xL^x(x+1)+\delta$ и $(n, T)=1$. А поэтому выражение $f(x^2+x+\delta)$ может оказаться либо неприводимым, либо разлагаться на произведение двух

[1] В более общем случае поля характеристика p, аналогом квадратичных отображений является преобразование $f(x)\to f(x^p-x+\delta)\,\delta\in GF(q)$

простых функций степени n. Кроме того, поскольку $p(n)=\frac{1}{2}(2\nmid n)$, $T=2$ и $2^m-1=1$, то, согласно (10), приходим к выражению

$$I_n(x^2+x+\delta)=I_n(x)I_{2n}^{(p(n)t(n))}(x),$$

или, что то же самое

$$I_n(x^2+x+\delta)=I_n^{(\bar{\Delta}(n))}(x)I_{2n}^{(\Delta(n))}(x), \tag{18}$$

где

$$\Delta(n)=p(n)t(n), \quad \bar{\Delta}(n)=2(1-p(n))t(n).$$

Из (18) непосредственно следует, что вероятность того, что функция $f(x^2+x+\delta)$ является простой, равна 1/2. Этот результат находится в согласии с утверждением теоремы. Таким образом, для любого натурального нечетного числа n теорему можно считать доказанной.

Пусть теперь $n=2^hN$ $(h>0, 2\nmid N)$. Предположим, что для любого натурального числа k, удовлетворяющего условиям $k<n$ и $k|2n$ формула (18) верна. Рассмотрим равенство

$$x^{2^n}+x=x(x+1)I_n(x)\prod_{1<v|n,\, v\neq n}I_v(x),$$

которое после замены в нем x на $x^2+x+\delta$ даст соотношение

$$L^x(x^n+1)(x+1)=\big(L^x(x^2+1)\big)I_n(x^2+x+\delta)\prod_{1<v|n,\, v\neq n}I_v(x^2+x+\delta). \tag{19}$$

Как легко понять, квадратичные отображения двух произвольных (но не равных между собой) простых функций не могут иметь общего делителя, отличного от 1. Поэтому в силу принятого соглашения и учитывая, что $\Delta(k)+\bar{\Delta}(2k)=t(2k)$ $(k>0)$, можно написать формулу

$$I_{\frac{k}{2}}(x^2+x+\delta)I_k(x^2+x+\delta)=I_{\frac{k}{2}}^{\left(\bar{\Delta}\left(\frac{k}{2}\right)\right)}(x)I_k(x)I_{2k}^{(\Delta(k))}(x), \tag{20}$$

справедливую для любого четного k, удовлетворяющего условиям $k<n$ и $k|2n$.

Из (10), (18), (19) и принимая во внимание, что $\bar{\Delta}(k)=t(k)$ $(2\nmid k)$, $\bar{\Delta}(2)=0$ находим

$$L^x(x^n+1)(x+1)=$$

$$=\big(L^x(x^2+1)\big)I_n(x^2+x+\delta)I_n^{\left(\Delta\left(\frac{n}{2}\right)\right)}(x)\prod_{u|N,\, u\neq N}I_{2^{h+1}u}^{(\Delta(2^hu))}(x)\prod_{2<u|n,\, u\neq n}I_u(x). \tag{21}$$

Дальнейшие рассуждения разделяются на два варианта в зависимости от того $N>1$ или $N=1$.

При $N>1$ ясно, что для любого u, удовлетворяющего условия $u|N$ и $u\neq N$, будем иметь $2^{h+1}u<n$ и $2^{h+1}u|2n$. А это, согласно (20), позволяет заключить, что

$$I_{2^{h+1}u}(x)|I_{2^hu}(x^2+x+\delta)I_{2^{h+1}u}(x^2+x+\delta)$$

и следовательно,

$$(I_n(x^2+x+\delta),\ \prod_{u|N,\ u\neq N} I_{2^{h+1}u}(x)) = 1. \tag{22}$$

Согласно (21), (22), лемме 2 и с учетом, что

$$L^x(x^n+1)(x+1)|L^x(x^{2n}+1),\quad L^x(x^2+1) = (x+1)I_2(x),$$

можно написать соотношение

$$I_n(x^2+x+\delta)\, I_n^{\left(\Delta\left(\frac{n}{2}\right)\right)}(x) = I_u(x)\, I_{2n}^{(A)}(x),$$

откуда, ввиду

$$A = \frac{1}{2}\left(t(n)+\Delta\left(\frac{n}{2}\right)\right) = \Delta(n),$$

окончательно получим

$$I_n(x^2+x+\delta) = I_n^{(\Delta(n))}(x)\, I_{2n}^{(\Delta(n))}(x),$$

т. е. формулу (18).

Если же $N=1$, то тогда $\prod_{u|N,\ u\neq N} I_{2^{h+1}u}(x) = 1$ и поэтому автоматически будет выполняться соотношение (22), которое, в свою очередь, также приведет нас к выражению (18).

Между тем, соотношение (18) справедливо для любого нечетного n. Следовательно, оно имеет место для любого натурального числа.

Рассмотрим теперь уравнение $2 = \alpha+2\beta$. Как легко видеть, оно имеет всего лишь два различных неотрицательных целочисленных решения: $\alpha=0$, $\beta=1$ и $\alpha=2$, $\beta=0$. Поэтому, согласно (18), заключаем, что функция $f(x^2+x+\delta) = I_n^{(\alpha)}(x)I_{2n}^{(\beta)}(x)$ может быть либо неприводимой, либо разлагается на произведение двух различных простых сомножителей степени n. При этом вероятность, с которой функция $f(x^2+x+\delta)$ может оказаться неразложимой, равна

$$p(n) = \frac{1}{2nt(n)} \sum_{u|N} \mu(u) 2^{\frac{n}{u}}.$$

Итак теорема 11 полностью доказана.

Условимся обозначать через $\pi(f)$ коэффициент при неизвестном x^{n-1} полинома $f(x)$, n-й степени.

Теорема 12. *Квадратичное отображение любой неприводимой функции $f(x)$ степени n неразложимо, тогда и только тогда, если выполняется условие*

$$\pi(f)+n\delta = 1. \tag{23}$$

Доказательство. Предположим, что ζ является корнем полинома $f(x)$. Но тогда $\zeta, \zeta^2, \ldots, \zeta^{2^{n-1}}$ суть все корни $f(x)$ и поскольку $\pi(f) = \sum_{v=0}^{n-1} \zeta^{2^v}$, то будем иметь равенство

$$\sigma_{\varrho_n}(\zeta, \delta^1) = \pi(f)+\delta^1, \tag{24}$$

где $\varrho_n(x) = \sum_{u=0}^{n-1} x^u$, δ^1 — произвольный элемент $GF(2)$. Из (24) в частности следу-

ет, что выражение $\sigma_{\varrho_n}(x, \delta^1)$ делится без остатка на функцию $f(x)\big(f(\zeta)=0\big)$ в том и только в том случае, если

$$\pi(f)+\delta^1 = 0. \tag{25}$$

Между тем, как легко заметить,

$$\sigma_{\varrho_n}(x^2+x+\delta, \delta^1) = \sigma_{x^n+1}(x, n\delta+\delta^1). \tag{26}$$

Далее, согласно лемме 1, имеем

$$xL^x(x^n+1)(x+1) = x\big(L^x(x^n+1)\big)\big(xL^x(x^n+1)+1\big) = \sigma_{x^n+1}(x, 0)\sigma_{x^n+1}(x, 1)$$

и поэтому $\sigma_{x^n+1}(x, 1)|L^x(x^{2n}+1)$, в то время как

$$\big(\sigma_{x^n+1}(x, 0), \sigma_{x^n+1}(x, 1)\big) = 1, \quad \text{значит} \quad \big(\sigma_{x^n+1}(x, 1), L^x(x^n+1)\big) = 1.$$

Это позволяет сделать следующий вывод: для того, чтобы функция $f\big(\sigma_g(x, \delta)\big)$ была неразложимой, необходимо и достаточно, чтобы она являлась делителем выражения $\sigma_{x^n+1}(x, 1)$. Но для этого как следует из (26), полином $\sigma_{\varrho_n}(x, \delta^1)$ должен быть кратным функции $f(x)$, или, что тоже самое

$$\pi(f)+\delta^1 = 0. \tag{27}$$

Кроме того, очевидно должно удовлетворяться также и условие $\sigma_{\varrho_n}(x^2+x+\delta, \delta^1) = \delta_{x^n+1}(x, 1)$, которое в свою очередь, в силу (26), эквивалентно равенству

$$n\delta+\delta^1 = 1. \tag{28}$$

Сопоставляя между собой (27) и (28), получаем (23). Таким образом, убеждаемся в правильности утверждения теоремы. Итак, теорема 12 доказана полностью.

Рассмотрим следующий пример.

Пусть $f(x) = x^5+x^2+1$. Решая уравнение $\pi(f)+5\delta = 0+\delta = 1$ относительно δ, находим, что $\delta=1$. Следовательно, полином десятой степени $(x^2+x+1)^5+(x^2+x+1)^2+1 = x^{10}+x^9+x^8+x^6+x^5+x+1$ неприводим.

Сопоставляя между собой теоремы 10 и 11 можно получить следующий результат

$$N_n\big(\pi(f) = 1\big) = \frac{1}{2n}\sum_{u|n\,2\nmid u}\mu(u)2^{\frac{n}{u}}, \tag{29}$$

где $N_n(a_{u_i}=\delta_i)$ — общее число всевозможных неприводимых в поле $GF(2)$ полиномов $f(x) = \sum_{u=0}^{n} a_u x^u$ степени n, коэффициенты a_{u_i}, которых принимают заранее заданные значения $\delta_i \in GF(2)$.

Исследование числовых функций $N_n(\cdot)$ является чрезвычайно важной и актуальной задачей теории приводимости. Однако несмотря на это не существуют по-видимому конечных формул позволяющих вычислять точное значение функции $N_n(\cdot)$ при различных их параметрах, (кроме, разве лишь, тривиального случая ($t=0$) т. е. формула Дедекинда). Причем, получение таких формул представляет интерес также и в связи с возможностью использовать их косвенно в решении задачи синтеза неприводимых полиномов методом

„проб и ошибок". Однако, этот вопрос лежит несколько в стороне от затронутой нами темы. Поэтому, не задерживаясь на нем приведем без доказательства одну рекуррентную формулу общего характера, связывающую между собой эти функции.

Для любых натуральных $t, n \geqq t$ и произвольного элемента δ поля $GF(2)$, выполняется соотношение

$$\sum_{i=0}^{t-1} 2^i \left(2^{n-t} - \sum_{km=n-i} kN_k\left(M(m, t-j) = \delta(j)(j = 0, i)\right)\right) = 0 \tag{30}$$

где

$$M(m, u) = \sum_{\alpha_1+2\alpha_2+\ldots+u\alpha_u=u} \frac{m! \prod_{i=1}^{u} a_i^{\alpha_i}}{(m-\alpha_1-\ldots-\alpha_u)!\,\alpha_1!\ldots\alpha_u!}; \quad \delta(j) = \begin{cases} \delta & \text{если } j=0, \\ 0 & \text{если } j>0. \end{cases}$$

Опираясь на формулу (30) можно получать соответствующие выражения для различных значений функций $N_n(\cdot)$. Так, например, полагая в формуле (30) $t=1$ и, используя формулу обращения Дедекинда—Лиувилля, получим соотношение (29) или же при $t=2$, $2 \nmid n$ будем иметь

$$N_n(a_1=\delta_1, a_2=\delta_2) = \frac{1}{n} \sum_{su=n} \mu(u)\left(2^{\frac{s-1}{2}} + (-1)^{\frac{s^2-1}{8}+\delta_1\frac{n-1}{2}+\delta_2}\right)2^{\frac{s-3}{2}}.$$

В заключении отметим, что в произвольном поле $GF(q)$ содержание теорем 11 и 12 можно будет выразить в следующей форме.

Теорема 13. Функция $f(x^p-x+\delta)$ либо неприводимо в поле $GF(q)$, либо разлагается на произведение p неприводимых множителей степени n. Причем вероятность того, что выражение $f(x^p-x+\delta)$ может оказаться неприводимым, равна

$$\frac{p-1}{p}\left(\sum_{u|n,\ p\nmid u} \mu(u) q^{\frac{n}{u}}\right)\left(\sum_{u|n} \mu(u) q^{\frac{n}{u}}\right)^{-1}.$$

Теорема 14. Для того, чтобы полином $f(x^p-x+\delta)$ был неприводим в поле $GF(q)$, необходимо и достаточно, чтобы выполнялось условие

$$\sum_{u=0}^{s-1} \left(n\delta + \pi(f)\right)^{p^u} \neq 0 \qquad (q=p^s).$$

В частности, теорема 14 является существенным уточнением известного результата (см. [1] теорема (5—23)) который устанавливает лишь достаточное условие неприводимости функции $f(x^p-x)$ $(\delta=0)$.

ЛИТЕРАТУРА

[1] Albert, A. A.: *Fundamental contcepts of higher algebra.* Univ. of Chicago Press, 1956.
[2] Zierler, N.: Linear recurring sequency *I. Soc. Ind. Appl. Math.* **7** (1959), № 1, 31—48.

Ереван, Вычислительный Центр АН АРМ СССР

(Поступила 12-ого Января, 1971 г.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 21—28.

ON THE RAREFACTION OF RENEWAL PROCESSES. III

by

J. MOGYORÓDI

In this paper the definitions, notions, notations and results of the previous ones will be used.

5. If the mean-value μ of the original process exists, then in every step of the rarefaction we have

$$M\left(\frac{\tau_i^{(n)} - \tau_{i-1}^{(n)}}{M^n}\right) = \mu,$$

($n=1, 2, \ldots$; $i=1, 2, \ldots$). (See paper I.) But the variance, if exists, and a fortiori the distribution function is, in general, different from those of the original process.

The problem arises what are the original processes for which

$$\frac{\tau_i^{(1)} - \tau_{i-1}^{(1)}}{M} \qquad (i = 1, 2, \ldots)$$

has the same distribution function as the original process. In that case we say that the original process is invariant with respect to the rarefaction.

It is obvious that if the original process is invariant, then for every n the distribution function of the random variables

$$\frac{\tau_i^{(n)} - \tau_{i-1}^{(n)}}{M^n} \qquad (i = 1, 2, \ldots)$$

will be the same as that of the original renewal process.

Let us denote by $F(x)$, resp. $\psi(s)$, the distribution function and the Laplace transform of the original process. If this is invariant, then

$$F(x) = \sum_{k=1}^{\infty} p_k F^{(*k)}(Mx),$$

or, in the language of the Laplace transforms,

$$\psi(s) = f\left(\psi\left(\frac{s}{M}\right)\right).$$

This can also be written in the form

$$\psi(s) = f_n\left(\psi\left(\frac{s}{M^n}\right)\right), \qquad (n = 1, 2, \ldots)$$

which means that

$$\psi(s) = \lim_{n \to +\infty} f_n\left(\psi\left(\frac{s}{M^n}\right)\right).$$

By Theorem 7 we conclude that $F(x)$ necessarily has a finite mean-value μ and at the same time $F(x)$ is the limiting distribution. So,

$$\psi(s) = g(\mu s),$$

or, in the language of the distribution functions

$$F(x) = G\left(\frac{x}{\mu}\right),$$

where $G(x)$ is defined as in Lemma 1. (See paper I.)

Conversely, we show that the original process, the distribution function of which is $G\left(\frac{x}{\mu}\right)$, where μ is an arbitrary positive number, is invariant. In fact, the Laplace transform of

$$\frac{\tau_i^{(1)} - \tau_{i-1}^{(1)}}{M} \tag{27}$$

is in this case

$$f\left(g\left(\frac{\mu s}{M}\right)\right).$$

Now $g(s)$ satisfies the so called Poincaré equation:

$$g(s) = f\left(g\left(\frac{s}{M}\right)\right),$$

which implies that the Laplace transform of (27) is $g(\mu s)$.

We have proved

THEOREM 9. *Let us suppose that* $0 < \sigma^2 < +\infty$. *The original process is invariant — i.e. the distribution function of* (27) *is equal to that of the original process — if and only if its distribution function is of the form* $G\left(\frac{x}{\mu}\right)$, *where* $\mu > 0$ *is an arbitrary fixed number and* $G(x)$ *is defined by Lemma 1.*

If the original process is invariant, then for every $n = 1, 2, \ldots$ the distribution function of the random variables

$$\frac{\tau_i^{(n)} - \tau_{i-1}^{(n)}}{M^n}$$

is also $G\left(\frac{x}{\mu}\right)$.

The assertion of Theorem 9 holds also in case $\mu = 0$.

The notion of the invariance can be extended for that case when we do not prescibe the normalizing factor to be $\frac{1}{M^n}$. We call that the original renewal process

is invariant in the general case, if there exists a sequence δ_n of positive numbers, such that for every n the random variables

$$\delta_n(\tau_i^{(n)} - \tau_{i-1}^{(n)})$$

have the same distribution as that of the original renewal process.

If the original renewal process $\tau_0 \equiv 0 \leqq \tau_1 \leqq \tau_2 \leqq \cdots$, with distribution function $F(x)$, is invariant in the general sense, then the Laplace transform $\psi(s)$ of $F(x)$ must satisfy for every n the equality

$$\psi(s) = f_n(\psi(\delta_n s)).$$

From this and from Theorem 4 it follows that

$$\psi(s) = g(Cs^\alpha),$$

where $C>0$ and $0<\alpha \leqq 1$, or

$$\psi(s) \equiv 1.$$

The last case is not interesting. In the first one we see that $F(x)$ is necessarily that distribution function, the Laplace transform of which is $g(Cs^\alpha)$. Let us now deal with the problem of choice of the normalizing factors δ_n. By Theorem 4 the relation

$$\lim_{n \to +\infty} M^n(1 - g(C\delta_n^\alpha s^\alpha)) = Cs^\alpha \quad (s \geqq 0)$$

must hold in case of invariance. For $s=1$ we obtain from this

$$1 - g(C\delta_n^\alpha) \sim \frac{C}{M^n} \quad (n \to +\infty),$$

which means that

$$g(C\delta_n^\alpha) \sim 1 - \frac{C}{M^n}.$$

It follows that

$$C\delta_n^\alpha \sim g^{-1}\left(1 - \frac{C}{M^n}\right). \quad (n \to +\infty).$$

By Taylor expansion of $g^{-1}(x)$ at the point $x=1$, we obtain

$$C\delta_n^\alpha \sim \frac{C}{M^n}, \quad (n \to +\infty).$$

So

$$\delta_n \sim \frac{1}{M^{\frac{n}{\alpha}}}.$$

Conversely we show that if the distribution function of the original renewal process has the Laplace transform $g(Cs^\alpha)$ $(C>0,\ 0<\alpha \leqq 1)$, then this process is invariant in the general sense if we choose the normalizing factors δ_n to be $\frac{1}{M^{\frac{n}{\alpha}}}$. I. e. the distribu-

tion function of the random variables

$$\frac{1}{M^{\frac{n}{\alpha}}}(\tau_i^{(n)}-\tau_{i-1}^{(n)})$$

will be for every n the same as that of the original process. In fact, we have for $g(s)$ $(s\geqq 0)$ the Poincaré equation

$$g(s)=f_n\left(g\left(\frac{s}{M^n}\right)\right).$$

Substituting Cs^α we obtain

$$g(Cs^\alpha)=f_n\left(g\left(\frac{Cs^\alpha}{M^n}\right)\right)=f_n\left(g\left(C\left(\frac{s}{M^{\frac{n}{\alpha}}}\right)^\alpha\right)\right).$$

These considerations imply the following.

THEOREM 10. *Let us suppose that* $0<\sigma^2<+\infty$. *The original renewal process is invariant in the general sense if and only if the Laplace transform of his distribution function is of the form* $g(C\mu s^\alpha)$, *where* $\mu>0$ *is an arbitrary number and* $g(s)$ *is the Laplace transform of the distribution function* $G(x)$, *defined in Lemma 1. Here* $C>0$ *and* $0<\alpha\leqq 1$. *In case of invariance the normalizing factor must be determined by the formula*

$$\delta_n=\frac{1}{\mu^{\frac{1}{\alpha}}M^{\frac{n}{\alpha}}}.$$

Theorem 9 was proved first by the author in [1] under the assumption that the original process has finite mean-value. Using the techniques of functional equations the assertion of Theorem 9 is proved in his general form in Szántai's paper [2]. Our proof is simpler. The result of Theorem 10 is new.

6. Consider now some examples, which are more general than the case of section 3. (See paper I.)

Let for this purpose

$$u_q(s)=\frac{q^s}{1-(1-q)s},\quad |s|\leqq 1,\quad 0<q<1,$$

be the probability generating function of the geometric distribution. Further let us consider a function $h(s)$ such that the function

$$f(s)=h^{-1}(u_q(h(s))),\qquad 0\leqq s\leqq 1,$$

be again a probability generating function. It is obvious that such a $h(s)$ exists.

Since $f(s)$ is a generating function, it follows that $f(s)$ is absolutely monotonic:

$$f^{(n)}(s)\geqq 0\qquad (n=1,2,\ldots)$$

for $0\leqq s\leqq 1$. Consequently, $h(s)$ must have derivatives of every order in $0\leqq s\leqq 1$. By some calculations we can also show that $h(1)=1$ and $h(0)=0$. Further, since $h(s)$ is continuous and has an inverse $h^{-1}(s)$, it follows that $h(s)$ is strictly monotonically increasing in $[0, 1]$ and $h(0)=0$, $h(1)=1$.

Let us put for example $h(s)=s^k$, k being a positive integer. ($k=1, 2, \ldots$.) Then

$$f(s) = \left(u_q(s^k)\right)^{\frac{1}{k}} = \left(\frac{qs^k}{1-(1-q)s^k}\right)^{\frac{1}{k}}$$

is the probability generating function of a discrete distribution. The corresponding random variable takes on the values $nk+1$ with the probability

$$\sqrt[k]{q}(1-q)^n(-1)^n\binom{-\frac{1}{k}}{n},$$

$n=0, 1, 2, \ldots$. (Cf. HARRIS [5], page 23.)

It can be shown that if

$$f(s) = h^{-1}\left(u_q(h(s))\right)$$

is a generating function, then

$$f_n(s) = h^{-1}\left(u_{q^n}(h(s))\right) \qquad (n = 1, 2, \ldots),$$

where $f_n(s)$ is the n-th iterate of the generating function $f(s)$.

Let $v_i^{(n)}$ ($i, n=1, 2, \ldots$) be a sequence of independent and identically distributed random variables, taking on as values the positive integers, with the common probability generating function $f(s)$ ($0 \leqq s \leqq 1$). Consider the rarefaction model, described in Paper I, with these random variables.

Let $\psi(s)$ be the Laplace transform of the distribution function of the original renewal process $\tau_0 \equiv 0 \leqq \tau_1 \leqq \tau_2 \leqq \cdots$. Then the Laplace transform of the random variable

$$\delta_n(\tau_i^{(n)} - \tau_{i-1}^{(n)})$$

will be of the form

$$f_n\left(\psi(\delta_n s)\right) = h^{-1}\left(\frac{q^n h\left(\psi(\delta_n s)\right)}{1-(1-q^n)h\left(\psi(\delta_n s)\right)}\right),$$

or

$$f_n\left(\psi(\delta_n s)\right) = h^{-1}\left(\frac{h\left(\psi(\delta_n s)\right)}{\frac{1-h\left(\psi(\delta_n s)\right)}{1-\psi(\delta_n s)}\,\frac{1-\psi(\delta_n s)}{q^n} + h\left(\psi(\delta_n s)\right)}\right).$$

Since for $\delta_n \to 0$ we have

$$h\left(\psi(\delta_n s)\right) \to 1,$$

further

$$\frac{1-h\left(\psi(\delta_n s)\right)}{1-\psi(\delta_n s)} \to h'(1),$$

we conclude that

$$\lim_{n\to+\infty} f_n\left(\psi(\delta_n s)\right)$$

exists if and only if

$$\lim_{n\to+\infty} \frac{1-\psi(\delta_n s)}{q^n}$$

exists. Now $M=\frac{1}{q}$. So, by Theorem 4 (paper II.) we can say that

$$\lim_{n\to+\infty}\frac{1-\psi(\delta_n s)}{q^n}$$

exists if and only if

$$\lim_{n\to+\infty}\frac{1-\psi(\delta_n s)}{q^n}=Cs^\alpha \quad (C>0,\ 0<\alpha\leqq 1),$$

or

$$\lim_{n\to+\infty}\frac{1-\psi(\delta_n s)}{q^n}=0.$$

In the last case $\lim_{n\to+\infty} f_n(\psi(\delta_n s))=1$, while in the previous one

$$\lim_{n\to+\infty} f_n(\psi(\delta_n s))=h^{-1}\left(\frac{1}{1+Ch'(1)s^\alpha}\right) \quad (s\geqq 0). \tag{28}$$

If, especially, $h(s)=s^k$, then

$$\lim_{n\to+\infty} f_n(\psi(\delta_n s))=\left(\frac{1}{1+kCs^\alpha}\right)^{\frac{1}{k}}. \tag{29}$$

The case $k=1$ corresponds to the rarefraction model, described in section 3, paper I, but now the mean-value μ of the original renewal process is not supposed to be finite. The case $k=1$ was discussed in papers [3], [4]. In [3] KOVALENKO has dealt with the problem of getting an exponential limit distribution, which corresponds in (29) to the case $k=1$, and $\alpha=1$. He stated an assertion, which is not true in general as it has been remarked in [4]. Namely, he asserted that for the existence of an exponential limit distribution in the case $k=1$ it is necessary and sufficient that the original renewal process have finite mean-value. This condition is only sufficient (cf. the result of section 3, paper I). But if we choose the normalizing factor δ_n to be $\frac{1}{M^n}=q^n$, then the finiteness of the mean-value of the original renewal process is at the same time necessary. This is the assertion of Theorem 7. (See paper II.) If the value of δ_n is not prescribed, then the condition of Theorem 6 is the necessary and sufficient condition. Comparing the results of section 4, Theorem 6 and Theorem 7 we can make the following assertion:

THEOREM 11. *If $\nu_i^{(n)}$ is distributed for every i and n according to the geometric distribution with the parameter $0<q<1$, then the random variables*

$$\delta_n(\tau_i^{(n)}-\tau_{i-1}^{(n)})$$

have for fixed i and $n\to+\infty$ an exponential limit distribution if and only if for the distribution function $F(x)$ of the original renewal process the limit relation holds:

$$\lim_{x\to+\infty}\frac{x(1-F(x))}{\int_0^x(1-F(z))\,dz}=0.$$

If we prescibe the value of δ_n to be q^n, then the variables

$$q^n(\tau_i^{(n)} - \tau_{i-1}^{(n)})$$

have a limiting distribution if and only if the mean-value of the distribution function is finite and positive. In this case the limit distribution will also be exponential.

In proving Theorem 3 (paper I) we have used the condition that the distribution function $G(x)$ has bounded density function. The example of section 3 shows that such a $G(x)$ exists. The following example shows another distribution with the same property. Let for this purpose

$$h^{-1}(s) = se^{\lambda(s-1)} \quad (0 < s \leqq 1),$$

where $\lambda > 0$ is a fixed number. The function $h^{-1}(s)$ is the probability generating function of the random variable $\xi + 1$, where ξ is a random variable distributed according to the Poisson law with parameter λ. $h^{-1}(s)$ is absolutely monotonic and for every n the derivative $d^n h^{-1}(s)/ds^n$ is positive for $s > 0$. $h(s)$, being the inverse of $h^{-1}(s)$, has thus positive derivatives for $0 < s < 1$. From these it follows that

$$h^{-1}(u_q(h(s)))$$

is a generating function. The corresponding limit distribution, according to (28) is such that its characteristic function at the point t is

$$\frac{1}{1 - ih'(1)t} e^{\lambda\left[\left(\frac{1}{1-ih'(1)}t\right)-1\right]},$$

where $h'(1) = \dfrac{1}{1+\lambda}$. Expanding this in a series we obtain

$$e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!\left(1 - \dfrac{i}{1+\lambda}t\right)^{k+1}},$$

from which the corresponding density function is

$$e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k (\lambda+1)^k x^k e^{-(\lambda+1)x}}{(k!)^2} = e^{-\lambda} e^{-(\lambda+1)x} I_0\left(2\sqrt{\lambda(\lambda+1)x}\right),$$

where

$$I_0(x) = \sum_{k=0}^{\infty} \frac{1}{(k!)^2}\left(\frac{x}{2}\right)^{2k}$$

is the modified Bessel function of order 0. ($x \geqq 0$.) This density function, is for $\lambda < 1$, bounded. In fact, it is less than

$$(\lambda+1)e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda(\lambda+1)x)^k}{k!} e^{-(\lambda+1)x} = (\lambda+1)e^{-\lambda} e^{-(\lambda+1)x} e^{\lambda(\lambda+1)x},$$

which, in turn, is less than $(\lambda+1)e^{-\lambda}$, if $\lambda < 1$.

REFERENCES

[1] MOGYORÓDI, J.: A rekurrens folyamat ritkításáról. *MTA. III. Osztály Közleményei.* **19** (1969), 25—31.
[2] SZÁNTAI, T.: Véletlen pontfolyamatok ritkítása. Thesis. Budapest, 1970.
[3] Коваленко, И. Н: О классе предельных распределений для редеющих потоков однородных событий. *Литовский Математический сборник* **5** (1965), 569—573.
[4] Гнеденко. Б. В, Фрайер, Б.: Несколько замечаний к одной задаче И. Н. Коваленко. *Литовский математический сборник* **9** (1969), 463—470.
[5] Харрис, Т.: *Теория ветвящихся случайных процессов.* Издательство „Мир". Москва, 1966.

Eötvös L. University, Budapest

(Received March 11, 1971.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 29—38.

ON THE RAREFACTION OF RENEWAL PROCESSES. IV

by

J. MOGYORÓDI

7. Now we shall be interested in the rarefaction of the process

$$\tau_n = \sum_{i=1}^{n} \xi_i \qquad (n = 1, 2, \ldots)$$

where the random variables ξ_i are independent and identically distributed variables which can take on arbitrary (not only non-negative) values. The rarefaction will be made according to the model described in paper I, section 2. So, we can define the random variables

$$\tau_0^{(n)} \equiv 0, \quad \tau_i^{(n)} = \tau_{Z_1^{(n)} + \ldots + Z_i^{(n)}}, \qquad (i = 1, 2, \ldots).$$

But the intuitional meaning of the rarefaction is lost in this case. Of course, the random variables

$$\tau_i^{(n)} - \tau_{i-1}^{(n)} \qquad (i = 1, 2, \ldots)$$

are for a fixed n independent and identically distributed. So, we can deal with only the random variable

$$\tau_1^{(n)} = \tau_{Z_1^{(n)}}.$$

The first assertion we prove is

THEOREM 12. *If $M(\xi_i)=\mu$ exists, then*

$$\lim_{n \to +\infty} \mathsf{P}\left(\frac{\tau_{Z_1^{(n)}}}{M^n} < \mu x\right) = \begin{cases} G(x), & \text{if} \quad \mu > 0 \\ 1 - G(-x), & \text{if} \quad \mu < 0, \end{cases}$$

further, if $\mu=0$ and $D^2(\xi_i)=d^2$ exists, we have

$$\lim_{n \to +\infty} \mathsf{P}\left(\frac{\tau_{Z_1^{(n)}}}{M^n} < x\right) = \int_0^{+\infty} \Phi\left(\frac{x}{d\sqrt{z}}\right) dG(z),$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{y^2}{2}\right) dy$$

is the standard normal distribution and $G(x)$ is the distribution function, defined in Lemma 1, Paper I.

PROOF. Taking into account the formula (cf. paper I, section 2)

$$\frac{\tau_{Z_1^{(n)}}}{M^n} = \sum_{j=1}^{Z_1^{(n)}} \xi_j / M^n,$$

our first limit relation follows immediately from Theorem 1. Concerning the second limit relation we refer the reader to [1]. This proves the theorem.

Now we shall be concerned in the rate of convergence in Theorem 12. The following assertion can be proved similarly as that of Theorem 3. (Paper I.)

THEOREM 13. *Let us suppose that the distribution function $F(x)$ of the random variables ξ_i is such that*

$$\int_{-\infty}^{+\infty} x^2 dF(x) < +\infty$$

and $\mu \neq 0$, further $G(x)$ has bounded density. Then for the rarefaction model described in section 2 (paper I) we have

$$\left| \mathsf{P}\left(\frac{\tau_{Z_1^{(n)}}}{\mu M^n} < x \right) - H(x) \right| \leqq K M^{-\frac{n}{3}} \quad (x \geqq 0),$$

where K is a constant not depending on n and $H(x)$ denotes the distribution function $G(x)$ or $1-G(-x)$ according to the sign of μ.

If $\mu = 0$, then we have the following

THEOREM 14. *If $\mu = 0$, $D^2(\xi_i) = 1$ and*

$$\int_{-\infty}^{+\infty} |x^3| dF(x) < +\infty$$

further $G(x)$ has bounded density, then there exists a constant $K' > 0$ such that

$$\left| \mathsf{P}\left(\frac{\tau_{Z_1^{(n)}}}{\sqrt{M^n}} < x \right) - \int_0^{+\infty} \Phi\left(\frac{x}{\sqrt{z}} \right) dG(z) \right| \leqq K' \max\left(M\left(\frac{1}{\sqrt{Z_1^{(n)}}} \right), M^{-\frac{n}{3}} \right).$$

PROOF. By the theorem of total probabilities we have

$$\mathsf{P}\left(\frac{\tau_{Z_1^{(n)}}}{\sqrt{M^n}} < x \right) = \sum_{k=1}^{\infty} \mathsf{P}(z_1^{(n)} = k) \mathsf{P}\left(\frac{\tau_k}{\sqrt{k}} < \frac{x}{\sqrt{k}} \sqrt{M^n} \right).$$

For every k we can now apply the Berry—Esseen theorem according to which for every fixed $k = 1, 2, \ldots$ we have

$$\sup_{-\infty < y < +\infty} \left| \mathsf{P}\left(\frac{\tau_k}{\sqrt{k}} < y \right) - \Phi(y) \right| \leqq \frac{C}{\sqrt{k}},$$

where C is a positive constant. Applying this we have

$$\left| \mathsf{P}\left(\frac{\tau_k}{\sqrt{k}} < \frac{x}{\sqrt{k}} \sqrt{M^n} \right) - \Phi\left(\frac{x\sqrt{M^n}}{\sqrt{k}} \right) \right| \leqq \frac{C}{\sqrt{k}}.$$

Thus

$$\sum_{k=1}^{\infty} \mathsf{P}(Z_1^{(n)} = k)\Phi\left(\frac{x\sqrt{M^n}}{\sqrt{k}}\right) - \sum_{k=1}^{\infty} \mathsf{P}(Z_1^{(n)} = k)\frac{C}{\sqrt{k}} \leqq \mathsf{P}\left(\frac{\tau_{Z_1^{(n)}}}{\sqrt{M^n}} < x\right) \leqq$$

$$\leqq \sum_{k=1}^{\infty} \mathsf{P}(Z_1^{(n)} = k)\Phi\left(\frac{x\sqrt{M^n}}{\sqrt{k}}\right) + \sum_{k=1}^{\infty} \mathsf{P}(Z_1^{(n)} = k)\frac{C}{\sqrt{k}}.$$

Let us introduce the notation

$$G_n(z) = \mathsf{P}\left(\frac{Z_1^{(n)}}{M^n} < z\right).$$

Then

$$\int_0^{+\infty} \Phi\left(\frac{x}{\sqrt{z}}\right) dG_n(z) - CM\left(\frac{1}{\sqrt{Z_1^{(n)}}}\right) \leqq \mathsf{P}\left(\frac{\tau_{Z_1^{(n)}}}{\sqrt{M^n}} < x\right) \leqq$$

$$\leqq \int_0^{+\infty} \Phi\left(\frac{x}{\sqrt{z}}\right) dG_n(z) + CM\left(\frac{1}{\sqrt{Z_1^{(n)}}}\right).$$

Now $Z_1^{(n)}$ converges in probability to $+\infty$ which ensures that $M\left(\frac{1}{\sqrt{Z_1^{(n)}}}\right) \to 0$. Let us further remark that

$$\int_0^{+\infty} \Phi\left(\frac{x}{\sqrt{z}}\right) dG_n(z) = \frac{1}{2} - \int_0^{+\infty} G_n(z)\, d_z \Phi\left(\frac{x}{\sqrt{z}}\right),$$

further that

$$\int_0^{+\infty} \Phi\left(\frac{x}{\sqrt{z}}\right) dG(z) = \frac{1}{2} - \int_0^{+\infty} G(z)\, d_z \Phi\left(\frac{x}{\sqrt{z}}\right).$$

Comparing these we obtain

$$\int_0^{+\infty} (G(z) - G_n(z))\, d_z \Phi\left(\frac{x}{\sqrt{z}}\right) - cM\left(\frac{1}{\sqrt{Z_1^{(n)}}}\right) \leqq$$

$$\leqq \mathsf{P}\left(\frac{\tau_{Z_1^{(n)}}}{\sqrt{M^n}} < x\right) - \int_0^{+\infty} \Phi\left(\frac{x}{\sqrt{z}}\right) dG(z) \leqq$$

$$\leqq \int_0^{+\infty} (G(z) - G_n(z))\, d_z \Phi\left(\frac{x}{\sqrt{z}}\right) + cM\left(\frac{1}{\sqrt{Z_1^{(n)}}}\right).$$

By Theorem 3 of paper I we see that

$$|G(z) - G_n(z)| \leqq KM^{-\frac{n}{3}},$$

and so, because

$$\left| \int_0^{+\infty} d_z \Phi\left(\frac{x}{\sqrt{z}}\right) \right| \leqq 1,$$

we obtain,

$$\left| \mathsf{P}\left(\frac{\tau_{Z_1^{(n)}}}{\sqrt{M^n}} < x \right) - \int_0^{+\infty} \Phi\left(\frac{x}{\sqrt{z}} \right) dG(z) \right| \leqq \max\left(KM^{-\frac{n}{3}},\ cM\left(\frac{1}{\sqrt{Z_1^{(n)}}} \right) \right).$$

This proves our theorem.

If the existence of the first moment of the random variables is not supposed, then we are not able to give a complete characterization of the possible limiting distributions. In this case the apparatus of the Laplace transforms cannot be applied, because the random variables ξ_i can take on arbitrary values. So the apparatus of the characteristic functions will be applied. Now the method of proof of Theorem 4 of paper I is not an appropriate advice, since the inversion of the generating functions for complex arguments establishes some difficulties. We shall give some informations about the possible limit distributions without trying to the complete characterization.

Let δ_n be a sequence of positive numbers tending to zero.

LEMMA 2. *Let us consider the random variables*

$$\delta_n \tau_{Z_1^{(n)}} = (\xi_1 + \cdots + \xi_{Z_1^{(n)}})\delta_n.$$

The non-degenerate distribution function $H(x)$ is a limit distribution for $\delta_n \tau_{Z_1^{(n)}}$ if and only if there exists a positive constant A such that

$$h(t) = f\left(h\left(\frac{t}{A} \right) \right),$$

where $h(t)$ is the characteristic function of $H(x)$ and $f(z)$ denotes the generating function of the variables $v_i^{(n)}$ $(i, n = 1, 2, \ldots)$.

PROOF. Let $F(x)$ and $\psi(t)$ denote the distribution and the characteristic function, respectively, of the random variables ξ_i. Then the characteristic function $\psi_n(t)$ of the random variables $\delta_n \tau_{Z_1^{(n)}}$ is

$$\psi_n(t) = f_n(\psi(\delta_n t)).$$

If $H(x)$ is the limit distribution of the sequence $\delta_n \tau_{Z_1^{(n)}}$, then

$$\lim_{n \to +\infty} f_n(\psi(\delta_n t)) = h(t).$$

We have

$$\lim_{n \to +\infty} f_{n+1}(\psi(\delta_n t)) = f(h(t)).$$

$f(h(t))$ is continuous. Since $f_{n+1}(\psi(\delta_n t))$ is the characteristic function of $\delta_n \tau_{Z_1^{(n+1)}}$, we see that $f(h(t))$ is also a characteristic function. It is easy to see that $f(h(t))$ is at the same time non-degenerate. It follows by a lemma of KHINCHIN (see [2], page 308) that

$$\lim_{n \to +\infty} \frac{\delta_n}{\delta_{n+1}} = A$$

exists, it is positive and the equality

$$H\left(\frac{x}{A}\right) = \sum_{k=1}^{\infty} p_k H^{(k*)}(x)$$

holds, where $H^{(k*)}(x)$ is the k-fold convolution of $H(x)$ and

$$f(z) = \sum_{k=1}^{\infty} p_k z^k.$$

Thus we have in the language of the characteristic functions

$$h(At) = f(h(t)).$$

Let us suppose, conversely, that the last functional equation is satisfied. Then

$$h(t) = f_n\left(h\left(\frac{t}{A^n}\right)\right).$$

Consider now the sequence $\xi_1, \xi_2, \ldots$ of independent and identically distributed random variables with the common distribution function $H(x)$ and put

$$\delta_n = \frac{1}{A^n},$$

further

$$\tau_{Z_1^{(n)}} = \xi_1 + \cdots + \xi_{Z_1^{(n)}}.$$

Then the characteristic function of $\delta_n \tau_{Z_1^{(n)}}$ is $f_n\left(h\left(\frac{t}{A^n}\right)\right) = h(t)$, for every $n = 1, 2, \ldots$.

This proves completely our assertion.

In the special case, when the characteristic function of the random variables ξ_i is non-negative we formulate the following

LEMMA 3. *Let us consider the random variables*

$$\delta_n \tau_{Z_1^{(n)}} = \delta_n(\xi_1 + \cdots + \xi_{Z_1^{(n)}}),$$

where the characteristic function of the variables ξ_i is non-negative. In order that the characteristic function $\varphi(t)$ be the characteristic function of the limiting distribution it is necessary and sufficien that his form be

$$\varphi(t) = g(C|t|^{\alpha}) \quad (C > 0), \ (0 < \alpha \leqq 2).$$

where $g(t)$ is the Laplace transform of the distribution function $G(x)$, defined in Lemma 1, or

$$\varphi(t) \equiv 1.$$

PROOF. Since now $\psi(t)$ is non-negative, we can show in the same manner as in the proof of Theorem 4 that the characteristic function $\varphi(t)$ of the limiting distribution is of the form

$$\varphi(t) = g(C|t|^{\alpha}),$$

or

$$\varphi(t) \equiv 1,$$

where we have $C>0$. Now we prove that $0<\alpha\leqq 2$. The inequality $\alpha<0$ is not possible, since $\varphi(0)=1$ and $\lim_{|t|\to+\infty} g(t)=0$. Further $g''(0)$ is finite. From this it follows that $\alpha>2$ is also impossible. In fact, we have

$$\frac{dg(C|t|^\alpha)}{dt} = g'(C|t|^\alpha)C\alpha(\operatorname{sign} t)|t|^{\alpha-1}$$

and

$$\frac{d^2g(C|t|^\alpha)}{dt^2} = g''(C|t|^\alpha)(C\alpha(\operatorname{sign} t)|t|^{\alpha-1})^2 + g'(C|t|^\alpha)C\alpha(\alpha-1)(\operatorname{sign} t)^2|t|^{\alpha-2}$$

and this, in case $\alpha>2$, would imply

$$\left.\frac{d^2\varphi(t)}{dt^2}\right|_{t=0} = 0.$$

But this means that the distribution corresponding to the characteristic function $\varphi(t)$ is degenerate. This is not our case. The equality $\alpha=0$ cannot occur, too. In that case we would have

$$\varphi(t) = g(C).$$

But C is positive and the Laplace transform $g(t)$ is strictly monotonically decreasing. This proves the necessity.

Now the proof of the sufficiency part is the same as that of Lemma 2.

We also give a sufficient condition expressed in terms of the characteristic function of the random variables ξ_i. This ensures the existence of the limit distribution.

THEOREM 15. *Let us suppose that the distribution function of the random variables ξ_i belongs to the domain of attraction of the stable distribution with the characteristic exponent α $(0<\alpha\leqq 2)$, i.e. the logarithm of their characteristic function for $t\to 0$ is of the form*

$$\ln\psi(t) = -C|t|^\alpha L\left(\frac{1}{|t|}\right)\left(1+i\beta\frac{|t|}{t}\omega(t,\alpha)\right),$$

where $C>0$, $|\beta|\leqq 1$ are constants and

$$\omega(t,\alpha) = \begin{cases} \operatorname{tg}\dfrac{\pi}{2}\alpha, & \text{if } \alpha\neq 1, \\ \dfrac{2}{\pi}\ln|t|, & \text{if } \alpha = 1, \end{cases}$$

further $L(x)$ is a slowly oscillating function at $x=+\infty$. Then, for a suitable choice of the normalizing factors the limit distribution of the random variables

$$\delta_n\tau_{Z_1^{(n)}}$$

exists. The characteristic function of the limit distribution is

$$g\left(C|t|^\alpha\left(1+i\beta\frac{|t|}{t}\omega(t,\alpha)\right)\right),$$

where $g(s)$ is the Laplace transform of the distribution function, defined in Lemma 1.

PROOF. The supposition concerning $\ln \psi(t)$ implies that the function

$$\lambda(t) = |t|^{\alpha} L\left(\frac{1}{|t|}\right)$$

converges to 0, as $t \to 0$. (Cf. [3].) Let us choose the factors δ_n according to the definition

$$\delta_n = \inf\left\{t : t > 0,\ \lambda(t) = \frac{1}{M^n}\right\}.$$

For large n the such defined δ_n exists, because in the neighborhood of $t=0$ the function $\lambda(t)$ is continuous together with $\psi(t)$. From this it follows also that $\lambda(\delta_n) = \frac{1}{M^n}$. Let us consider first the case $\alpha \neq 1$. Then

$$\psi(\delta_n t) - 1 = -C\delta_n^{\alpha}|t|^{\alpha} L\left(\frac{1}{\delta_n |t|}\right)\left(1 + i\beta \frac{|t|}{t} \operatorname{tg} \frac{\pi}{2}\alpha\right).$$

We can put $L\left(\frac{1}{\delta_n}\right)$ instead of $L\left(\frac{1}{\delta_n |t|}\right)$, because $L(x)$, for, $x \to +\infty$, is a slowly oscillating function. Thus we have

$$\psi(\delta_n t) - 1 = -C\frac{|t|^{\alpha}}{M^n}\left(1 + i\beta \frac{|t|}{t} \operatorname{tg} \frac{\pi}{2}\alpha\right)(1 + o(1)), \quad n \to +\infty.$$

Further, if $\alpha = 1$, then

$$\psi(\delta_n t) - 1 = -C\delta_n |t| L\left(\frac{1}{\delta_n |t|}\right)\left(1 + i\beta \frac{|t|}{t}\frac{2}{\pi} \ln |t| + i\beta \frac{|t|}{t}\frac{2}{\pi} \ln \delta_n\right).$$

Now $\delta_n L\left(\frac{1}{\delta_n |t|}\right) \ln \delta_n$ converges to zero, if $n \to +\infty$. This follows from the property of the slowly oscillating functions, according to which for every fixed $\varepsilon > 0$ we have

$$\frac{L(x)}{x^{\varepsilon}} \to 0,$$

if $x \to +\infty$. Thus for $\alpha = 1$ we have also

$$\psi(\delta_n t) - 1 = -\frac{C}{M^n}|t|\left(1 + i\beta \frac{|t|}{t}\frac{2}{\pi} \ln |t|\right)(1 + o(1)), \quad n \to \infty.$$

So, for every $0 < \alpha \leqq 2$ we have

$$\psi(\delta_n t) - 1 = -\frac{C|t|^{\alpha}}{M^n}\left(1 + i\beta \frac{|t|}{t}\omega(t, \alpha)\right)(1 + o(1)), \quad n \to +\infty.$$

From this point on the proof is the same as that of the sufficiency part of Theorem 4, paper II. We obtain that the limit distribution exists and his characteristic function is

$$g\left(C|t|^{\alpha}\left(1 + i\beta \frac{|t|}{t}\omega(t, \alpha)\right)\right).$$

COROLLARY. For $0<\alpha\leqq 2$, $\alpha\neq 1$ the limit distribution is a mixture of the stable distribution function of characteristic exponent α with the weight distribution $G(x)$, defined in Lemma 1, paper I; i.e. the limit distribution has the form

$$\int_0^{+\infty} F_\alpha(x/y^{1/\alpha}) \quad dG(y),$$

where $F_\alpha(x)$ is the stable distribution function of characteristic exponent α ($\alpha\neq 1$).

REMARK. If the random variables ξ_i are non-negative then the characteristic number α can take on only the values lying in the interval $0<\alpha\leqq 1$. In fact, for $\alpha>1$ we would have

$$\frac{\psi(t)-1}{t} = -C\frac{L\left(\frac{1}{|t|}\right)}{\frac{1}{|t|^{\alpha-1}}}\operatorname{sign} t\left(1+i\beta\frac{|t|}{t}\operatorname{tg}\frac{\pi}{2}\alpha\right)(1+o(1)),$$

which converges to zero, if $t\to 0$. This means that $\psi'(0)=0$. From this by an assertion of FELLER [2], page 635, it follows that the variables ξ_i have zero mean-value. This not an interesting case.

Having thus investigated the form of the characteristic function of the random variables ξ_i, we are in the position to prove a rate of convergence theorem, similar to Theorem 3, for the rarefaction of renewal processes. The case of Theorem 3 corresponds to $\alpha=1$, while now we shall investigate the case $0<\alpha\leqq 1$. We have proved in Theorem 15 that if $\psi(t)$ and δ_n are such that

$$\psi(\delta_n t)-1 = -C\frac{|t|^\alpha}{M^n}\left(1+i\beta\frac{|t|}{t}\omega(t,\alpha)\right)(1+o(1)), \quad n\to+\infty,$$

then the limit distribution of the variables

$$\delta_n \tau_{Z_1^{(n)}}$$

exists. Now we shall suppose a little more, namely that

$$\psi(\delta_n t) = 1-\frac{C}{2}\frac{|t|^\alpha}{M^n}\left(1+i\beta\frac{|t|}{t}\omega(t,\alpha)\right)+O\left(\frac{|t|^{2\alpha}}{M^{2n}}\right) \quad (n\to+\infty). \tag{30}$$

By the aid of this condition we can prove the following rate of convergence theorem for the rarefaction model described in section 2, paper I.

THEOREM 16. *Let us suppose that* (30) *holds for* $0<\alpha\leqq 1$ *and let us denote the limiting distribution of*

$$\delta_n \tau_{Z_1^{(n)}}$$

by $G_\alpha(x)$. *If* $G_\alpha(x)$ *has a bounded density function, then*

$$|\mathsf{P}(\delta_n \tau_{Z_1^{(n)}} < x) - G_\alpha(x)| \leqq KM^{-\frac{n}{2\alpha+1}},$$

where $K>0$ *is a constant, not depending on* n.

PROOF. Let us consider again the Laplace transform $g(s)$ of the distribution function $G(x)$ defined in Lemma 1, paper I. Its expansion about $s=0$ gives

$$g(s) = 1-s+\frac{\sigma^2+M^2-M}{M^2-M}s^2+o(s^2) \quad (s \to 0).$$

The characteristic function of $G_\alpha(x)$ is

$$g\left(\frac{C}{2}|t|^\alpha\left(1+i\beta\frac{|t|}{t}\omega(t,\alpha)\right)\right).$$

This satisfies the functional equation

$$g\left(\frac{C}{2}|t|^\alpha\left(1+i\beta\frac{|t|}{t}\omega(t,\alpha)\right)\right) = f_n\left(g\left(\frac{C}{2}\frac{|t|^\alpha}{M^n}\left(1+i\beta\frac{|t|}{t}\omega(t,\alpha)\right)\right)\right)$$

for every $n=1, 2, \ldots$. Further the characteristic function of

$$\delta_n \tau_{Z_1^{(n)}}$$

is

$$f_n(\psi(\delta_n t)).$$

So, using the Lagrange theorem and (30), we have

$$\left|f_n(\psi(\delta_n t))-g\left(\frac{C}{2}|t|^\alpha\left(1+i\beta\frac{|t|}{t}\omega(t,\alpha)\right)\right)\right| =$$

$$= \left|f_n'(\vartheta_n(t))\left(\psi(\delta_n t)-g\left(\frac{C}{2}\frac{|t|^\alpha}{M^n}\left(1+i\beta\frac{|t|}{t}\omega(t,\alpha)\right)\right)\right)\right| \leqq$$

$$\leqq M^n\left|\psi(\delta_n t)-g\left(\frac{C}{2}\frac{|t|^\alpha}{M^n}\left(1+i\beta\frac{|t|}{t}\omega(t,\alpha)\right)\right)\right| \leqq L\frac{|t|^{2\alpha}}{M^n},$$

where $0 \leqq |\vartheta_n(t)| \leqq 1$ and L is a positive constant not depending on n. We have thus

$$\left|\frac{f_n(\psi(\delta_n t))-g\left(\frac{C}{2}|t|^\alpha\left(1+i\beta\frac{|t|}{t}\omega(t,\alpha)\right)\right)}{t}\right| \leqq L\frac{|t|^{2\alpha-1}}{M^n},$$

which means that for $T>0$

$$I = \int_{-T}^{T}\left|\frac{f_n(\psi(\delta_n t))-g\left(\frac{C}{2}|t|^\alpha\left(1+i\beta\frac{|t|}{t}\omega(t,\alpha)\right)\right)}{t}\right|dt \leqq L\frac{T^{2\alpha}}{\alpha M^n},$$

since now $2\alpha-1 > -1$.

Let A denote an upper bound of the density function of $G_\alpha(x)$. Then by Esseen's theorem for an arbitrary $k>1$ we have

$$|\mathrm{P}(\delta_n \tau_{Z_1^{(n)}} < x)-G_\alpha(x)| \leqq \frac{k}{2\pi}\frac{LT^{2\alpha}}{\alpha M^n}+\frac{AC(k)}{T},$$

$C(k)$ being a positive constant depending on k. Let

$$T = M^{\frac{n}{2\alpha+1}}$$

and denote by K the number

$$\frac{Lk}{2\pi\alpha} + AC(k).$$

Then from the preceding inequality our assertion follows.

REFERENCES

[1] MOGYORÓDI, J.: A remark on stable sequences of random variables and a limit distribution theorem for a random sum of independent random variables. *Acta Mathematica Acad. Sci. Hung.* **17** (1966) 401—409.

[2] Феллер, В: *Введение в теорию вероятностей и ее приложения*. 2. Издательство „Мир". Москва, 1967.

[3] Ибрагимов, И. А, Линник, Ю. В.: *Независимые и стационарно связанные величины*. Издательство Наука. Москва 1965.

[4] Гнеденко, Б. В., Колмогоров, А. Н.: *Предельные распределения для сумм независимых величин*. Гостехиздат. Москва, 1949.

Eötvös L. University, Budapest

(Received March 11, 1971.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 39—41.

THE RADICAL OF A LOCALLY COMPACT SEMIALGEBRA

by

R. SCOZZAFAVA

1. Introduction

A *semialgebra* is a non-empty subset A of a real Banach algebra B, such that

(i) $x \in A,\ y \in A,\ \alpha \in R^+ \Rightarrow x+y \in A,\ xy \in A,\ \alpha x \in A.$

The semialgebra A is said to be *strict* if

$$x \in A,\ -x \in A \Rightarrow x = 0.$$

BONSALL introduced in [2] the concept of *locally compact semialgebra*, i.e. a semialgebra $A \neq \{0\}$ such that

(ii) the set of elements $x \in A$, with $\|x\| \leqq 1$, is a compact subset of B.

He proved the existence and properties of idempotents and minimal ideals in a locally compact semialgebra. In particular, his main results (which lead to theorems analogue of classical theorems of WEDDERBURN on the structure of algebras [1]) concern *semisimple* locally compact semialgebras, i.e. semialgebras A for which the zero ideal is the only closed two-sided ideal J such that $J^2 = \{0\}$.

In [4] TOMIUK introduced the concept of *radical* for a *left completely continuous semialgebra*, i.e. a strict semialgebra A such that for each element $a \in A$ the multiplication operator $T_a : x \to ax$ $(x \in A)$ is completely continuous. Every *strict* locally compact semialgebra is left completely continuous, and Tomiuk gave also an example of a left completely continuous semialgebra which is not locally compact. Much of the technique developed by BONSALL in [2] for locally compact semialgebras can be carried over to left completely continuous semialgebras.

The radical of A is defined by TOMIUK via the concept of *quasi-regular* element (i.e. an element $a \in A$ such that there exists $b \in A$ with $a+ba = a+ab = b$). Namely, he defines the radical as the set

$$\{a \in A : a(\beta + x) \text{ is quasi-regular, } (x \in A), (\beta \in R^+)\},$$

and the definition he gives of semisimplicity, i.e. radical $= \{0\}$, is different from that given by BONSALL, but the two definitions are equivalent for *strict* locally compact semialgebras.

In this note we give a simple definition of the radical $\mathscr{R}(A)$ for a locally compact semialgebra A, which is analogue of that of Jacobson radical for a ring R as the set of all elements of R which annihilate all the irreducible R-modules (cfr. [3], p. 4). Moreover, we show that $\mathscr{R}(A) = \{0\}$ *if and only if* A is semisimple according to BONSALL's definition.

2. The radical and semisimplicity

If P and Q are subsets of a semialgebra A, PQ will denote the set of all finite sums of elements of the form pq, with $p \in P$ and $q \in Q$; and P^2 will denote PP.

A semialgebra J contained in A is called a *right ideal* of A if $JA \subset J$. Left and two-sided ideals are similarly defined.

An ideal J of A is called a *closed* ideal if it is a closed subset of A in the relative topology induced by the norm topology in B.

A closed right (left, two-sided) ideal J of A is called *minimal* if $J \neq \{0\}$, and the only closed right (left, two-sided) ideals of A contained in J are $\{0\}$ and J.

BONSALL proved in [2] that each non-zero closed right ideal of a locally compact semialgebra contains a *minimal* closed right ideal (and a similar statement holds for left and two-sided ideals).

Let us recall that, given a subset E of a semialgebra A, the *right annihilator* of E in A is the set

$$(E)_r = \{x : x \in A \quad \text{and} \quad ux = 0 \ (u \in E)\}.$$

Given an index set Λ, let us denote by $\{M_\lambda : \lambda \in \Lambda\}$ the set of all minimal closed two-sided ideals of a locally compact semialgebra A, such that $M_\lambda^2 \neq \{0\}$.

DEFINITION 1. *The radical $\mathscr{R}(A)$ of a locally compact semialgebra A is the intersection of the right annihilators of all minimal closed two-sided ideals M_λ of A such that $M_\lambda^2 \neq \{0\}$, i.e.*

$$\mathscr{R}(A) = \bigcap_{\lambda \in \Lambda} (M_\lambda)_r .$$

(We agree that the intersection of an empty collection of subsets of A is A.)

DEFINITION 2. *We call semisimple a locally compact semialgebra A with $\mathscr{R}(A) = \{0\}$.*

We show now that for locally compact semialgebras our definition of semisimplicity is equivalent to BONSALL's definition.

THEOREM. *In a locally compact semialgebra A we have $\mathscr{R}(A) = \{0\}$ if and only if the zero ideal is the only closed two-sided ideal J with $J^2 = \{0\}$.*

PROOF. Suppose $\mathscr{R}(A) = \{0\}$, and let J be a closed two-sided ideal of A, with $J^2 = \{0\}$. If $J \not\subset \mathscr{R}(A)$, then there exists a minimal closed two-sided ideal M_λ, with $M_\lambda^2 \neq \{0\}$, such that $M_\lambda J \neq \{0\}$. By the minimality of M_λ, it follows $M_\lambda \cap J = M_\lambda$, since $M_\lambda J \subset M_\lambda \cap J$. Hence $M_\lambda \subset J$ and so $M_\lambda^2 = \{0\}$, which is a contradiction. Therefore $J \subset \mathscr{R}(A)$ and so $J = \{0\}$.

Conversely, given any closed two-sided ideal J, suppose that $J^2 = \{0\}$ implies $J = \{0\}$. The right annihilators of the minimal closed two-sided ideals M_λ are closed two-sided ideals, and so $\bigcap_{\lambda \in \Lambda} (M_\lambda)_r = \mathscr{R}(A)$ is a closed two-sided ideal. If $\mathscr{R}(A) \neq \{0\}$ it contains a minimal closed two-sided ideal M_α (and $M_\alpha^2 \neq \{0\}$, by hypothesis). By the definition of $\mathscr{R}(A)$ it follows that M_α is a right annihilator of each element of the set $\{M_\lambda : \lambda \in \Lambda\}$. In particular $M_\alpha^2 = \{0\}$, contradicting the hypothesis. Therefore $\mathscr{R}(A) = \{0\}$.

REFERENCES

[1] ALBERT, A. A.: Structure of algebras, *Amer. Math. Soc. Colloq. Publ.*, XXIV (1939).
[2] BONSALL, F. F.: Locally compact semialgebras, *Proc. London Math. Soc.*, **13** (1963), 51—70.
[3] JACOBSON, N.: Structure of rings, *Amer. Math. Soc. Colloq. Publ.*, XXXVII (1956).
[4] TOMIUK, B. J.: Left completely continuous semialgebras, *Fundam. Mathem.*, **64** (1969), 123—145.

Edinburgh, on leave from „Istituto Matematico U. Dini", Universita di Firenze (Italy).

(Received July 20, 1971.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 43—48.

ON THE HAMILTONIAN INDEX OF A GRAPH

by

GARY CHARTRAND and CURTISS E. WALL

Introduction. In [1] the following theorem was proved.

THEOREM A. *If G is a connected graph of order p (i.e., G has p vertices) which is not a path, the iterated line-graph $L^n(G)$ is hamiltonian for all $n \geqq p-3$.*

The *hamiltonian index* $h(G)$ of a graph G is defined as the smallest non-negative integer n such that $L^n(G)$ is hamiltonian. According to Theorem A, $h(G)$ is defined for all connected graphs G which are not paths, and, furthermore, $h(G) \leqq p-3$, where p is the order of G. It was also shown in [1] that the bound $p-3$ is sharp since for each $p \geqq 3$ there exists a graph of order p whose hamiltonian index is $p-3$. In particular, each of the graphs of Figure 1 has $p-3$ as its hamiltonian index.

In this article a formula is presented which gives the hamiltonian index of all trees (connected graphs without cycles) which are not paths. In addition, an improved bound is established for connected graphs containing cycles.

Before proceeding further, it is convenient to present a few definitions and state some results which will be useful in what follows.

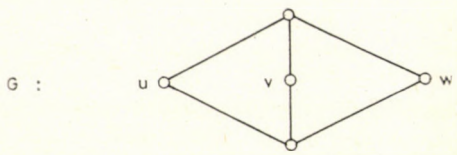


Figure 1. Graphs of order p with hamiltonian index $p-3$

Premilinaries. All graphs considered in this article are ordinary graphs, i.e., graphs without loops or multiple edges.

An *end-vertex* is a vertex having degree one. A *cut-vertex (bridge)* of a connected graph G is a vertex (edge) the removal of which from G results in a disconnected graph. A *block of a graph* G is a maximal connected subgraph which contains no cut-vertices of itself. A block may contain a single edge *(an acyclic block)*, otherwise it contains cycles and is called a *cyclic block*. If a cyclic block has a cycle containing all vertices of that block, then it is referred to as a *hamiltonian block*. An *end-block* of a graph is a block possessing exactly one cut-vertex of the graph. It is well known that every connected graph with two or more blocks has at least two end-blocks.

With every graph G having at least one edge there is associated another graph $L(G)$, called the *line-graph* of G, whose vertex set can be put in one-to-one correspondence with the edge set of G such that two vertices of $L(G)$ if and only if the corresponding edges of G are adjacent. By $L^2(G)$ we shall mean $L(L(G))$ and for $n \geqq 3$, $L^n(G) = L(L^{n-1}(G))$.

The symbols $L^1(G)$ and $L^0(G)$ denote $L(G)$ and G, respectively.

A *hamiltonian graph* is a graph G which has a cycle containing all vertices of G; such a cycle is also called *hamiltonian*. A graph G is called *sequential* if its q edges,

$q \geqq 3$, can be ordered $e_1, e_2, \ldots, e_q, e_{q+1}=e_1$ such that consecutive edges in the sequence are adjacent. The hamiltonian and sequential graphs are related by the following theorem.

THEOREM B. *The line-graph $L(G)$ of a graph G is hamiltonian if and only if G is sequential.*

The hamiltonian graphs form a subset of the set of sequential graphs, so by Theorem B, the line-graph of a hamiltonian graph is hamiltonian. Another class of sequential graphs was determined in [1].

THEOREM C. *If G is a graph containing a cycle C such that every edge of G is incident with at least one vertex of C, then G is sequential.*

The Hamiltonian Index of a Tree. In this section a formula for the hamiltonian index of any tree is developed. For the purpose of doing this we introduce some new concepts and prove two lemmas.

A graph G is *totally sequential* if its q edges, $q \geqq 3$, can be ordered $e_1, e_2, \ldots, e_q, e_{q+1}=e_1$ such that (i) every two consecutive edges in the sequence are adjacent and (ii) for every vertex v of G which is not an end-vertex, there exists i, $1 \leqq i \leqq q$, such that v is incident with e_i and e_{i+1}. Such an ordering is called a *total edge sequence* of G. Certainly, every totally sequential graph is sequential; however, the converse is not true. For example, the graph G of Figure 2 is sequential but not totally sequential since there exists no appropriate ordering of the edges of G having each of the vertices u, v, w incident with consecutive edges in the sequence.

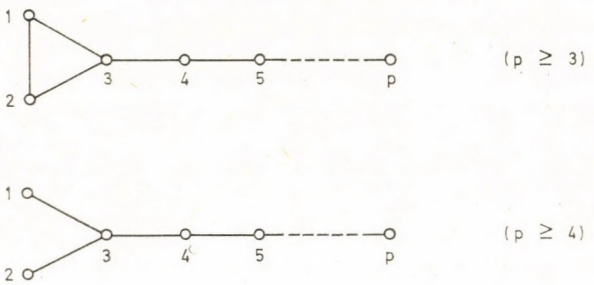


Figure 2. A sequential graph which is not totally sequential

We have already noted that every hamiltonian graph is sequential; we show now that each such graph is, in fact, totally sequential.

LEMMA 1. *Every hamiltonian graph is totally sequential.*

PROOF. Let G be a hamiltonian graph and C: $u_0, u_1, \ldots, u_p, u_0$ be a hamiltonian cycle of G. We now construct an ordering s of the edges of G. Begin s with those edges of G incident with u_0 which do not belong to C, i.e., the diagonals of C incident with u_0. (These edges may be taken in any order.) Next select the edge u_0u_1 of C and then the diagonals of C incident with u_1. After this edge u_1u_2 is placed into s, followed by the diagonals of C incident with u_2 which have not already appeared in s. Continuing in this manner, we complete the sequence s with the edge

$u_p u_0$. It is easily observed that s satisfies the two properties required for it to be a total edge sequence of G.

This fact will prove useful in verifying the lemma which follows.

LEMMA 2. *If G is a connected graph which is not a path such that any acyclic block of G is an endblock while every cyclic block of G is hamiltonian, then $L(G)$ is hamiltonian.*

PROOF. Suppose first that G contains no cyclic blocks. In this case every block of G is acyclic and G is a star graph with three or more edges. Hence $L(G)$ is complete and thus hamiltonian.

We may therefore assume that G contains at least one cyclic block. In this case, we prove that G is totally sequential (implying that $L(G)$ is hamiltonian) by employing induction on the number n of blocks of G. Suppose $n=1$ so that G consists of exactly one block, namely a cyclic block. Since every cyclic block of G is a hamiltonian block, G itself is hamiltonian and therefore totally sequential by Lemma 1.

Assume now that any graph with k blocks, $k \geqq 1$, at least one of which is cyclic and which satisfies the hypothesis of the lemma is totally sequential. Let G be a graph with $k+1$ blocks having at least one cyclic block such that G satisfies the hypothesis of the lemma. We show that G is totally sequential. Let B be an endblock of G, and let v be the cut-vertex of G belonging to v. Remove from G all vertices of B different from v, calling the resulting graph H.

It may occur that B is the only cyclic block of G, in which case H is a star graph containing one or more edges. Since B is a hamiltonian graph, it is totally sequential; therefore, there exists a total edge sequence t of B.

Let e and f be the consecutive edges of t incident with v. By inserting the edges of H between e and f in t, a total edge sequence of G is obtained so that G is totally sequential.

If H contains one or more cyclic blocks, then H satisfies the inductive hypothesis. Hence H is totally sequential and therefore there exists a total edge sequence s of H. Let a and b be two consecutive edges of s incident with v. If B is acyclic, its edge may be inserted between a and b in s obtaining a total edge sequence s_1, of G. Otherwise, B is a hamiltonian block so that there exists a total edge sequence t of B. Let c and d be two consecutive edges of t incident with v. The terms of t may be permuted cyclically (if necessary) producing a total edge sequence t' whose first two terms are c and d (or d and c). By deleting the first term of t' and inserting the remaining sequence between a and b, a total edge sequence of G is obtained. Therefore G is totally sequential and $L(G)$ is hamiltonian.

Before presenting the main result of this section we need a few additional terms. A *poly-vertex* is one having degree at least three. A *poly-path* in a graph G is a path joining two polyvertices any other vertex of which has degree two in G, while and *end-path* in G is a path joining a poly-vertex with an endvertex such that all other vertices of the path have degree two in G. Denote the length of a path P (the number of edges in P) by $l(P)$.

THEOREM 1. *If G is a tree which is not a path, then*

$$h(G) = \max\{\{l(Q)\}, \{1+l(P)\}\},$$

over all end-paths Q and poly-paths P of G.

PROOF. From the manner in which the line-graph is defined, to any end-path or poly-path in G there corresponds in $L(G)$ the same type of path having length one less than that in G. Furthermore, each cyclic block of $L(G)$ is produced from a star subgraph of G; hence each cyclic block of $L(G)$ is complete and therefore hamiltonian. In $L^2(G)$ every end-path and poly-path has length two less than the corresponding paths in G. Also observe that the cyclic blocks of $L^2(G)$ are produced from those maximal subgraphs of $L(G)$ which satisfy the hypothesis of Lemma 2. Thus each cyclic block of $L^2(G)$ is hamiltonian. The same statement can now be made about the graphs $L^n(G)$, $n \geqq 3$. Let

$$k = \max\{\{l(Q)\}, \{1+l(P)\}\}$$

over all end-paths Q and poly-paths P of G.

We show $h(G) \geqq k$ by considering two cases.

Case 1. Suppose there exists an end-path Q of G such that $l(Q)=k$. Then in $L^{k-1}(G)$, there is an end-path of length one (but no end-paths having length exceed-one) so that $L^{k-1}(G)$ is not hamiltonian.

Case 2. Assume there is a poly-path P of G such that $l(P) = k-1$, then in $L^{k-2}(G)$ there is a poly-path of length one (but no poly-path of length exceeding one) implying the existence of a cut-vertex in $L^{k-1}(G)$. Therefore $L^{k-1}(G)$ is not hamiltonian.

Hence $h(G) \geqq k$. We have already observed that in $L^{k-1}(G)$ there are no end-paths having length more than one. Also, $L^{k-1}(G)$ contains no poly-paths. Thus the end-paths of $L^{k-1}(G)$ are precisely the acyclic blocks so that each such block is an end-block. Therefore the graph $L^{k-1}(G)$ satisfies the hypothesis of Lemma 2 and $L^k(G)$ is hamiltonian. This implies $h(G)=k$.

The preceding theorem further implies that the tree of Figure 1 is the only tree of order p having hamiltonian index $p-3$.

On the Hamiltonian Index of Graphs with Cycles. As a consequence of Theorem 1, only the hamiltonian index of graphs with cycles need to be determined. Of course, if a graph is hamiltonian, it has zero as its hamiltonian index. If G is a connected graph with two or more blocks, each of which is hamiltonian, then by Lemma 2, $h(G)=1$. By employing the same technique of proof as in Theorem 1, the following result can be established.

THEOREM 2. *If G is a connected graph containing cyclic and acyclic blocks such that each cyclic block is hamiltonian, then*

$$h(G) = \max\{\{l(Q)\}, \{1+l(P)\}\}$$

over all end-paths Q and poly-paths P of G.

It therefore follows that $h(G)$ can be determined for any graph in which each cyclic block is hamiltonian. In other cases, an upper bound can be provided, however, depending on the lengths of its cycles.

THEOREM 3. *If G is a connected graph of order p having a cycle of length n, then $h(G) \leqq p-n$.*

PROOF. If $p=n$, then G is hamiltonian and $h(G)=0$. Hence we may assume $n<p$.

Let H be the subgraph induced (determined) by the set of edges incident with vertices of C, and let Q be the subgraph induced by the vertices of G not on C (i.e., the edges of Q are those edges of G which join two vertices of Q). Denote by $G_1, G_2, \ldots, G_k$ the components of Q. We write $G = H \oplus Q$ to indicate that the edge set of G is partitioned by H and Q. Let $H_1 = L(H)$, which is a hamiltonian graph by Theorem C, and let J_1 be the subgraph induced by those edges incident with one or more vertices of H_1. Thus

$$L(G) = J_1 \oplus L(G_1) \oplus \cdots \oplus L(G_k),$$

where $L(G_i)$ and $L(G_j)$ are vertex-disjoint for $i \neq j$. Let $H_2 = L(J_1)$ and J_2 be the subgraph induced by those edges incident with one or more vertices of H_2 so that

$$L^2(G) = J_2 \oplus L^2(G_1) \oplus \cdots \oplus L^2(G_k).$$

Continuing in this manner, we have for each positive integer m,

$$L^m(G) = J_m \oplus L^m(G_1) \oplus \cdots \oplus L^m(G_k).$$

We consider two cases.

Case 1. Each of the components $G_1, G_2, \ldots, G_k$ of Q is a path. None of these paths have more than $p-n$ vertices so that for each $i = 1, 2, \ldots, k$, $L^{p-n-1}(G_i)$ consists of a single vertex or does not exist. Hence $L^{p-n-1}(G) = J_{p-n-1}$; therefore, $L^{p-n}(G) = L(J_{p-n-1}) = H_{p-n}$, which is hamiltonian.

Case 2. Assume the first t subgraphs, $1 \leq t \leq k$, of $G_1, G_2, \ldots, G_k$ are not paths. Since every G_i, $1 \leq i \leq t$, has at least 3 vertices, each of $G_{t+1}, G_{t+2}, \ldots, G_k$ has at most $p-n-3$ vertices. Hence

$$L^{p-n-3}(G) = J_{p-n-3} \oplus L^{p-n-3}(G_1) \oplus \cdots \oplus L^{p-n-3}(G_t).$$

Since none of the graphs $G_1, G_2, \ldots, G_t$ is a path and since each of these graphs has at most $p-n$ vertices, all the graphs $L^{p-n-3}(G_1), L^{p-n-3}(G_2), \ldots, L^{p-n-3}(G_t)$ are hamiltonian. Thus, for each $i = 1, 2, \ldots, t$, there is an edge joining H_{p-n-3} and $L^{p-n-3}(G_i)$. This implies that for each $i = 1, 2, \ldots, t$, there is a vertex u_i in H_{p-n-2} adjacent to the two endpoints of an edge in $L^{p-n-2}(G_i)$. Also $u_i \neq u_j$ for $i \neq j$. Denote by e_{i1} and e_{i2} the edges of $L^{p-n-2}(G)$ which join u_i to the distinct endpoints of an edge f of $L^{p-n-2}(G_i)$.

We now show that the graph $L^{p-n-2}(G)$ is sequential, implying that $L^{p-n-1}(G)$ is hamiltonian and that $h(G) < p-n$. For $1 \leq i \leq t$, $L^{p-n-2}(G_i)$ is a hamiltonian subgraph of G and, as such, the edges of $L^{p-n-2}(G_i)$ can be arranged in a sequence s_i such that each pair of consecutive edges in s_i are adjacent as are the first and last edges. Let g_i be the term following f_i in s_i (or the first term of s_i if f_i is the last term). The edge f_i is adjacent to both e_{i1}, and e_{i2} while q_i, being adjacent to f_i, is adjacent to one of e_{i1} and e_{i2}. Therefore, by cyclically permuting the terms of s_i (if necessary) and reversing their order (if necessary), we can obtain a sequence s_i' whose first and last terms are adjacent to e_{i1} and e_{i2}, respectively.

Since H_{p-n-2} is a hamiltonian subgraph of $L^{p-n-2}(G)$ and every edge of J_{p-n-2} is incident with a hamiltonian cycle C of H_{p-n-2}, we may order the edges of H_{p-n-2} according to the procedure given in the proof of Lemma 1. If in this sequence we insert, for each u_i in C, those edges incident with u_i which are in J_{p-n-2} but not in

H_{p-n-2} immediately after the diagonals of C at u_i, we obtain a sequence s of the edges of J_{p-n-2} with the property that consecutive edges of s are adjacent as are the first and last edges. Moreover, without loss of generality, we may construct s so that for each $i=1, 2, \ldots, t$, the edge e_{i2} will immediately follow e_{i1}. Hence if for each $i=1, 2, \ldots, t$, we insert s_i' between e_{i1} and e_{i2} in s, we obtain a sequence s' which has properties essential to verify that $L^{p-n-2}(G)$ is a sequential graph.

Obviously, the bound provided by Theorem 3 is most useful if n is the length of a longest cycle in G (sometimes referred to as the circumference of G). The bound $p-n$ is, in a sense, best possible since for every n and p, $3 \leqq n \leqq p$, there exists a graph G of order p having a cycle of length n such that $h(G) = p-n$. Indeed, let H be a hamiltonian graph of order n, let P be a path of length $p-n$, and let G be that graph obtained by identifying an end-vertex of P with any vertex of H.

There is another bound which can be given in case the minimum degree $\delta(G)$ of a graph G is sufficiently large. A few additional terms and a lemma will be convenient at this point.

A subgraph H of a graph G is called a *spanning subgraph* if H has the same vertex set as G. An *eulerian circuit* of a graph G is a circuit (a closed walk in which no edge is repeated but in which vertices may be repeated) containing all vertices and edges of G. An *eulerian graph* is one containing an eulerian circuit.

LEMMA 3. *If G is a graph containing a spanning eulerian subgraph, then $h(G) \leqq 1$.*

PROOF. Let H be a spanning eulerian subgraph of G, and let C: $u_0, u_1, u_2, \ldots, u_n, u_0$ be an eulerian circuit of H. Every edge of G belongs to C or joins two vertices of C. We show that G is sequential which, by Theorem B, implies that $h(G) \leqq 1$. We begin an ordering s of the edges of G with those edges of G incident with u_0 which do not belong to C. (These may be taken in any order.) Next take the edge u_0u_1 of C and then the edges of G incident with u_1 which do not belong to C. The next edge in s is u_1u_2, followed by all edges of G not in C which are incident with u_2 but which have not previously appeared in s. Continuing in this fashion, we complete the sequence s with the edge u_nu_0.

THEOREM 4. *If G is a connected graph such that $\delta(G) \geqq 3$, then $h(G) \leqq 2$.*

PROOF. Let v be a vertex of G. There are three or more edges of G incident with v. These edges correspond to mutually adjacent vertices in $L(G)$ so that there exists a cycle C_v containing all such vertices in $L(G)$. The union of the cycles C_v, over all vertices v of G, is a spanning eulerian subgraph of $L(G)$. Thus, by Lemma 3, $h(L(G)) \leqq 1$, which implies that $h(G) \leqq 2$.

As with Theorem 3, the bound given in Theorem 4 is best possible in general, for if G is any connected graph with $\delta(G) \geqq 3$ such that G contains a bridge, then $h(G) = 2$.

REFERENCES

[1] CHARTRAND, G.: On hamiltonian line-graphs, *Trans. Amer. Math. Soc.* **134** (1968), 559—566.

Western Michigan University. Michigan State University (now at Old Dominion University).

(Received August 25, 1972)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 49—52.

A MORE GENERAL FARKAS-TYPE THEOREM

by

L. REJTŐ

This paper presents a generalized version of the theorems of paper [3]. The main result (Theorem 1) is the analogue of Farkas theorem for the case of an arbitrary locally convex topological vector space. Finally an application of the results is given.

Let E be a topological vector space. E' denotes the space of continuous linear functionals defined on E.

The next theorem is the analogue of Theorem B of paper [3].

THEOREM 1. *Let E be a locally convex topological vector space over the real numbers, let $P(\neq\emptyset)$ be a weakly compact subset of E, let q denote a fixed element of E. Then the following conditions are equivalent:*

(1) *There exists a finite measure μ on P for which*

$$\int_P f(p)\mu(dp) = f(q) \quad \text{for all} \quad f \text{ in } E';$$

(2) *If $g(p)\geqq 0$ holds for some g in E' and for all p in P, then $g(q)\geqq 0$.*

The proof of this theorem needs the following lemmas:

LEMMA A. *Let P be a nonvoid subset of a locally convex topological vector space E. Let*

$$H = \{f: f\in E', \quad f(p)\geqq 0 \quad \text{for all} \quad p\in P\}$$

$$L = \{y: y\in E, \quad f(y)\geqq 0 \quad \text{for all} \quad f\in H\}.$$

Then L is equal to the closed convex hull of the cone generated by P.

PROOF. L is a closed convex cone: $L = \bigcap_{f\in H} \{y: f(y)\geqq 0\}$. $P\subset L$ implies $L\neq\emptyset$. Let K be the convex hull of the cone generated by P, and $\overline{K}$ be the closure of K. Then $\overline{K}\subset L$ holds as well. In the case $\overline{K}=E$, $L=\overline{K}$ obviously holds.

Let us suppose that $\overline{K}\neq E$, let z_0 be an element of $E\setminus\overline{K}$.

We are going to prove that z_0 belongs also to $E\setminus L$, and this yields $\overline{K}=L_0$. There exists $g\in E'$ for which

$$g(z_0) < \inf_{y\in K} g(y)$$

([1] p. 118). $\overline{K}$ is a cone containing the null element of E, hence $\inf_{y\in K} g(y) = 0$, therefore $0\leqq g(y)$ for all y in $\overline{K}$, so that $g(p)\geqq 0$ for all p in P, and $g\in H$. Since $g(z_0)<0$ and g is in H, $z_0\notin L$, which proves Lemma A.

Remark: The assertion of Lemma A holds also in the weak topology. This means that L is equal to the weak closure of K. This is the consequence of the fact that in a locally convex topological vector space each closed convex subset is weakly closed. ([1] p. 119.)

LEMMA B. *Let F be an arbitrary subset of the locally convex topological vector-space E. If the cone V generated by the closed convex hull of F is closed, then V is the closure of the convex hull of the cone generated by F.*

PROOF. The proof is similar to that of Lemma 2. in [3].

PROOF of Theorem 1. (1) obviously implies (2). If (2) holds then q belongs to L, using the notations of Lemma A. The weak closure of the convex hull of the cone, generated by P, satisfies the conditions of Lemma B. Hence L is equal to the cone generated by the weak closure of the convex hull of P, and there exists an $\alpha>0$ real number such that αq belongs to the weak closure of the convex hull of P. Then, according to Proposition 1.2 of [2] (p. 5) there exists a finite regular measure μ on P such that

$$\int_P f(p)\mu(dp) = f(q) \quad \text{for all } f \text{ in } E'.$$

Thus the proof of Theorem 1. is complete.

Remark: In the case when E is the k-dimensional Euclidean space and P is the set of the vectors

$$(a_{11}, \ldots, a_{k1}); \ldots; (a_{1n}, \ldots, a_{kn})$$

Theorem 1. is reduced to the original Farkas theorem.

THEOREM 2. *Let $\mathcal{X}$ be an arbitrary set and let $\{Y; \mathfrak{S}; \lambda\}$ be a σ-finite measure space. Let us suppose that the functions $F(x; y)=F_x(y)=F_y(x)$ for all fixed x are elements of the space $L^p(Y; \lambda)$ for a fixed $1 \leqq p < +\infty$, and moreover let the set*

$$P = \{F_x(y) : x \in X\}$$

be weakly compact in the space $L^p(Y, \lambda)$ and $\varphi(y)$ be an element of $L^p(Y; \lambda)$. Then the next two conditions are equivalent:

(1) *There exists a finite measure μ on X for which*

$$\int_X F_y(x)\mu(dx) = \varphi(y) \quad \textit{for all } y \textit{ in } Y.$$

(2) *If*

$$\int_Y F_x(y)g(y)\lambda(dy) \geqq 0$$

holds for some $g(y)$ in $L^q(Y; \lambda)$ and for all x in X, where $L^q(Y; \lambda)$ is the conjugate space of $L^p(Y; \lambda)$, then

$$\int_Y \varphi(y)g(y)\lambda(dy) \geqq 0.$$

PROOF. This theorem is a consequence of Theorem 1. similarly as in paper [3] where Theorem A is the special case of Theorem B.

Example: Let $f(x)$ and $g(x)$ be continuous functions on the interval $[A; B]$ $(-\infty < A < B < +\infty)$.

Let

$$C = \min\Big(\min_{x\in[A;B]} f(x); \min_{x\in[A;B]} g(x)\Big),$$

$$D = \max\Big(\max_{x\in[A;B]} f(x); \max_{x\in[A;B]} g(x)\Big).$$

$$T = \{(x; y): x\in[A; B],\ f(x) \leqq y \leqq g(x) \quad \text{or} \quad g(x) \leqq y \leqq f(x)\}.$$

One wishes to prescribe the (one dimensional) measure of the horizontal intersections of T, i.e. the measure μ of the sets

$$A_y = \{x: (x; y)\in T\}, \qquad y\in[C; D].$$

In this case

$$F_y(x) = \begin{cases} 1 & \text{if} \quad x\in A_y, \\ 0 & \text{elsewhere} \end{cases}$$

and

$$P = \{F_x(y): x\in[A; B]\}.$$

The functions $F_x(y)$ are elements of the function space $L^1([C; D]; \lambda)$, where λ denotes the Lebesgue measure.
The set of the functions $F_r(y)$, where r is a rational number in $[A, B]$, is everywhere dense in P. Hence P is separable in the weak topology as well.

Let $F^*(y)$ be the weak limit of a sequence

$$\{F_{r_n}(y)\}_{n=1}^{\infty} \quad \text{where} \quad \{r_n\}_{n=1}^{\infty}$$

is a sequence of rational numbers in $[A, B]$ and $F_{r_n}(y)$ is an element P for all n. It will be proved that $F^*(y)$ belongs to P. Then the uniform boundedness of the functions of P implies that P is compact.

$$F_{r_n}(y) = \begin{cases} 1 & \text{if} \quad y\in[a_n; b_n], \\ 0 & \text{elsewhere.} \end{cases}$$

If the sequence $\{a_n\}_{n=1}^{\infty}$ has no limit point then there exist two subsequences of it for which $a_{n_k} \to a$ and $a'_{n_l} \to c$. Similarly, if the sequence $\{b_n\}$ has no limit point, then there exist $b_{n_k} \to b$ and $b'_{n_l} \to d$ subsequences of $\{b_n\}$. Let

$$h(y) = \begin{cases} 1 & \text{if} \quad y\in[a; c]\cup[b; d], \\ 0 & \text{elsewhere.} \end{cases}$$

Then

$$\int_C^D [F_{r_{n_k}}(y) - F_{r_{n'_l}}(y)]h(y)\,dy \to |a-c| + |b-d|,$$

$$\text{if} \quad n_k \to \infty \quad \text{and} \quad n'_l \to \infty.$$

The functions $F_{r_n}(y)$ converging weakly to $F(y)$, there exists an $n_0(\varepsilon)$ for all positive ε, that

$$\left| \int_C^D [F_{r_n}(y) - F_{r_m}(y)] h(y)\, dy \right| < \varepsilon$$

for all $n, m > n_0$. Hence $|a-c|+|b-d| = 0$ and there exist a^* and b^* for which

$$F^*(y) = \begin{cases} 1 & \text{if} \quad y \in [a^*; b^*], \\ 0 & \text{elsewhere.} \end{cases}$$

$F^*(y)$ belongs to P, because $f(x)$ and $g(x)$ are continuous.
In this case Theorem 2 is applicable.

REFERENCES

[1] EDWARDS, R. E.: *Functional Analysis*, Holt, Rinehart and Winston 1965.
[2] PHELPS, R. R.: *Lectures on Choquet's theorem*, Van Nostrand, Princeton, 1966.
[3] REJTŐ, L.: On a generalization of Farkas theorem. *Studia Sci. Math. Hung.* **6** (1971), 361—365.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received October 18, 1971.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 53—57.

ON THE FACE VECTOR OF A 4-VALENT 3-POLYTOPE

by

E. JUCOVIČ

By the *face vector* of a 3-polytope P we mean the vector $(p_3(P), p_4(P), \ldots, p_m(P))$, where $p_i(P)$ denotes the number of i-gonal faces of P; $p_j(P)=0$ for all $j>m$. For the face vector of a 3-polytope P with a regular 3-valent or 4-valent graph from Euler's formula follows

$$\sum_{k=3}^{m}(6-k)p_k(P) = 12, \text{ or}$$

$$\sum_{k=3}^{m}(4-k)p_k(P) = 8, \text{ respectively.}$$

These equalities do not impose restrictions on the numbers $p_6(P)$ or $p_4(P)$, respectively.

However, the converse of the equalities above is not true in general. Not all sequences $p=(p_3, p_4, \ldots, p_m)$ satisfying the condition

$$\sum_{k=3}^{m}(6-k)p_k = 12, \tag{1}$$

or

$$\sum_{k=3}^{m}(4-k)p_k = 8 \tag{2}$$

are 3-realizable or 4-realizable, respectively. (A sequence p of non-negative integers is called *j-realizable*, $j=3, 4$, if it is the face vector of a 3-polytope with a regular j-valent graph. If none such 3-polytope exists, p is called non-j-realizable.)

Moreover, there are sequences $p=(p_3, p_4, \ldots, p_m)$ satisfying condition (1) which are non-3-realizable with an infinite number of values of p_6. E.g. the sequence $(4, 0, 0, p_6)$ is 3-realizable only if p_6 is even. (B. GRÜNBAUM [2].) In this note we show that this phenomenon does not occur when 4-realizability of sequences is considered. Given a sequence $p'=(p_3, p_5, \ldots, p_m)$ satisfying condition (2) there is a *finite* number of values p_4 only for which $p=(p_3, p_4, p_5, \ldots, p_m)$ is non-4-realizable.

In constructive parts of the proofs below STEINITZ's [6] characterization of 3-polytopal graphs is used: A graph is 3-polytopal if and only if it is planar and 3-connected. — So graphs with the properties mentioned, instead of polytopes, are constructed.

THEOREM 1. *Every sequence* $(p_3, p_4, \ldots, p_m)$ *satisfying the conditions*

$$\sum_{k=3}^{m}(4-k)p_k = 8, \quad \textit{and}$$

$$p_4 \geqq 13 + \sum_{k=5}^{m}(3k-10)p_k$$

is 4-realizable.

The PROOF consists of constructing, for every sequence satisfying the given conditions, its realization M. If $\sum_{k\geq 5} p_k = 0$, the construction of M is indicated in GRÜNBAUM [2, p. 282]. In the other cases M will be the union of three submaps; for simplicity we shall speak of them as of the lid, the belt, and the base.

At the beginning of the construction *the base* B is a $\left(\sum_{k=5}^{m} p_k + 4\right)$-gon or a $\left(\sum_{k=5}^{m} p_k + 3\right)$-gon dissected by its diagonals into 4 triangles and $\sum_{k=5}^{m} p_k$ or $\sum_{k=5}^{m} p_k - 1$ quadrangles according as whether $\sum_{k=5}^{m} p_k$ is even or odd, respectively (see Fig. 1.)

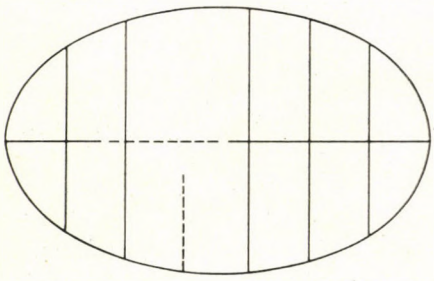


Fig. 1

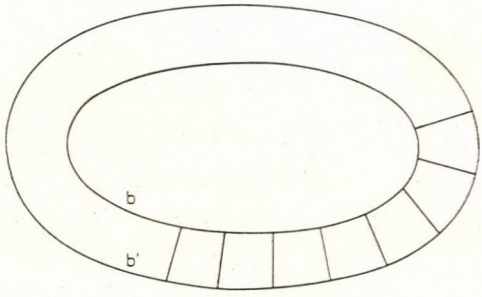


Fig. 2

Notice that only the vertices on the boundary of B are trivalent; all other vertices are 4-valent.

The starting graph for the *belt* T is an annulus decomposed into the same number of quadrangles as the number of vertices on the boundary of the base B (Fig. 2).

Next let us transform the quadrangles (and possibly one triangle) of the base into the required k-gons, $k \geqq 5$. Suppose $\sum_{k=5}^{m} p_k \equiv 0 \pmod 2$. First join the base B to the belt T so that the vertices and edges on the boundary of B are identified with those on the boundary b of T. Let Q be a quadrangle belonging to B and Q' its adjacent quadrangle belonging to T, and suppose that for some k be $p_k \neq 0$. Insert into the quadrangle Q' $(k-4)$ distinct triangles having each one vertex on the common edge of Q and Q', and having the remaining two vertices on the opposite edge. The polygon Q becomes a k-gon, in Q' we get $k-3$ quadrangles. (See Fig. 3 for

$k=6$.) Analogously all required k-gons, $k \geqq 5$, are constructed from the quadrangles belonging to B. If $\sum_{k=5}^{m} p_k \equiv 1 \pmod 2$ the same construction is applied to one triangle of B and a quadrangle of T which is adjacent to the triangle.

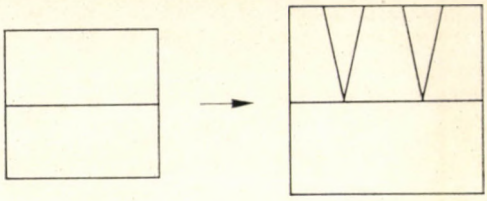


Fig. 3

Up to now all vertices of the map constructed are 4-valent, — except the $2n\left(=\sum_{k=5}^{m}(2k-7)p_k+4 \text{ or } =\sum_{k=5}^{m}(2k-7)p_k+5\right)$ vertices on the second boundary b' of the belt T. The next step consists in adding the *lid*; this is accomplished by dissecting this $2n$-gon by its diagonals, analogously as in the base, into four triangles and $2n-4$ quadrangles. The map M thus obtained realizes the sequence p with $\leqq 5+\sum_{k=5}^{m}(3k-10)p_k$ quadrangles.

Increasing the number of quadrangles by four is performed by inserting, before adding the lid, one triangle into such a quadrangle of the belt which is adjacent to a triangle of the base (Fig. 4).

Increasing the number of quadrangles by five is performed by replacing two neighbouring triangles in the belt and the quadrangle separating them by two triangles and four quadrangles as depicted in Fig. 5. If there is no such pair of triangles in

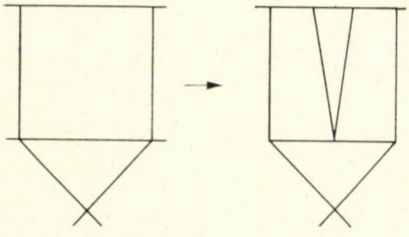


Fig. 4

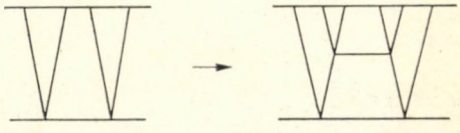


Fig. 5

the map, i.e. if an even number of pentagons is required and $p_i=0$ for all $i>5$, the above construction is changed in such a way that two pentagons are created from two adjacent triangles in the base, and not from quadrangles. Again all this is done before adding the lid.

Increasing the number of quadrangles by six is performed in such a way that in the original base we construct two quadrangles more than we need k-gons, $k \geqq 5$. This increases the number of quadrangles by two in the belt as well as in the lid.

By a suitable combining of these three procedures we get a 4-realization of the sequence p with any $p_4 \geqq 13+\sum_{k\geqq 5}^{m}(3k-10)p_k$. The proof of Theorem 1 is finished.

The bound in Theorem 1 can probably be decreased for all sequences satisfying (2). However, the question about characterizing face vectors of 4-valent 3-polytopes is far from being answered. Sequences of non-negative integers satisfying (2)

which are non-4-realizable do exist. MALKEVITCH [4] has proved: The sequence $(p_3, p_4, \ldots)$ is non-4-realizable if for it $\sum p_i = 1$ for $i \not\equiv 0 \pmod 3$ holds. — From a result of GALLAI [1] (cf. MOTZKIN [5]) it follows: The sequence p is non-4-realizable if for it $\sum p_i = 0$ for $i \not\equiv 0 \pmod 3$ and $\sum_{k \geqq 3} k p_k \not\equiv 0 \pmod{24}$ holds. — A simple reasoning yields: No sequence p with $\sum_{i \geqq 4} p_i = 1$ is 4-realizable.

On the other hand the following Theorem 2 defines quite a large class of sequences which are 4-realizable with *all* values p_4.

THEOREM 2. *Every sequence of non-negative integers* $(p_3, p_4, \ldots, p_m)$ *satisfying the conditions*

$$\sum_{k \geqq 3} (4-k) p_k = 8,$$

$$p_5 = 12 + \sum_{k \geqq 7} (k-6) p_k,$$

$$p_6 \geqq 16$$

is 4-realizable.

In the proof the following statement by GRÜNBAUM [3] will be employed:

LEMMA. *Every sequence* $q = (q_3, q_4, \ldots, q_m)$ *satisfying the conditions* $q_3 = q_4 = 0$, $q_6 \geqq 8$, *and* $\sum_{k \geqq 3} (6-k) q_k = 12$ *is 3-realizable.*

Further the notion of the *medial map* $\mu(M)$ of a planar map M is needed (cf. ORE [7]). If $P_1, P_2, \ldots, P_n$ and $f_1, f_2, \ldots, f_m$ are all vertices and faces of the map M, respectively, all $\mu(P_i)$ and all $\mu(f_i)$ are faces of $\mu(M)$: two faces α and β in $\mu(M)$ are adjacent if and only if $\alpha = \mu(P_i)$, $\beta = \mu(f_j)$ where P_i and f_j are an incident pair face f_j-vertex P_i. — If the graph of M is 3-connected, then so is the graph of $\mu(M)$, which is moreover regular 4-valent. — We are now ready to prove Theorem 2.

Let $c > 0$ be the minimum integer such that $p_4 \equiv c \pmod 4$ holds, and let P be the 3-realization according to Lemma of the sequence $(p_5, p_6^1, p_7, \ldots, p_m)$ with $p_6^1 = p_6 - 2c$. Since in the medial map $\mu(P)$ of P there are at least four face-aggregates consisting of two pentagons and two triangles each (s. Fig. 6a) we can replace each of these c face-aggregates by that depicted in Fig. 6b. (It contains two hexagons, one quadrangle, two pentagons, and six triangles.) The quadrangle is surrounded by four triangles. To obtain the $p_4 - c$ quadrangles insert into one such quadrangle

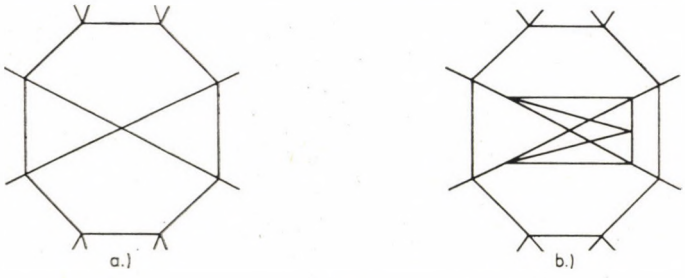


Fig. 6

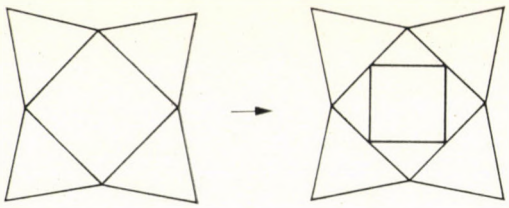


Fig. 7

a new quadrangle and four triangles as depicted in Fig. 7 changing the triangles into quadrangles. The procedure just described is performed $\frac{1}{4}(p_4-c)$ times. For the map N obtained $\sum_{k\geq 3}(4-k)p_k(N) = 8$ holds, and this is why $p_3(N)=p_3$. This completes the proof of Theorem 2.

REFERENCES

[1] GALLAI, T.: Signierte Zellenzerlegungen I. *Acta Math. Hungar.* **22** (1971), 51—63.
[2] GRÜNBAUM, B.: *Convex Polytopes.* Interscience, New York 1967.
[3] GRÜNBAUM, B.: Some analogues of Eberhard's theorem on convex polytopes. *Israel J. Math.* **6** (1968), 398—411.
[4] MALKEVITCH, J.: *Properties of planar graphs with uniform vertex and face structure.* Mem. Amer. Math. Soc. No 99 (1970).
[5] MOTZKIN, T. S.: The edge number of regular-homomorphic 3-polyhedra. *Proc. Colloq. on Convexity, Copenhagen 1965* (Kobenhaven 1967), 212—213.
[6] STEINITZ, E. und RADEMACHER, H.: *Vorlesungen über die Theorie der Polyeder.* Springer, Berlin 1934.
[7] ORE, O.: *The Four Color Problem.* Academic Press, New York 1967.

University of P. J. Šafárik, Košice, Czechoslovakia

(Received November 9, 1971; revised January 25, 1973.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 59—78.

ГЕОДЕЗИЧЕСКИЕ ПОТОКИ НА КОМПАКТНЫХ РИМАНОВЫХ ПОВЕРХНОСТЯХ БЕЗ ФОКАЛЬНЫХ ТОЧЕК

А. КРАМЛИ

Введение и формулировка результатов

Пусть M компактная, двумерная, ориентируемая поверхность класса C^∞, гиперболического типа. Это значит, что на M можно ввести риманову метрику, конформно эквивалентную метрике постоянной отрицательной кривизны. Универсальная накрывающая $\tilde{M}$ может быть реализована, как внутренность единичного круга $u^2+v^2<1$ с метрикой

$$ds^2 = \frac{4f^2(u,v)(du^2+dv^2)}{(1-u^2-v^2)^2} \quad (0<a\leqq f\leqq b).$$

Пространство линейных элементов многообразия $(\tilde{M})$ обозначим через $(\tilde{W})$, а естественную проекцию на $(\tilde{M})$ через $\pi(w)$.

Пусть $\{T_t\}$ геодезический поток на M. $\{T_t\}$ является однопараметрической группой диффеоморфизмов пространства W. Для любого $w\in W$ траектория $T_t w$ является кривой на W, а ее носитель $\pi(T_t w)$ — геодезической на M. Обозначим эту геодезическую, как кривую на M, через $\gamma_w(t)$. Параметр t есть параметр длины дуги, а $\gamma_w(0)=\pi(w)$.

Пусть ν мера, индуцированная римановой метрикой, и σ_x равномерная мера на окружности линейных элементов, имеющих общий носитель $x\in M$. Рассмотрим в W меру μ такую, что индуцированная ею при отображении π мера на M совпадает с ν, а условная мера на $\pi^{-1}(x)$ совпадает с σ_x. Хорошо известно, что μ инвариантная мера для геодезического потока $\{T_t\}$. В случае метрики постоянной отрицательной кривизны в работе [5] Э. Хопфа было доказано, что геодезический поток с мерой μ эргодичен. Из результатов работ Д. В. Аносова [1], Я. Г. Синая [3] и Д. В. Аносова и Я. Г. Синая [2] следует, что в этом случае геодезический поток $\{T_t\}$ будет K-потоком. В работе Э. Хопфа [7] рассматривались поверхности с кривизной переменного знака, но такие, что всякий отрезок времени, при котором геодезическая находится в области положительной кривизны достаточно мал, и сменяется достаточно большим отвезком времини, когда геодезическая находится в области отрицательной кривизны. Методы упомянутых работ позволяют показать, что и в этом случае геодезичический поток буцет K-потоком.

В данной работе рассматриваются такие метрики на M, для которых отсутствуют фокальные точки. Для этого случая мы доказываем следующую теорему:

Теорема: *Пусть W_--множество таких линейных элементов w, что определенные ими геодезические $\gamma_w(t)$ не проводят все время в области нулевой кривизны. Тогда W_- — инвариантное множество относительно геодезического потока, и $\mu(W_-)>0$. Почти каждая эргодическая компонента потока $\{T_t\}$ на*

множестве W_- — имеет положительную меру. На каждой эргодической компоненте, принадлежащей множеству W_-, поток $\{T_t\}$ является K-потоком.

Не исключено, что всегда $\mu(W \setminus W_-)=0$. Мы доказываем это утверждение при некоторых дополнительных условиях.

Метод исследования основан на построении и изучении свойств трансверсальных слоений, введенных в работах [1], [2] и [3]. Оказывается, что из условия отсутствия фокальных точек вытекает для траекторий, вдоль которых временное среднее от гауссовой кривизны отрицательно, выполнение условия типа условия Y неустойчивости Аносова. Неявно это свойство было отмечено в работах [6], [9] Л. Грина, и М. Морса и Г. Хедлунда, посвященных исследованиям топологических свойств геодезических на многообразиях с метрикой без сопряженных точек. (См. также недавную работу [8] В. Клингенберга.)

Особенность рассматриваемой ситуации заключается в том, что условие Y выполняется здесь неравномерно. Поэтому трансверсальные слоения, которые мы строим, существуют почти всюду не обладают хорошими свойствами регулярности. Тем не менее, удается доказать, что специальное отображение, являющееся аналогом канонического изоморфизма для трансверсальных слоев в теории Y-систем (см. [1], [2]), абсолютно непрерывно. Отсюда, применяя методы работ [3] и [4], легко получить утверждение теоремы.

§ 1. Предварительные сведения

1° Определение фокальных точек

Пусть M-двумерная C^∞ риманова поверхность с метрикой ds^2 (см. Введение). Возьмем какой-либо линейный элемент $w \in W$ и рассмотрим геодезическую линию $\gamma_w(t)$ $(\pi(w)=\gamma_w(0))$. Вдоль этой геодезической гауссова кривизна является функцией параметра t, которую мы обозначим через $K(t)$. Рассмотрим вдоль $\gamma_w(t)$ уравнение Якоби

$$s''(t)+K(t)s(t) = 0. \tag{1}$$

Мы скажем, что $\gamma_w(t_0)$ есть фокальная точка для w, если существует решение $s(t)(s(t)\not\equiv 0)$ уравнения Якоби с граничными условиями $s(0)=0$ и $s'(t_0)=0$.

Оснавное предположение о метрике ds^2: Для всякого $w \in W$ не существует ни одной фокальной точки.

2° Координаты Ферми и геометрическое определение фокальных точек

Известно (см. [9]), что на накрывающей поверхности $\tilde{M}$ для любого геодезического отрезка $\gamma_w(t)(0 \leqq t \leqq T)$ можно найти такую полосу ширины ϱ (ϱ универсальная постоянная для M) в которой перпендикулярные к $\gamma_w(t)$ геодезические образуют гладкое семейство кривых, между собой не пересекаются и заполняют всю полосу. Для линейного элемента w_0 носитель $x_0=\pi(w_0)$ которого лежит в этой полосе мы можем ввести следующие координаты s_{w_0} и u_{w_0}. Проведем через x_0 геодезический отрезок $\delta_{x_0}(s)$ ортогональный к $\gamma_w(t)$ и введем на нем параметр длины дуги s, отсчитываемый в положительном направлении от точки пересечений $\gamma_w(t_0)$. Обозначим через s_{w_0} значение параметра s, для которого $\delta_{x_0}(s)=x_0$, и через u_{w_0} тангенс угла между вектором $T_{t_0}w$ и параллельно перенесенным линейным элементом w_0 вдоль $\delta_{x_0}(s)$ до $\gamma_w(t_0)$.

В дальнейшем мы будем называть s_{w_0} и u_{w_0} координатами Ферми линейного элемента w_0 вдоль геодезической $\gamma_w(t)$. Предположим, что геодезическая $\gamma_0(t)=\pi(T_t w_0)$ для $t\in[t_0,T]$ целиком лежит в ϱ-окрестности $\gamma_w(t)$. Тогда решение $s_0(t)$ уравнения Якоби с начальными условиями $s_0(t_0)=s_{w_0}$ и $s_0'(t_0)=u_{w_0}$ определяет с точностью до малого второго порядка функции $s_{T_t w_0}$ и $u_{T_t w_0}$, т. е.

$$
\begin{aligned}
s_{T_t w_0} &= s_0(t)+o(|s_{w_0}|+|u_{w_0}|),\\
u_{T_t w_0} &= s_0'(t)+o(|s_{w_0}|+|u_{w_0}|), \quad t\in[t_0,T].
\end{aligned}
\tag{2}
$$

Естественно, о-малое зависит от t, но для фиксированного значения $|t-t_0|$ оно равномерно на всем многообразии M.

Исходя из формул (2) мы можем дать определения фокальных точек чисто геометрическим путем:

Пусть $\gamma_w(t)$ и $\delta_{x_0}(s)$ те же самые, как в определении координат Ферми, а $\gamma_s(t)$ ортогональная к $\delta_{x_0}(s)$ геодезическая, проходящая через точку $\delta_{x_0}(s)$. Мы скажем, что $\gamma_w(t_1)$ есть фокальная точка для $T_{t_0}w$, если она является пределом точек пересечений геодезических $\gamma_s(t)$ и $\gamma_w(t)$ при $s\to 0$.

Пусть $s(\sigma,t)$ такое однопараметрическое семейство геодезических, что $s(0,t)=\gamma_w(t)$, а $s(t)=\dfrac{\partial}{\partial\sigma}s(\sigma,t)|_{\sigma=0}$ является решением уравнения Якоби вдоль $\gamma_w(t)$. Отношение $\varkappa(t)=s'(t)/s(t)$ имеет следующий геометрический смысл: оно равно геодезической кривизне ортогональной траектории к пучку геодезических $s(\sigma,t)$ при $\sigma=0$. Очевидно, что $\varkappa(t)$ удовлетворяет уравнению Риккати: $\varkappa'(t)+\varkappa^2(t)=K(t)$.

3° Сопряженные точки

Пусть $\gamma_w(t)$ геодезическая линия; точки $\gamma_w(t_0)$ и $\gamma_w(t_1)$ являются сопряженными, если существует решение $s(t)\big(s(t)\not\equiv 0\big)$ уравнения Якоби, с граничными условиями $s(t_0)=0, s(t_1)=0$.

Очевидно, что из отсутствия фокальных точек вытекает отсутствие сопряженных точек. Подобно случаю фокальных точек, мы можем дать геометрическое описание сопряженных точек; возьмем геодезическую $\gamma_w(t)$ и семейство $\gamma_\varphi(t)$ геодезических, проходящих через точку $x=\pi(w)$, где φ параметр равный углу между геодезическими $\gamma_w(t)$ и $\gamma_\varphi(t)$ в точке x. Точка $\gamma_w(t_1)$ сопряженная к точке $x=\gamma_w(0)$, если она является пределом точек пересечений геодезических $\gamma_\varphi(t)$ и $\gamma_w(t)$ при $\varphi\to 0$.

§ 2. Предельное решение уравнения Якоби

В этом параграфе мы построим предельные решения $s^+(t)$ и $s^-(t)$ уравнения Якоби вдоль геодезической $\gamma_w(t)$. Зафиксируем параметр $t: t=t_0$ — и рассмотрим решение $s_{t_1}(t)$ уравнения (1.1) удовлетворяющее граничным условиям $s_{t_1}(t_0)=1$ и $s_{t_1}(t_1)=0$ для $t_1>t_0$. Из линейности уравнения Якоби и отсутствия сопряженных точек вытекает существование такого решения. Из этих условий также вытекает, что для $t_2>t_1$ $s_{t_2}(t)>s_{t_1}(t)$ и $s_{t_2}'(t_0)>s_{t_1}'(t_0)$. Таким образом $s_t'(t_0)$ монотонно не убывает при $t\to\infty$, и благодаря отсутствию фокальных точек $s_{t_1}'(t)<0$ при $t_0\leqq t\leqq t_1$, т. е. существует неположительный предел $s_{t_0}^+=$

$= \lim_{t\to\infty} s_t'(t_0)$. Назовем решение $s^+(t)$, удовлетворяющее начальным условиям $s^+(t_0)=1, \frac{d}{dt}s^+(t)|_{t=t_0} = s_{t_0}^+$ положительно предельным. Так как $s^+(t)$ на каждом конечном отрезке $[t_0, t]$ является равномерным пределом последовательности невозрастающих функций, то она сама невозрастающая и нигде не обращается в 0 при $t > t_0$. Из результатов работы Грина [6] следует, что решение $\varkappa^+(t) = \frac{ds^+(t)}{dt}(s^+(t))^{-1}$ уравнения Риккати ограничено при $t\to\infty$. Там же доказывается единственность такого решения. Из этого вытекает, что определение $s^+(t)$ не зависит от выбора t_0 — разные предельные решения могут отличаться только мультипликативной константой.

Аналогичным образом строится отрицательно предельное решение $s^-(t)$.

Построенное нами решение можно рассматривать как функцию на W. Для того, чтобы это подчеркнуть, обозначим его через $s^+(w)$. Так как функция $s^+(w)$ является пределом непрерывных функций, измерима. Применим к ней эргодическую теорему Биркгофа—Хинчина. Нам будет легче работать с решением $\varkappa^+(t) = \frac{ds^+(t)}{dt}(s^+(t))^{-1}$ уравнения Риккати. Как мы уже доказали, $\varkappa^+(t)$ — неположительная функция. Интегрируя обе части уравнения Риккати по t, от 0 до T, получим:

$$\frac{1}{T}\int_0^T (\varkappa^+(t))'dt + \frac{1}{T}\int_0^T (\varkappa^+(t))^2 dt = \frac{1}{T}\int_0^T K(t)dt.$$

Первый член $-\frac{1}{T}(\varkappa^+(t)|_0^T)$, для почти всех траекторий сходится к 0, благодаря ограниченности $\varkappa_T^+(t)$. Второй член неотрицателен, поэтому предел $\lim_{T\to\infty}\frac{1}{T}\int_0^T K(t)dt$ вдоль почти каждой геодезической неположителен. Множество тех линейных элементов $w\in W$, для которых вдоль $\gamma_w(t)$ $\lim_{T\to\infty}\int_0^T K(t)dt = 0$ образует инвариантное множество относительно геодезического потока $\{T_t\}$.

Далее мы будем заниматься множеством W_- тех линейных элементов для которых этот предел существует и отрицателен. Из предыдущего замечания вытекает, что для почти каждого линейного элемента $w\in W\setminus W_-$, $\gamma_w(t)$ лежит целиком в множестве $\{x\in M: K(x) = 0\}$. Если мы предполагаем, что множество $\{\tilde{x}\in\tilde{M}: K(\tilde{x}) = 0\}$ состоит из непересекающихся ограниченных множеств, нетрудно доказать, что $W\setminus W_-$ имеет меру 0.

Так как $\varkappa^+(t)$ ограничена, положительность предела $\lim_{T\to\infty}\frac{1}{T}\int_0^T (\varkappa^+(t))^2 dt$ влечет за собой отрицательность предела $\lim_{T\to\infty}\int_0^T (\varkappa^+(t))^2 dt$. Отсюдасл едует, что решение уравнения Якоби $s^+(t) = s^+(0)\exp\left\{\int_0^t \varkappa(\tau)d\tau\right\}$ убывает с экспоненциальной скоростью.

Для дальнейших нам понадобится оценка разности двух предельных решений $\varkappa_1^+(t)$ и $\varkappa_2^+(t)$ уравнения Риккати для двух различных геодезических $\gamma_1(t)$ и $\gamma_2(t)$. С этой целью мы докажем следующую лемму:

Лемма 2.1. *Пусть* $\gamma_\sigma(t)$ *— такой гладкий пучок геодезических* ($1 \leqq \sigma \leqq 2$, $0 \leqq t \leqq T$), *у которого* t *есть параметр длины дуги, и кривые* $t=\text{const}$ *образуют ортогональные траектории к пучку* $\gamma_\sigma(t)$. *Возьмем два каких-либо отрицательных решения* $\varkappa_1(t)$ *и* $\varkappa_2(t)$ *уравнения Риккати вдоль* $\gamma_1(t)$ *и* $\gamma_2(t)$ *соответственно. Тогда*

$$|\varkappa_1(t)-\varkappa_2(t)| \leqq \text{const}\,\pi_t^T + |\varkappa_1(T)+\varkappa_2(T)| \exp\left\{\int_t^T (\varkappa_1(\tau)+\varkappa_2(\tau))\,d\tau\right\},$$

где const *абсолютная для многообразия* M *и* π_t^T *обозначает площадь, заключенную между* $\gamma_1(t)$ *и* $\gamma_2(t)$ *в отрезке от* t *до* T.

Доказательство. Сравнивая два уравнения Риккати мы получаем „линейное" уравнение для $\Delta(t) = \varkappa_1(t)-\varkappa_2(t)$:

$$\Delta'(t)+(\varkappa_1(t)+\varkappa_2(t))\Delta(t)+K_1(t)-K_2(t) = 0. \tag{1}$$

($K_1(t)$ и $K_2(t)$ гауссовские кривизны вдоль $\gamma_1(t)$ и $\gamma_2(t)$.) Общее решение уравнения (1) имеет вид:

$$\Delta(t) = \left[\int_t^T (K_1(\tau)-K_2(\tau)) \exp\left\{-\int_\tau^T (\varkappa_1(\vartheta)+\varkappa_2(\vartheta))\,d\vartheta\right\} d\tau + \right.$$
$$\left. + \Delta(T)\right] \exp\left\{\int_t^T (\varkappa_1(\tau)+\varkappa_2(\tau))\,d\tau\right\}, \quad (t \leqq \tau \leqq T). \tag{2}$$

Если $\Delta(T)$ ограничено при $T\to\infty$, тогда для фиксированного значения t, например для $t=0$ в предположении, что хотя бы для одной из двух решений, $\overline{\lim}\,\frac{1}{T}\int_0^T \varkappa(t)\,dt < 0$, при $T\to\infty$ второй член убывает с экспоненциальной скоростью. Для оценки первого слагаемого перепишем формулу (2):

$$\Delta(0) = \int_0^T [K_1(\tau)-K_2(\tau)] \exp\left\{\int_0^\tau (\varkappa_1(\vartheta)+\varkappa_2(\vartheta))\,d\vartheta\right\} d\tau + \tag{2'}$$
$$+ \Delta(T) \exp\left\{\int_0^T (\varkappa_1(\tau)+\varkappa_2(\tau))\,d\tau\right\}.$$

Заметим, что поскольку множитель $\exp\left\{\int_0^\tau (\varkappa_1(\vartheta)+\varkappa_2(\vartheta))\,d\vartheta\right\} \leqq 1$, то первый член оценивается сверху интегралом $\int_0^T |K_1(\tau)-K_2(\tau)|\,d\tau \leqq \text{const}\,\pi_0^T$. Лемма доказана.

Замечание 1.1. Если $\gamma_1(t)$ и $\gamma_2(t)$ экспоненциально сближаются при $t\to\infty$, то π_0^t остается ограниченным. Даже и $\int_0^T \pi_t^\infty dt$ остается ограниченным.

Замечание 1.2. Аналогичные рассуждения справедливы для разности положительных решений $\varkappa_1(t)$ и $\varkappa_2(t)$ при $T\to-\infty$.

§ 3. Оценка расстояния между сближающимися геодезическими

Пусть линейный элемент $w\in W_-$ и для него при некотором $\varepsilon>0$ и всех $T>\frac{1}{\varepsilon}$ интеграл

$$\frac{1}{T}\int_0^T K(t)\,dt<-\varepsilon. \tag{1}$$

Вдоль геодезической $\gamma_w(t)$ введем координаты Ферми s и u. Допустим, что задан геодезический отрезок $\gamma^*(t)$ $(0\leqq t\leqq T)$, который лежит в области определения этих координат Ферми, т. е. расстояние между $\gamma_w(t)$ и $\gamma^*(t)$ всюду меньше, чем ϱ. Предположим, что $\gamma^*(t)$ описывается функциями $s^*(t)$ и $u^*(t)$, где $s^*(t)$ положительна, а $u^*(t)$ неположительна на отрезке $[0, T]$.

Лемма 3.1. *Существуют такие положительные постоянные C_1, C_2 и d, зависящие только от ε, что из неравенства $s^*(0)<d$ вытекает неравенство $\frac{s^*(t)}{s^*(0)}<C_1e^{-C_2t}$ для $0\leqq t\leqq T$.*

Доказательство. Из условия (1) вытекают при $T>\frac{1}{\varepsilon}$ неравенства

$$\begin{aligned}\frac{1}{T}\int_0^T\left(K(t)+\frac{\varepsilon}{2}\right)dt&<-\frac{\varepsilon}{2},\\ \frac{1}{T}\int_0^T\left(K(t)+\frac{\varepsilon}{4}\right)dt&<-\frac{3\varepsilon}{4}.\end{aligned} \tag{2}$$

Обозначим через $M_{\varepsilon/2}$ $(M_{\varepsilon/4})$ множество тех точек $x\in M$ для которых $K(x)\leqq-\frac{\varepsilon}{2}$ $\left(K(x)\leqq-\frac{\varepsilon}{4}\right)$ и обозначим через $\chi_{\varepsilon/2}(x)$ $\left(\chi_{\varepsilon/4}(x)\right)$ отвечающие им характеристические функции. Пусть $\delta=\varrho(\overline{M\setminus M_{\varepsilon/4}}, M_{\varepsilon/2})$. Из компактности M следует, что $\delta>0$. Из неравенств (2) и ограниченности кривизны вытекает, что при $T>\frac{1}{\varepsilon}$

$$\frac{1}{T}\int_0^T\chi_{\varepsilon/2}(\gamma_w(t))\,dt\geqq\tau_{\varepsilon/2}.$$

Обозначим через v_T максимальное число непересекающихся интервалов (t_i, t_{i+1}) длины δ, лежащих внутри отрезка $[0, T]$, на которых $\chi_{\varepsilon/4}(\gamma_w(t))=1$. Очевидно, что

$$\frac{1}{T} v_T > \frac{1}{T} \int_0^T \chi_{\varepsilon/2}(\gamma_w(t))\,dt \geqq \tau_{\varepsilon/2} \quad \text{при} \quad T > \frac{1}{\varepsilon}. \tag{3}$$

Применим теперь формулы (1.2) к нашей геодезической $\gamma^*(t)$ при $t = t_0+\delta$, предполагая, что $\gamma_w(t)$ на отрезке $[t_0, t_0+\delta]$ целиком лежит в множестве $M_{\varepsilon/4}$. Это дает нам:

$$\begin{aligned} s^*(t_0+\delta) &= s_{t_0}(t_0+\delta) + o\big(|s^*(t_0)| + |u^*(t_0)|\big), \\ u^*(t_0+\delta) &= u_{t_0}(t_0+\delta) + o\big(|s^*(t_0)| + |u^*(t_0)|\big), \end{aligned} \tag{4}$$

где $s_{t_0}(t)$ и $u_{t_0}(t)$ получаются из решения $\varkappa_{t_0}(t)$ уравнения Риккати вдоль геодезической $\gamma_w(t)$ с начальным условием $\varkappa_{t_0}(t_0) = \dfrac{u^*(t_0)}{s^*(t_0)}$ при помощи формул

$$s_{t_0}(t) = s_{t_0}(t_0) \exp\Big\{ \int_{t_0}^t \varkappa_{t_0}(\tau)\,d\tau \Big\} \quad \text{и} \quad u_{t_0}(t) = s'_{t_0}(t).$$

Заметим, что из локальной евклидовой структуры многообразия M вытекает ограниченность отношения $\left|\dfrac{u^*(t_0)}{s^*(t_0)}\right|$ при достаточно малом d. Поэтому в формулах (4) $|s^*(t_0)|+|u^*(t_0)|$ можно заменить на $s^*(t_0)$. Теперь мы докажем, что существует такое $\varepsilon_1>0$, что $\int_{t_0}^{t_0+\frac{\delta}{2}} \varkappa_{t_0}(\tau)\,d\tau < -\dfrac{\varepsilon_1\delta}{2}$. Если ε_1 достаточно мало, например меньше, чем $\varepsilon\delta/100$ и $\varkappa_{t_0}(t) > -\varepsilon_1$, для некоторого $t \in \left[t_0,\ t_0+\dfrac{\delta}{2}\right]$, найдется такое положительное число $\delta' < \delta$, что

$$\varkappa_{t_0}(t_0+\delta') > \varepsilon_1, \tag{5}$$

потому, что из уравнения Риккати вытекает неравенство

$$\varkappa'_{t_0}(t) > \frac{\varepsilon}{4} - \varepsilon_1^2 > 0 \quad \text{для} \quad t \in \left[t_0,\ t_0+\frac{\delta}{2}\right].$$

Мы можем предполагать, что $s^*(t_0+\delta') > \dfrac{1}{2} s^*(t_0)$-иначе из соотношения (3) следовало бы утверждение леммы. Но тогда из неравенства (5) и формул (4) для $\delta=\delta'$:

$$u^*(t_0+\delta') \geqq \varepsilon_1 s^*(t_0+\delta') + o\big(s^*(t_0)\big),$$

т. е.

$$u^*(t_0+\delta') \geqq \frac{1}{2}\varepsilon_1 s^*(t_0) + o\big(s^*(t_0)\big),$$

если $s^*(0)$ достаточно мало.

Этим мы доказали неравенство $\int_{t_0}^{t_0+\frac{\delta}{2}} \varkappa_{t_0}(t)dt < -\varepsilon_1 \frac{\delta}{2}$, где ε_1 не зависит от t_0.

А из формул (4) мы получаем:

$$\frac{s^*\left(t_0+\frac{\delta}{2}\right)}{s^*(t_0)} = \frac{s^*(t_0)e^{-\varepsilon_1\frac{\delta}{2}} + o\left(s^*(t_0)\right)}{s^*(t_0)} < e^{-\varepsilon_1\frac{\delta}{4}} < 1 \tag{6}$$

при достаточно малом $s^*(0)/d/$. (6) вместе с (3) уже доказывает лемму. Здесь мы воспользовались тем обстоятельством, что благодаря компактности многообразия M, при фиксированном δ формулы (1.2) выполняются равномерно.

§ 4. Построение орициклов

В этом параграфе мы построим асимптотический пучок геодезических, и покажем, что предельное решение уравнения Риккати является геодезической кривизной ортогональной траектории этого пучка. Эти ортогональные траектории, оснащенные единичными нормальными векторами — как кривые в пространстве W — мы будем называть орициклами, по аналогии с поверхностями с отрицательной кривизны.

Пусть $\gamma_w(t)$ геодезический луч, проходящий через линейный элемент $w \in W$ и $s^+(t)$, $\varkappa^+(t)$ предельные решения уравнений Якоби и Риккати на полупрямой $0 \leqq t \leqq \infty$ отвечающих этой геодезической. Введем в окрестности $\gamma_w(t)$ координаты Ферми.

Теорема 1. *Существует такое однопараметрическое семейство* $\{s(\sigma, t), u(\sigma, t)\}$ *функций, определенных при* $0 \leqq \sigma \leqq \delta$, $0 \leqq t < \infty$ *(где* δ *зависит от* w, *и* $s(\sigma, 0) = \sigma$*), что при фиксированном* $\sigma = \sigma_0$ *кривая заданная в координатах Ферми функциями* $s(\sigma_0, t)$ *и* $u(\sigma_0, t)$, *будет геодезической, а при фиксированном* $t = t_0$ *выполнены соотношения*

$$s^+(t_0) = \left.\frac{\partial s(\sigma, t)}{\partial \sigma}\right|_{\sigma=0} \quad \text{и} \quad \varkappa^+(t_0) = \left(s^+(t)\right)^{-1} \frac{\partial u(\sigma, t_0)}{\partial \sigma}.$$

Доказательство. Возьмем значение $t = T > 0$ и проведем через точку $\gamma_w(T)$ ортогональную кривую $x_T(\sigma')$ на M с постоянной геодезической кривизной $\varkappa$ в окрестности точки $\gamma_w(T)$ фиксированного радиуса. Значение $\varkappa$ мы можем выбрать так, чтобы для любой геодезической решение $\varkappa(t)$ уравнения Риккати с начальным условием $\varkappa(T) = \varkappa$ осталось отрицательным при $t < T$. Приведем далее через $x_T(\sigma')$ пучок ортогональных геодезических

$$x_T(\sigma', t) \quad (0 \leqq t \leqq T).$$

Покажем, что при любом достаточно большим T все рассматриваемые нами пучки лежат в области определения координат Ферми, построенных в окрестности геодезического чула $\gamma_w(t)$. Мы можем считать t параметром длины дуги

для всех геодезических этого пучка. Тогда при фиксированном $t=t_0$, кривые $x_T(\sigma', t_0)$ являются ортогональными траекториями этого пучка. Мы можем в дальнейшем считать, σ' параметром длины дуги на кривой $x_T(\sigma', 0)$. Обозначим через $\Sigma_T(\sigma', t)$ длину дуги кривой $x_T(\sigma', t)$ на отрезке $[0, \sigma']$ и через $\varkappa_T(\sigma', t)$ ее геодезическую кривизну в точке с координатой σ'. Из уравнения Якоби для каждой геодезической рассматриваемого пучка следует соотношение:

$$\Sigma'_T(\sigma', t) = \int_0^{\sigma'} \exp\left\{\int_0^t \varkappa_T(\xi, \tau)\, d\tau\right\} d\xi.$$

Так как $\varkappa_T(\xi, t) < 0$ в силу отсутствия фокальных точек и выбора $\varkappa$, то экспонента под знаком интеграла не превосходит 1. Отсюда следует, что $\Sigma'_T(\sigma', t) \leqq \leqq \delta$, и, если δ достаточно мало, то наш пучок будет лежать в области определения координат Ферми, и применимы леммы 2.1 и 3.1.

Найдется такое ξ_0, $0 \leqq \xi_0 \leqq \sigma'$, зависящее от σ', t, что

$$\Sigma'_T(\sigma', t) = \sigma' \exp\left\{\int_0^t \varkappa_T(0, \tau) d\tau\right\} \left[\exp\left\{\int_0^t (\varkappa_T(\xi_0, \tau) - \varkappa_T(0, \tau)) d\tau\right\}\right] =$$

$$= \sigma' \exp\left\{\int_0^t \varkappa^+(\tau)\, d\tau\right\} \exp\left\{\int_0^t (\varkappa_T(0, \tau) - \varkappa^+(\tau) + \varkappa_T(\xi_0, \tau) - \varkappa_T(0, \tau)) d\tau\right\}.$$

Положим

$$\beta_T(\sigma', t) = \max_{0 \leqq \xi \leqq \sigma'} |\varkappa_T(0, t) - \varkappa^+(t)| + |\varkappa_T(\xi, t) - \varkappa_T(0, t)|.$$

Мы имеем неравенство

$$\exp\left\{-\int_0^t \beta_T(\sigma', \tau) d\tau\right\} \leqq \frac{\Sigma'_T(\sigma', t)}{\sigma' \exp\left\{\int_0^t \varkappa^+(\tau)\, d\tau\right\}} \leqq \exp\left\{\int_0^t \beta_T(\sigma', \tau) d\tau\right\}.$$

Мы докажем теперь, что $\varkappa_T^2(\sigma', t)$ для $t < T$ не превосходит максимума C_0 гауссовой кривизны $K(t)$. Пусть ε произвольное положительное число, и предположим, что для некоторого $t=t_0$ $\varkappa_T^2(\sigma', t_0) = C_0 + \varepsilon$. Из непрерывности решения $\varkappa_T(\sigma', t)$ следует, что при некотором t_1, $t_0 < t_1 < T$, $\varkappa_T^2(\sigma', t_1) = C_0 + \frac{\varepsilon}{2}$. Тогда из уравнения Риккати вытекает неравенство $\varkappa'_T(\sigma', t_1) \leqq \frac{\varepsilon}{2}$, т. е. при убывании t, $\varkappa_T(\sigma', t)$ строго убывает начиная с момента t_1, по крайней мере до того момента t_2, когда $\varkappa_T^2(\sigma', t_2) = C_0$. Последовательным применением этого рассуждения мы получим неравенство $\varkappa_T^2(\sigma', t_0) \leqq C_0 + \frac{\varepsilon}{2}$, которое противоречит предположению. Так как ε произвольно, наше утверждение доказано. Поэтому, если геодезические $x_T(\sigma', t)$ в силу леммы 3.1 экспоненци-

ально сближаются с $\gamma_w(t)$ когда расстояние измеряется по ортогональным геодезическим — то они экспоненциально сближаются и тогда, когда расстояние измеряется по ортогональным траекториям рассматриваемого пучка. Это расстояние равно как раз $\Sigma'_T(\sigma', t)$. По лемме 3.1

$$\sum'_T(\sigma', t) \leqq C_1\sigma' e^{-C_2 t} \quad \text{и} \tag{1}$$

$$\Pi_t^T \leqq C_3\sigma' e^{-C_2 t} \tag{2}$$

где константы C_1, C_2, C_3 не зависят от выбора T. Имея в виду, что $|\varkappa^+(T) - \varkappa_T(0, T)| < \varkappa$, и применяя лемму 2.1 мы можем написать

$$\beta_T(\sigma', t) \leqq C_3 e^{-C_2 t} + \varkappa e^{\int_t^T \varkappa^+(\tau)\, d\tau}. \tag{3}$$

Из оценки (3) получается неравенство

$$\int_0^t \beta_T(\sigma', t) \leqq C_5\sigma', \quad \text{т. е.} \quad \frac{\sum'_T(\sigma', t)}{\sigma'} = e^{\int_0^t \varkappa^+(\tau)\, d\tau} + O(\sigma') + \varepsilon_T. \tag{4}$$

Если мы запишем функции, описывающие пучок $x_T(\sigma', t)$ в координатах Ферми — то для так полученного пучка $s_T(\sigma', t)$ справедливо соотношение аналогичное (4):

$$\frac{\Sigma(\sigma, t)}{\sigma} = e^{\int_0^t K^+(\tau)\, d\tau} + O(\sigma) + \varepsilon_T. \tag{4'}$$

где σ и Σ измеряются по ортогональным к $\gamma_w(t)$ геодезическим, и $\sigma = s_T(\sigma, 0)$. (Заметим, что из ограниченности $\varkappa_T(\sigma', t)$ вытекают равенства $\frac{d\sigma}{d\sigma'} = 1$ и $\frac{d\Sigma}{d\Sigma'} = 1$.)

Так как кривые $s_T(\sigma, t)$ не выходят из δ-окрестности $\gamma_w(t)$, для любого фиксированного $t = t_0$ и $\sigma = \sigma_0$ найдется такая последовательность $\{T_n\}$, что числовая последовательность $\{s_{T_n}(\sigma_0, t_0)\}$ сходится к пределу. Так как функции $s_{T_n}(\sigma_0, t)$ описывают геодезические лучи проходящие при $t = 0$ через точку многообразия M с координатой σ_0, последовательность $\{s_{T_n}(\sigma_0, t)\}$ равномерно сходится на любом конечном интервале $[0, T]$. ($[0, T]$ может и содержать t_0!) Предельная кривая $s(\sigma_0, t)$ будет геодезической, и для нее справедливо соотношение (4') с $\varepsilon_\infty = \lim_{n\to\infty} \varepsilon_{T_n} = 0$ потому, что константы в формулах (2) и (3) не зависят от выбора T. По лемме 3.1 $s(\sigma_0, t)$ экспоненциально сближается с $\gamma_w(t)$. В работе [9] Морса и Хедлунда было доказано, что при отсутствии сопряженных точек две пересекающиеся геодезические не сближаются друг с другом при $t \to \infty$. Иными словами для фиксированного σ_0. предельная кривая единственна. По такой же причине для двух различных значений σ_1, σ_2 кривие $s(\sigma_1, t)$ и $s(\sigma_2, t)$ между собой не пересекаются. Назовем семейство построенных геодезических положительно асимптотическим геодезическим пучком. Первое утверждение теоремы непосредственно вытекает из (4'). Второе утверждение получается применением теоремы Лагранжа. Так как многообразие

M гладкое, равномерная сходимость последовательности крывых, проходящих через одну и ту же точку $x \in M$ влечет за собой сходимость последовательности их касательных векторов в этой точке. Поэтому существует предел

$$u(\sigma_0, 0) = \lim_{T_n \to \infty} u_{T_n}(\sigma_0, 0).$$

Известно, что

$$u_{T_n}(\sigma_0, 0) = \int_0^{\sigma_0} \varkappa_{T_n}(\sigma, 0)\, d\sigma \quad \text{т. е.} \quad u_{T_n}(\sigma_0, 0) = \sigma_0 \varkappa_{T_n}(\sigma_1, 0)$$

где $0 \leqq \sigma_1 \leqq \sigma_0$ (теорема Лагранжа). Но $|\varkappa^+(0) - \varkappa_{T_n}(\sigma, 0)| = o(\sigma)$ по неравенству (4′). Поэтому существует производная, равная предельному решению уравнения Риккати. Этим теорема полностью доказана.

Замечание: из доказательства вытекает, что соотношение между геодезическими лучами „быть асимптотическим", симметрично потому, что роль $\gamma_w(t)$ состоит в том, что вдоль нее $\lim_{T \to \infty} \int_0^T \frac{1}{T} K(t)\, dt < 0$. Но из экспоненциальной сходимости асимптотических геодезических вытекает то же самое для любого члена этого семейства.

Ортогональные траектории к пучку $s_w^+(\sigma, t)$ снабженные положительно направленными нормальными единичными векторами мы будем называть сжимающимися орициклами, и обозначать через $S_w^+(\sigma, t)$. (Таким образом $S_w^+(\sigma, 0)$ является орициклом проходящим через линейный элемент w.)

Аналогичным образом с помощью отрицательно предельных решений уравнения Якоби строятся расширяющие орициклы $S_w^-(\sigma, 0)$.

§ 5. Построение трансверсальных слоений и их абсолютная непрерывность

Обозначим для $\varepsilon > 0$ через W_ε множество линейных элементов для которых выполняется следующее условие:

Интегралы от гауссовой кривизны $K(t)$ вдоль геодезической $\gamma_w(t)$ определенный линейным элементом w, удовлетворяют неравенствам:

$$\left.\begin{aligned} &\frac{1}{T} \int_0^T K(t)\, dt < -\varepsilon, \\ &\frac{1}{T} \int_{-T}^0 K(t)\, dt < -\varepsilon, \end{aligned}\right\} \quad \text{если} \quad T > \frac{1}{\varepsilon}.$$

Очевидно, что множество W_ε измеримо, и $\mu\left(\bigcup_{\varepsilon < 1} W_\varepsilon\right) = \mu(W_-)$. Покроем W_ε окрестностями U_δ радиуса δ, где δ зависит от ε и будет указано ниже. Для каждой точки $w \in U_\delta \cap W_\varepsilon$ мы построим отрезки орициклов $S_w^+(\sigma, 0)$ и $S_w^-(\sigma, 0)$ длины $2d$ $(-d \leqq \sigma \leqq d)$, где d константа, фигурирующая в условии леммы 3.1

и зависит только от ε. Положим, по определению,

$$C_w^+ = \bigcup_t T_t S_w^+(\sigma, 0) = S_w^+(\sigma, t),$$

и

$$C_w^- = \bigcup_t T_t S_w^-(\sigma, 0) = S_w^-(\sigma, t).$$

Напомним, что при фиксированном $\sigma=\sigma_0$ кривые $S_w^+(\sigma_0, t)$ $(S_w^-(\sigma_0, t))$ представляют собой положительно (отрицательно) направленные геодезические лучи, снабженные единичными касательными векторами, а при фиксированном $t=t_0$ кривые $S_w^+(\sigma, t_0)$ $(S_w^-(\sigma, t_0))$ являются сжимающимися (расширяющимися) орициклами. Так как орициклы $S_w^+(\sigma, t_0)$ и $S_w^-(\sigma, t_0)$ при $-d \leqq \sigma \leqq d$ — гладкие кривые, не имеющие точек самопересечения (их геодезические кривизны $\varkappa^+(\sigma, t_0)$ и $\varkappa^-(\sigma, t_0)$ ограничены по модулю константой K — см. теорему 4.1), связные компоненты пересечений слоев C_w^+ и C_w^- с окрестностью U_δ радиуса $\delta<d$ являются гладкими многообразиями.

Напомним также обозначения $\pi(S^+(\sigma, t)) = s^+(\sigma, t)$ и $\pi(S^-(\sigma, t)) = s^-(\sigma, t)$ (π естественная проекция W на M).

Из определения следует, что

$$C^+_{T_t w} = T_t C_w^+ \quad C^-_{T_t w} = T_t C_w^-.$$

Таким образом мы можем рассматривать C_w^+ и C_w^-, как слои некоторых измеримых слоений Ξ^+ и Ξ^-, инвариантных относительно потока $\{T_t\}$. В следующем мы придадим этому утверждению более четкий смысл.

Напомним в начале два определения из [3].

Определение 1. Слоение Ξ называется измеримым, если разбиение множества $W_\varepsilon \cap U_\delta$ на связные компоненты слоев $C \in \Xi$ является измеримым, и условная мера на них абсолютно непрерывна относительно внутренней римановой меры.

Обозначим через $\tilde{\Xi}^-$ и $\tilde{\Xi}^+$ слоения слои которых представляют собой сами полные расширяющиеся и сжимающиеся орициклы $(S_w^-(\sigma, 0)$ и $S_w^+(\sigma, 0))$.

Определение 2. Слоение Ξ^+ абсолютно непрерывно относительно слоения $\tilde{\Xi}^-$, если для почти всех $w \in U_\delta \cap W_\varepsilon$ множество $S_w^-(\sigma, 0) \cap U_\delta \cap W_\varepsilon$ имеет положительную меру, и определенная ниже мера ν абсолютно непрерывна относительно внутренней римановой меры орицикла $S_w^-(\sigma, 0)$ $(d\sigma)$.

Пусть A измеримое подмножество на полном расширяющемся орицикле, принадлежашее пересечению орицикла $S_w^-(\sigma, 0)$ с множеством $U_\delta \cap W_\varepsilon$. Мера ν определяется, как объем высеченного множества из $U_\delta \cap W_\varepsilon$ слоями сжимающегося слоения Ξ^+, проходящими через точки множества A.

Аналогичным образом определяется абсолютная непрерывность слоения Ξ^- относительно слоения $\tilde{\Xi}^+$.

Нетрудно построить такое гладкое двумерное слоение $\Xi^{\Gamma\Lambda}$ окрестности U_δ, слои которого имеют следующий вид: каждому слою $C_\tau^{\Gamma\Lambda}$ отвечает гладкая вогнутая кривая C на M с геодезической кривизной большей, чем $\varkappa$ (где $\varkappa$-константа для которой справеливо, что решение $\varkappa(t)$ уравнения Риккати с начальным условием $\varkappa(t_0)=\varkappa$ остается положительным при $t>t_0$, и если $\tilde{C}$-оснащение положительно направленными нормальными векторами, то $C_\tau^{\Gamma\Lambda} = \bigcup_{0 \leqq t \leqq f(\tau)} T_t \tilde{C}$. Переменное $\tau \in [0, 1]$ — вещественный параметр на факторпрост-

ранстве слоев, $f(\tau)$ — гладкая функция от τ, и слои $C_\tau^{\Gamma\Lambda}$ гладко зависят от τ. Иными словами $C_\tau^{\Gamma\Lambda}$ есть рассеивающий (но не обязательно асимптотический) геодезический пучок: $C_\tau^{\Gamma\Lambda}=S^\tau(\sigma, t)$ $(0\leqq t\leqq f(\tau))$, где σ — естественный параметр, а t-параметр времени. При фиксированном $\sigma=\sigma_0$, кривая $s^\tau(\sigma_0, t)$ есть геодезический отрезок оснащенный с единичными касательными векторами, а при фиксированном $t=t_0$, кривая $s^\tau(\sigma, t_0)=\pi(S^\tau(\sigma, t_0))$ ортогональная траектория к пучку геодезических составляющих элемент $C_\tau^{\Gamma\Lambda}$, имеющая геодезическую кривизну $\varkappa^\tau(\sigma, t_0)$. Так как через каждый элемент $w\in U_\delta$ проходит ровно один слой $C_\tau^{\Gamma\Lambda}$, мы можем сопоставить w — в качестве координат — тройку $\{\tau_w, t_w, \sigma_w\}$ вещественных чисел $(0\leqq\tau_w\leqq 1, 0\leqq t_w\leqq f(\tau_w), |\sigma_w|\leqq\delta)$.

Если радиус δ окрестности U_δ достаточно малый, можно построить отображение $p_{\tau_2}^{t_2}$ и $p_{\tau_1,\tau_2}^{t_1,t_2}$ на множествах $S_w^-(\sigma, 0)\cap U_\delta\cap W_\varepsilon$ и $S^{\tau_1}(\sigma, t_1)\cap U_\delta\cap W_\varepsilon$ отображающие эти множества в $S^{\tau_2}(\sigma, t_2)\cap U_\delta\cap W_\varepsilon$ таким образом, чтобы образы и прообразы лежали на одном и том же сжимающем листе. Для избежания недоразуменый мы обозначим через σ_i параметр длины дуги на кривой $s^{\tau_i}(\sigma, t_i)$. Таким образом $p_{\tau_2}^{t_2}(\sigma)=\sigma_2$ $(p_{\tau_1,\tau_2}^{t_1,t_2}(\sigma_1)=\sigma_2)$ тогда и только тогда, если геодезические $s_w^-(\sigma, t)$ и $s^{\tau_2}(\sigma_2, t)$ $(s^{\tau_1}(\sigma_1, t)$ и $s^{\tau_2}(\sigma_2, t))$ положительно асимптотические.

Наша ближайшая цель доказать абсолютную непрерывность отображений $p_{\tau_2}^{t_2}$ и $p_{\tau_1,\tau_2}^{t_1,t_2}$. Доказательство проводится по тому же пути, что и в [2], с некоторыми упрощениями, обьясняющимися одномерностью слоений.

Отображение $p_{\tau_2}^{t_2}$ на множестве $A=S_w^-(\sigma, 0)\cap U_\delta\cap W_\varepsilon$ можно описать чисто геометрически. Именно, возьмем сжимающийся орицикл $s_w^+(\sigma, t')$ касающийся $s_w^-(\sigma, 0)$ в точке с координатой $\sigma=\sigma_0$, и сдвинем его до того момента t'', когда он каснется $s^{\tau_2}(\sigma_2, t_2)$ в некоторой точке $\sigma_2=p_{\tau_2}^{t_2}(\sigma_0)$. Так как $s_w^+(\sigma, t')$ имеет ограниченную снизу, неположительную кривизну, а кривизны кривых $s^{\tau_2}(\sigma_2, t_2)$ больше, чем $\varkappa$ — то эта конструкция в достаточно малой окрестности M_δ осуществима. (см. рисунок 1.)

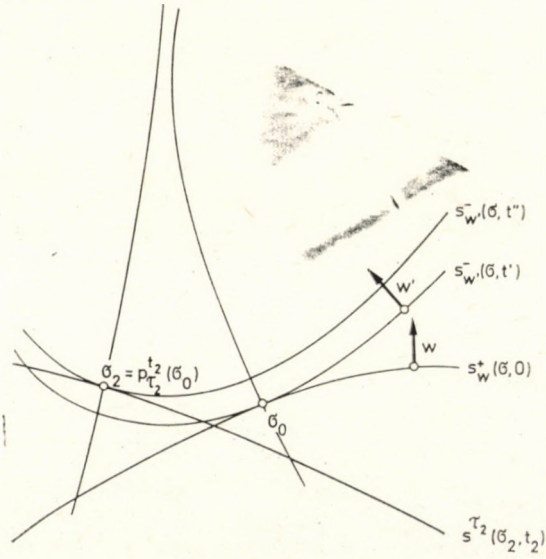


Рис. 1

Аналогичным образом можно описать геометрическим путем отображение $p_{\tau_1,\tau_2}^{t_1,t_2}(\sigma_1)$.

Перед доказательством абсолютной непрерывности отображения $p_{\tau_2}^{t_2}(\sigma)$ мы сделаем два замечания.

Замечание 1. Из компактности и локальной евклидовой структуры многообразия M вытекает существование таких универсальных (для M) констант r и α, что если w_1 и w_2 — два линейных элемента с носителями x_1 и x_2, $\varrho(x_1, x_2) < r$, а геодезическая $\gamma(t)$ соединяющая x_1 и x_2 ортогональна к w_1 а — угол β между $\gamma(t)$ и $\gamma_{w_2}(t)$ меньше, чем α, то тогда $\gamma_{w_1}(t)$ и $\gamma_{w_2}(t)$ пересекутся, следовательно не могут быть асимптотическими. Угол α может быть сколь угодно близким к $\frac{\pi}{2}$ если r достаточно мало. (см. рисунок 2.)

Замечание 2. Пусть $\gamma_{w_1}(t)$ и $\gamma_{w_2}(t)$ асимптотические ($t \to -\infty$) геодезические. Линейные элементы w_1, w_2 и геодезические $\gamma_{w_1}(t)$, $\gamma_{w_2}(t)$, $\gamma(t)$ те же, что в замечании 1, и угол $\beta > \alpha$. Проведем через точку x_1 геодезическую ортогональную к $\gamma_{w_2}(t)$, и назовем их точку пересечения x_3. Для длины дуги d любой выпуклой кривой без самопересечения, выходящей из точки x_1 и ортогональной к $\gamma_{w_1}(t)$ и $\gamma_{w_2}(t)$, справедлива оценка:

$$\varrho(x_1, x_3) \leqq d \leqq \varrho(x_1, x_2) + \varrho(x_2, x_3),$$

т. е.

$$\frac{\min d}{\max d} \geqq \frac{\varrho(x_1, x_3)}{\varrho(x_1, x_2) + \varrho(x_2, x_3)} \geqq \text{const} \tag{1}$$

так как $\beta > \alpha$ и $\varrho(x_1, x_2) < r$, где r и α универсальные константы. Если $\gamma_1'(t)$ и $\gamma_2'(t)$ отрезки геодезических, близкие (в окрестности радиуса $2r$ точек x_1, x_2) к $\gamma_{w_1}(t)$ и $\gamma_{w_2}(t)$ соответственно — например их расстояние в 8 раз меньше,

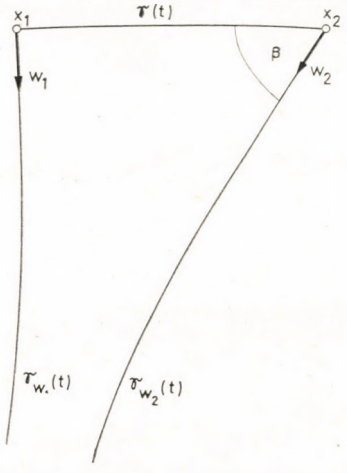


Рис. 2

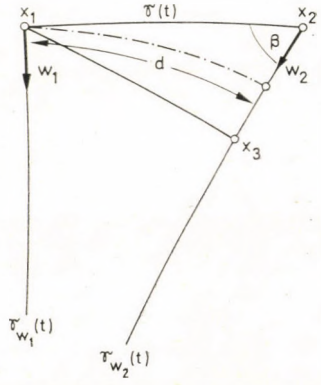


Рис. 3

чем $\varrho(x_1, x_2)$, то тогда существует такая константа K_1, что для длины дуги d' выпуклых, ортогональных к $\gamma_1'(t)$ и $\gamma_2'(t)$ кривых расположенных между $\gamma_1'(t)$ и $\gamma_2'(t)$,

$$K_1^{-1} \leqq \frac{\max d'}{\min d'} \leqq K_1$$

(см. рисунки 3. и 4.)

Перейдем теперь к доказательству абсолютной непрерывности отображения $p_{\tau_2}^{t_2}(\sigma)$. Пусть множество A выбрано, как и выше. Рассмотрим любое подмножество $A' \subset A$ и предположим, что оно содержится в отрезке Δ длины h. Покажем, что если $\sigma_0 \in \bar{A}'$, т. е. является предельной точкой множества A' — то и для $S_w^-(\sigma_0, 0)$ выполняется условие определения множества W_ε. Пусть $\sigma_0 = \lim_{n\to\infty} \sigma_n$ $(\sigma_n \in A')$. Фиксируем любой момент времени $T > \frac{1}{\varepsilon}$. Если для всех n $\frac{1}{T}\int_0^T K\left(s_w^-(\sigma_n, t)\right) \cdot dt \leqq -\varepsilon$, то и $\frac{1}{T}\int_0^T K\left(s_w^-(\sigma_0, t)\right) \cdot dt \leqq -\varepsilon$, благодаря непрерывности гауссовой кривизны $K(x)$ на M — и равномерной сходимости геодезических $s_w^-(\sigma_n, t)$ к $s_w^-(\sigma_0, t)$ на любом конечном отрезке времени.

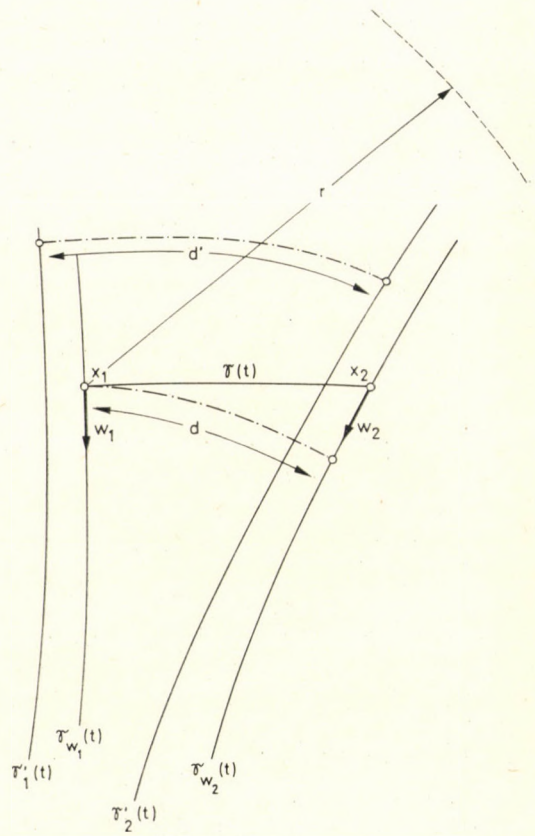


Рис. 4

Из теоремы Кантора вытекает, что кроме счетного числа изолированных точек, каждая точка множества $\bar{A}'$ является предельной точкой. Обозначим дополнение в $\bar{A}'$ к этому счетному множеству изолированных точек через $\bar{\bar{A}}'$.

Выберем теперь момент времени T столь большим, чтобы выполнялись следующие условия:

α) для всех фиксированных $\sigma_0 \in \bar{\bar{A}}'$ и $t > T$ геодезические $s_w^-(\sigma_0, t)$ и $s^{\tau_2}\left(p_{\tau_2}^{t_2}(\sigma_0), t\right)$ были бы ближе друг к другу, чем t^{-5}. Это возможно потому, что $s_w^-(\sigma_0, t)$ и $s^{\tau_2}\left(p_{\tau_2}^{t_2}(\sigma_0), t\right)$ положительно асимптотические, и в силу леммы 3.1 они сближаются с равномерной по $\bar{A}'$ экспоненциальной скоростью.

β) Для любых двух элементов σ', σ'' множества $\bar{\bar{A}}'$ из неравенства $d_t\left(s_w^-(\sigma', t), s_w^-(\sigma'', t)\right) < t^{-4}$ (d_t означает длину дуги сдвинутого на время t отрезка

$[\sigma', \sigma'']$ орицикла $s_w^-(\sigma, 0)$) вытекало бы, что

$$\left| \int_0^t (\varkappa^-(\sigma', \tau) - \varkappa^-(\sigma'', \tau)) d\tau \right| < \frac{1}{t}.$$

γ) Аналогичное неравенство было справедливо для разности

$$\int_0^t (\varkappa^{\tau_2}(p_{\tau_2}^{t_2}(\sigma'), \tau) - \varkappa^{\tau_2}(p_{\tau_2}^{t_2}(\sigma''), \tau)) d\tau,$$

где σ' и σ'' те же самые что в β), а функции $\varkappa^{\tau_2}(p_{\tau_2}^{t_2}(\sigma'), t)$ и $\varkappa^{\tau_2}(p_{\tau_2}^{t_2}(\sigma''), t)$ указаны в определении гладкого слоения $\Xi^{\Gamma\Lambda}$.

Возьмем любую точку $\sigma \in \bar{\bar{A}}'$ и для некоторого $t > T$ найдем точки $\sigma_l^t < \sigma$ и $\sigma_r^t > \sigma$ так, чтобы выполнялись точные равенства:

$$d_t(s_w^-(\sigma, t), s_w^-(\sigma_l^t, t)) = t^{-5} \quad \text{и} \quad d_t(s_w^-(\sigma, t), s_w^-(\sigma_r^t, t)) = t^{-5}.$$

В силу того, что $\bar{\bar{A}}'$ лежит на расширяющемся орицикле, $\sigma_l^t \nearrow \sigma$, $\sigma_r^t \searrow \sigma$ при $t \to \infty$. Поэтому для $t > T$ существуют T_r и T_l такие, что

$$\sigma_r^{T_r} \in \bar{\bar{A}}' \quad \text{или} \quad (\sigma, \sigma_r^{T_r}] \cap \bar{\bar{A}}' = \emptyset,$$

$$\sigma_l^{T_l} \in \bar{\bar{A}}' \quad \text{или} \quad [\sigma_l^{T_l}, \sigma) \cap \bar{\bar{A}}' = \emptyset.$$

Множество $\bar{\bar{A}}'$ мы можем покрыть открытыми интервалами вида $(\sigma_l^{T_l}, \sigma_r^{T_r})$. По теореме Гейне—Бореля уже конечное число из них покрывает $\bar{\bar{A}}'$. Мы можем перейти к конечной покрывающей системе замкнутых отрезков, которые могут иметь только общие концы. Эти отрезки мы будем обозначить через $[\sigma_l^{T_l^*}, \sigma_r^{T_r^*}]$. Возьмем два соседних отрезка $(\sigma_l^{T_l}, \sigma_r^{T_r})$ и $(\sigma_l'^{T_l'}, \sigma_r'^{T_r'})$. Предположим, что $\sigma_l'^{T_l'} < \sigma_r^{T_r}$. Если один из них не принадлежит к $\bar{\bar{A}}'$, тогда между σ и σ' нет точек множества $\bar{\bar{A}}'$, поэтому новыми концевыми точками могут служить сами σ и σ'. Если $\sigma_l^{T_l}$ и $\sigma_r'^{T_r'} \in \bar{\bar{A}}'$, тогда, увеличив время, найдем такое $T_l'^* > T$, что $\sigma_l'^{T_l'^*} = \sigma_r^{T_r}$. (Заметим, что $\sigma_r'^t \searrow \sigma'$). После этой поправки покрывающей системы множество $\bar{\bar{A}}'$ покрыт конечным числом интервалов Δ_i, концевые точки которых принадлежат к множеству $\bar{\bar{A}}_i$, и содержит еще счетное число точек, которыми с точки зрения абсолютной непрерывности можно пренебречь.

Оценим общую длину образов Δ_i при отображении $p_{\tau_2}^{t_2}(\sigma)$. Для простоты введем следующие обозначения:

$$T_i = \begin{cases} T_r & \text{если} \quad \Delta_i = [\sigma, \sigma_r^{T_r}], \\ T_l & \text{если} \quad \Delta_i = [\sigma_l^{T_l}, \sigma] \end{cases}$$

а $d_i^{T_i}$ обозначает длину дуги сдвинутого отрезка Δ_i на время T_i. Кроме того, пусть Δ_i^* образ отрезка Δ_i при отображении $p_{\tau_2}^{t_2}(\sigma)$, и $d_i^{*T_i}$ длина дуги сдвину-

того на время T_i отрезка Δ_i^*. Тогда

$$\frac{d_i^{*T_i}}{d_i^{T_i}} = \frac{\int\limits_{\Delta_i^*} \exp\left\{\int\limits_0^{T_i} \varkappa^{\tau_2}(\sigma_2, t)\,dt\right\} d\sigma_2}{\int\limits_{\Delta_i} \exp\left\{\int\limits_0^{T_i} \varkappa^{-}(\sigma, t)\,dt\right\} d\sigma}.$$

Вследствие теоремы о среднем, существуют такие $\bar{\sigma}$ и $\bar{\sigma}_2$ ($\bar{\sigma}_2$ не обязательно равно $p_{\tau_2}^{t_2}(\bar{\sigma})$!), что

$$\frac{d_i^{*T_i}}{d_i^{T_i}} = \frac{d(\Delta_i^*) \exp\left\{\int\limits_0^{T_i} \varkappa^{\tau_2}(\bar{\sigma}_2, t)\,dt\right\}}{d(\Delta_i) \exp\left\{\int\limits_0^{T_i} \varkappa^{-}(\bar{\sigma}, t)\,dt\right\}}.$$

Левая часть ограничена в силу замечания 2 и условия α),

а

$$\frac{\exp\left\{\int\limits_0^{T_i} \varkappa^{\tau_2}(\bar{\sigma}_2, t)\,dt\right\}}{\exp\left\{\int\limits_0^{T_i} \varkappa^{-}(\bar{\sigma}, t)\,dt\right\}} \geqq \text{const} \frac{\exp\left\{\int\limits_0^{T_i} \varkappa^{\tau_2}(p_{\tau_2}^{t_2}(\sigma), t)\,dt\right\}}{\exp\left\{\int\limits_0^{T_i} \varkappa^{-}(\sigma, t)\,dt\right\}}$$

для любой точки $\sigma \in A \cap \Delta_i$ потому, что вследствие условия

β) и γ), $\int\limits_0^{T_i} (\varkappa^{\tau_2}(p_{\tau_2}^{t_2}(\bar{\sigma}), t) - \varkappa^{\tau_2}(p_{\tau_2}^{t_2}(\sigma), t))\,dt$ и $\int\limits_0^{T_i} (\varkappa^{-}(\bar{\sigma}, t) - \varkappa^{-}(\sigma, t))\,dt$

имеют порядок $\frac{1}{T_i}$. Из леммы 2.1 следует, что

$$\left| \int\limits_0^{T_i} (\varkappa^{-}(\sigma, t) - \varkappa^{\tau_2}(\sigma_2, t))\,dt \right| < \text{const}\ (\text{const} — \text{абсолютная на } W_\varepsilon \cap U_\delta)$$

потому, что геодезические $s_w^{-}(\sigma, t)$ и $s^{\tau_2}(p_{\tau_2}^{t_2}(\sigma), t)$ экспоненциально сближаются при $t \to \infty$. Отсюда вытекает, что $\frac{d(\Delta_i^*)}{d(\Delta_i)} < \text{const}$ (const — абсолютная на A). Иными словами, если множество A' содержится в интервале Δ длины h, то тогда за исключением счетного числа точек, множество $p_{\tau_2}^{t_2}(A')$ можно вложить в обьединение конечного числа интервалов общей длины const h. Последнее утверждение влечет за собой абсолютную непрерывность отображения $p_{\tau_2}^{t_2}(\sigma)$.

Аналогичным образом доказывается абсолютная непрерывность отображений $p_{\tau_1,\tau_2}^{t_1,t_2}(\sigma_1)$. После того, как абсолютная непрерывность отображений $p_{\tau_1,\tau_2}^{t_1,t_2}(\sigma_1)$ установлена, доказательство измеримости слоения буквально совпадает с доказательством работы [2]:

Пусть $B \subset U_\delta \cap W_\varepsilon$ — произвольное измеримое множество положительной меры. Запишем инвариантную меру μ на U_δ в координатах τ, σ, t, связанных с гладким слоением $\Xi^{\Gamma\Lambda}$:

$$\mu(B) = \int_{U\delta} r(\tau, \sigma, t)\chi_B(\tau, \sigma, t)\,d\tau\,d\sigma\,dt =$$

$$= \int_{U\delta/\Xi^{\Gamma\Lambda}} d\tau\,dt \int_{C^{\Gamma\Lambda}_{\tau,t}} r(\tau, \sigma, t)\chi_B(\tau, \sigma, t)\,d\sigma =$$

$$= \int_{U\delta/\Xi^{\Gamma\Lambda}} d\tau\,dt \int_{C^{\Gamma\Lambda}_{\tau_0,t_0}} r\big(\tau, p^{t_0,t}_{\tau_0,\tau}(\sigma_0), t\big)\chi_B\big(\tau, p^{t_0,t}_{\tau_0,\tau}(\sigma_0), t\big)I(\sigma_0, t)\,d\sigma_0$$

$I(\sigma_0, t)$ — положительная, измеримая функция, возникающая из абсолютной непрерывности отображений $p^{t_0,t}_{\tau_0,\tau}(\sigma_0)$. При фиксированной точке w^* кривой $C^{\Gamma\Lambda}_{\tau_0,t_0} = S^{\tau_0}(\sigma_0, t_0)$ с координатой σ_0^* точки $p^{t_0,t}_{\tau_0,\tau}(\sigma_0^*)$ пробегают локальный слой $C^+_{\sigma_0^*}$ сжимающегося слоения Ξ^+. С помощью теоремы Фубини мы получаем равенство

$$\mu(B) = \int_{C^{\Gamma\Lambda}_{\tau_0,t_0}} d\sigma_0 \int_{C^{+*}_{\sigma_0}} r\big(\tau, p^{t_0,t}_{\tau_0,\tau}(\sigma_0), t\big)\chi_B\big(\tau, p^{t_0,t}_{\tau_0,\tau}(\sigma_0), t\big)I(\sigma|\tau, t)\,d\tau\,dt.$$

Эта последняя формула показывает, что на слоях $C^+_{\sigma_0^*}$ слоения Ξ^+ возникает условная мера, абсолютно непрерывная относительно естественной римановой меры.

Для доказательства абсолютной непрерывности слоения Ξ^+ относительно слоения $\tilde{\Xi}^-$ мы замечаем следующее: первое условие в определении 2 для множества $W_\varepsilon \cap U_\delta$ может быть не выполнено потому, что в некоторых точках σ слоя $S^-_w(\sigma, 0)$, хотя отображение $\big(p^0_{\tau_2}(\sigma)\big)^{-1}$ можно определить, но его образ не будет принадлежать к W_ε (не выполняется условие определения множества W_ε с константой ε). Поэтому мы пополним множество W_ε таким образом, чтобы избежать эту трудность, но доказательство абсолютной непрерывности отображений $p^{t_2}_{\tau_2}(\sigma)$ и $p^{t_1,t_2}_{\tau_1,\tau_2}(\sigma_1)$ осталось справедливым. Именно, пусть W^*_ε множество всех линейных элементов $w \in U_\delta$, для которых существуют такие линейные элементы w_1 и $w_2 \in U_\delta \cap W_\varepsilon$, что $\gamma_w(t)$ и $\gamma_{w_1}(t)\big(\gamma_w(t)$ и $\gamma_{w_2}(t)\big)$ будут положительно (отрицательно) асимптотическими геодезическими. Множество W^*_ε является замкнутым в следующем смысле: если для некоторого $w \in U_\delta$ существуют $w_1^* \in U_\delta \cap W_\varepsilon$ и $w_2^* \in U_\delta \cap W_\varepsilon$ такие, что $\gamma_w(t)$ и $\gamma^{(t)}_{w_1^*}$ $\big(\gamma_w(t)$ и $\gamma^{(t)}_{w_2^*}\big)$ будут положительно (отрицательно) асимптотическими — то найдутся и $w_1, w_2 \in U_\delta \cap W_\varepsilon$ с тем же свойством. Так как асимптотические лучи сходятся с равномерной на W_ε экспоненциальной скоростью, если линейные элементы, определяющие их лежат в достаточно малой окрестности U_δ, то условие определения множества W_ε выполняется и для W^*_ε, но быть может с другим ε'. (Если $\frac{1}{T}\int_0^T K(t)dt < -\varepsilon$ при $T > \frac{1}{\varepsilon}$, то существует такое ε', что $\frac{1}{T}\int_0^T \big(K(t) + e^{-\text{const}\,t}\big)dt < -\varepsilon'$ при $T > \frac{1}{\varepsilon'}$. Из леммы 3.1 следует, что ε' одно и то же для всего $W_\varepsilon \cap U_\delta$.) Таким образом доказательство абсолютной непрерывности

отображения $p_{t_2}^{t_2}(\sigma)$ проходит и для W_ε^*. Заметим, что $p_{\tau_2}^{t_2}(\sigma)$ определено и в новых точках потому, что в геометрической конструкции была важна только достаточно большая кривизна кривых $S^{\tau_2}(\sigma_2, t_2)$, составляющих слои гладкого слоения.

Предположим, что w является точкой плотности множества W_ε^*. Тогда существует такой слой $C_{\tau_2}^{\Gamma\Lambda}$ гладкого слоения $\Xi^{\Gamma\Lambda}$, что для некоторого t_2 кривая $S^{\tau_2}(\sigma_2, t_2)$ в достаточно малой окрестности $U_{\delta'}$ $(\delta' < \delta)$ столь близко к орициклу $S_w^-(\sigma, 0)$ (w-точка плотности), чтобы

$$[p_{\tau_2}^{t_2}]^{-1}\left(U_\delta \cap W_\varepsilon^* \cap S^{\tau_2}(\sigma_2, t_2)\right) \subset S_w^-(\sigma, 0) \cap W_\varepsilon^*.$$

Благодаря гладкости слоения $\Xi^{\Gamma\Lambda}$ мы можем предположить, что множество $U_{\delta'} \cap W_\varepsilon^* \cap S^{\tau_2}(\sigma_2, t_2)$ имеет положительную меру. Тогда из абсолютной непрерывности $p_{\tau_2}^{t_2}(\sigma)$ вытекает положительность меры множества $S_w^-(\sigma, 0) \cap W_\varepsilon^*$.

Абсолютная непрерывность меры ν относительно меры $d\sigma$ доказывается также, как в работе [2]. Так как для любого $\tau \in [0, 1]$ отображение $p_{\tau_2}^{t_2}(\sigma)$ абсолютно непрерывно, если множество $A \subset S_w^-(\sigma, 0) \cap W_\varepsilon^* \cap U_\delta$ имеет меру 0, то и для каждого $\tau \in [0, 1]$ и множество $p_\tau^t(A)$ имеет меру 0. Но слоение $\Xi^{\Gamma\Lambda}$ гладкое, поэтому из последнего утверждения следует, что длина множества, высеченного из множества $U_\delta \cap W_\varepsilon^*$ слоями сжимающегося слоения Ξ^+, проходящими через точки A равна 0.

Аналогичным образом доказывается измеримость слоения Ξ^- и абсолютная непрерывность слоения Ξ^- относительно $\tilde{\Xi}^+$.

Теперь мы уже в состоянии доказать основную теорему этой работы.

Теорема 5.1. *Почти каждая эргодическая компонента геодезического потока* $\{T_t\}$ *на множестве* W_- *имеет положительную меру.*

Доказательство. Возьмем минимальное инвариантное множество относительно потока $\{T_t\}$, содержащее $U_\delta \cap W_\varepsilon^*$:

$$\bigcup_t T_t(U_\delta \cap W_\varepsilon^*).$$

Исходя из измеримости слоений Ξ^+ и Ξ^-, и абсолютной непрерывности слоения Ξ^+ (Ξ^-) относительно слоения $\tilde{\Xi}^-$ ($\tilde{\Xi}^+$) на множестве $U_\delta \cap W_\varepsilon^*$, методом Э. Хопфа можно доказать, что каждая инвариантная относительно потока $\{T_t\}$ функция $f(w)$ постоянна почти всюду на $W_\varepsilon^* \cap U_\delta$ (см. работы [5], [7]). По построению W_ε^*, для почти каждого элемента $w \in W_\varepsilon^* \cap U_\delta$ существует такое положительное число τ_w, что $T_t w \in W_\varepsilon^* \cap U_\delta$ для $|t| < \tau_w$. Поэтому если $\mu(U_\delta \cap W_\varepsilon^*) > 0$ из инвариантности $f(w)$ вытекает, что она постоянна почти всюду на множестве $\bigcup_t T_t(U_\delta \cap W_\varepsilon^*)$.

Обозначим через W_0 обьединение эргодических компонент нулевой меры, и предположим, что $\mu(W_0) = \varepsilon_0$. При достаточно малом ε, мера $\mu(W_- \setminus W_\varepsilon) < \frac{\varepsilon_0}{2}$. Мы доказали, что для почти каждого w, из $w \in U_\delta \cap W_\varepsilon^*$ — где $\mu(U_\delta \cap W_\varepsilon^*) > 0$ — вытекает положительность меры эргодической компоненты содержащей w. Поэтому, если линейный элемент $w \in W_0$, и он не принадлежит к некоторому множеству меры 0, то тогда он не может быть точкой плотности множества

W_ε. Так как почти каждая точка множества W_ε является точкой плотности, $\mu(W_0) < \frac{\varepsilon_0}{2}$. Теорема доказана.

Тот факт, что на каждой эргодической компоненте положительной меры поток $\{T_t\}$ будет K-потоком, доказывается также, как и в [3]. Тем самым требуемое утверждение установлено.

Пользуясь случаем, автор хотел бы выразить свою глубокую благодарность Я. Г. Синаю за постановку задачи, которая позволила автору статьи ознакомиться с мощным методом трансверсальных слоений.

ЛИТЕРАТУРА

[1] Аносов, Д. В.: Геодезические потоки на замкнутых римановых многообразиях отрицательной кривизны, *Труды Матем. Инст. им. Стеклова* **90** (1967).
[2] Аносов, Д. В. и Синай, Я. Г.: Некоторые гладкие динамические системы, *УМН* **22** № 5 (1967), 107—172.
[3] Синай, Я. Г.: Классические динамические системы со счетнократным лебеговским спектром. II. *Изв. АН, сер. матем.* **30** № I (1966), 15—68.
[4] Синай, Я. Г.: Марковские разбиения и У-системы, *Функциональный анализ и его приложения* **2** № 1 (1968), 64—89.
[5] Хопф, Э.: Статистика геодезических линий на многообразиях отрицательной кривизны, *УМН* **4** № 2 (1949), 129—170.
[6] Green, L. W.: Surfaces without conjugate points, *Transactions of Amer. Math. Soc.* **62**, 5. (1953), 529—546.
[7] Hopf, E.: Statistik der Lösungen geodätischer Probleme von unstabilen Typus, II. *Math. Ann.* **117**, 4. (1940), 590—608.
[8] Klingenberg, W.: Fluss auf Mannigfaltigkeiten vom hyperbolischen Typ, preprint.
[9] Morse, M. and Hedlund, G. A.: Manifolds without conjugate points, *Transactions of Amer. Math. Soc.* **51**, 3. 362—386.

(Поступила: 4-ого января, 1972 г.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 79—89.

ON THE DISTRIBUTIONS OF R_{mn}^+ AND (D_{mn}^+, R_{mn}^+)*

by

G. P. STECK and G. J. SIMMONS

1. Summary. Consider the two-sample problem where F_m and G_n are the two empirical distribution functions. The difference $F_m(x)-G_n(x)$ changes only at an x_i, $i = 1, 2, \ldots, m+n$, corresponding to one of the observations. Let R_{mn}^+ denote the subscript i for which $F_m(x_i)-G_n(x_i)$ first achieves its maximum value D_{mn}^+. Let p denote the greatest common divisor of m and n and define $q = (m+n)/p$. We prove that the point probabilities for R_{mn}^+ are non-increasing and equal in blocks of length q and give a formula for their values. In particular, if m and n are relatively prime then R_{mn}^+ is uniformly distributed on the integers $1, 2, \ldots, m+n$. We also give a formula for the point probabilities of the vector (D_{mn}^+, R_{mn}^+).

2. Introduction. Let $X_1 \leqq X_2 \leqq \cdots \leqq X_m$ be the order statistics from a sample of m independent, identically distributed, random variables with a continuous distribution function F and let $mF_m(x)$ denote the number of X's less than or equal to x. Similarly, let $Y_1 \leqq Y_2 \leqq \cdots \leqq Y_n$ be the order statistics from a second sample of n independent, identically distributed, random variables with a continuous distribution function G and let $nG_n(y)$ denote the number of Y's less than or equal to y. Let R_i denote the rank of X_i in the ordered combined sample. Finally, consider the case $F=G$.

The Smirnov one-sided statistic is

$$D_{mn}^+ = \sup_t \{F_m(t)-G_n(t)\} = (1/mn) \max_{1 \leqq k \leqq m} \big(k(m+n)-mR_k\big).$$

If $mnD_{mn}^+=d$, let R_{mn}^+ be the smallest value of k for which $k(m+n)-mR_k = d$. The possible values of R_{mn}^+ are the integers $1, 2, \ldots, m+n$. Let

$$P_{mn}(r) = P(R_{mn}^+ = r),$$

$$Q_{mn}(d, r) = P(mnD_{mn}^+ = d, R_{mn}^+ = r).$$

Several authors have discussed the distribution of R_{mn}^+ in certain special cases. VINCZE (1957) found a formula for $Q_{nn}(d, r)$ and observed that $P_{nn}(2r-1) = P_{nn}(2r)$, $r=1, 2, \ldots, n$, and SARKADI (1961) showed that $P_{nn}(r) \geqq P_{nn}(r+1)$. STECK (1969) showed that if m and n are relatively prime then $mnD_{mn}^+=d$ implies at most one solution to the equation $k(m+n)-mR_k = d$; hence, in this case, at most one value of R_{mn}^+ is possible for a given value of d. GELLER (1971) proves something equivalent to

* This work was supported by the United States Atomic Energy Commission.

the fact that the limiting distribution of $R_{mn}^{+}/(m+n)$ is uniform on [0, 1] provided $\lim_{\substack{m\to\infty \\ n\to\infty}} m/n$ is positive and finite.

It is the purpose of this paper to give the exact distributions of $(mnD_{mn}^{+}, R_{mn}^{+})$ and R_{mn}^{+} for finite m and n. Our principal results are the following.

THEOREM. *Let* $p=\gcd(m, n)$ *and* $q=(m+n)/p$, *i.e.*, $m=ap$ *and* $n=bp$ *with* $\gcd(a, b)=1$. *Then* $P_{mn}(r) \geqq P_{mn}(r+1)$ *for* $r=1, 2, \ldots, m+n-1$, *and the* $\{P_{mn}(r)\}$ *are equal in blocks of length* q *with values* $\binom{m+n}{n} P_{mn}(r) = M_{mn}(r)$ *given by* (3. 2).

COROLLARY. If $\gcd(m, n)=1$ then R_{mn}^{+} is uniformly distributed on the integers $1, 2, \ldots, m+n$.

We will make no distinction between m and n except when we consider n to be a multiple of m. However, the computations will be easier when $m<n$ as the determinant for $M_{mn}(r)$ is $m-1$ by $m-1$.

The proof of the theorem consists of several lemmas which are conclusions about the random walk model for the two-sample problem. This random walk is a lattice path from (0, 0) to (n, m) with the kth step being one unit up or one unit to the right according as the kth observation in the ordered combined sample is an X or a Y. It is important to note that after the kth step up the path is at the point (R_k-k, k) and that $k(m+n)-mR_k$ is m times the horizontal distance from (R_k-k, k) to the diagonal. Thus mnD_{mn}^{+} is m times the maximum horizontal distance from the path to the diagonal. In what follows we shall use the word „distance" to mean horizontal distance. Distances to the diagonal will be considered positive if the point is to the left of the diagonal and negative otherwise.

The following results also have an interpretation in terms of the ballot problem. In that context A scores m votes, B scores n votes, all possible $\binom{m+n}{n}$ vote sequences are possible, and a path is interpreted as a vote sequence.

We will also wish to refer to the following result due to STECK (1969).

LEMMA 1. *Let* $b_1 \leqq b_2 \leqq \cdots \leqq b_m$ *and* $c_1 \leqq c_2 \leqq \cdots \leqq c_m$ *be sequences of integers such that* $i \leqq b_i \leqq c_i \leqq n+i$. *Then* $\binom{m+n}{n} P(b_i \leqq R_i \leqq c_i, \text{ all } i) = \det\left\{\binom{c_i-b_j+j-i+1}{j-i+1}_{+}\right\}$, *where* $\binom{x}{r}_{+} = \binom{\max(x, 0)}{r}$.

3. The distribution of R_{mn}^{+}. LEMMA 2a. *The number of paths from* (0, 0) *to* (n, m) *through the point* (x, y), $x \leqq ny/m$, *that are never farther from the diagonal than they are at* (x, y) *is the same as the number of paths from* (0, 0) *to* (n, m) *through* $(n-x, m-y)$ *that are never above diagonal.*

PROOF. Let P_1 be a path from (0, 0) to (x, y) and P_2 be a path from (x, y) to (n, m) such that the combined path P_1P_2 is never farther from the diagonal than it is at (x, y). Let P_1' be P_1 shifted up $m-y$ units and shifted right $n-x$ units. Then P_1' is a path from $(n-x, m-y)$ to (n, m). Similarly, let P_2' be P_2 shifted down y units and shifted left x units. Then P_2' is a path from (0, 0) to $(n-x, m-y)$ and the paths P_1P_2 and $P_2'P_1'$ are in 1—1 correspondence. Finally, $P_2'P_1'$ is never above the diagonal. This follows since if $P_2'P_1'$ were above the diagonal at (a, b) then P_1P_2 is farther from

the diagonal at the shifted image of (a, b) than it was at (x, y). But this is contrary to the initial assumption. ■

LEMMA 2b. *The number of paths from* $(0, 0)$ *to* (n, m) *through* (x, y) *that attain their maximum distance from the diagonal for the first time at* (x, y) *is the same as the number of paths from* $(0, 0)$ *to* (n, m) *through* $(n-x, m-y)$ *that are never above the diagonal before* $(n-x, m-y)$ *and never touch the diagonal afterwards. Here "touch" means a path reaches the diagonal and then leaves it; hence, in this sense, a path does not touch the diagonal at the point* (n, m).

PROOF. The proof is almost identical to that for Lemma 2a. ■

LEMMA 3a. *The number of paths from* $(0, 0)$ *to* (n, m) *which are never farther from the diagonal than they are on the* k^{th} *step is the same as the number of paths from* $(0, 0)$ *to* (n, m) *that are never above the diagonal.* (Note that this number is independent of k.)

PROOF. Consider all points (x, y) such that $x+y = k$, $x \leqq ny/m$, $1 \leqq x \leqq n$, $1 \leqq y \leqq m$. The set of required paths is the union of the disjoint subsets of paths through each of points (x, y). By Lemma 2a the paths in each of these subsets are in 1—1 correspondence with those in the (disjoint) sets of paths from $(0, 0)$ to (n, m) through $(n-x, m-y)$ that are never above the diagonal. Hence the elements in the set of required paths are in 1—1 correspondence with the set of paths that are on or below the diagonal on the $(m+n-k)^{\text{th}}$ step and are never above the diagonal. ■

LEMMA 3b. *The number of paths from* $(0, 0)$ *to* (n, m) *which attain their maximum distance from the diagonal for the first time on the* k^{th} *step is the same as the number of paths from* $(0, 0)$ *to* (n, m) *that are never above the diagonal and, moreover, never touch the diagonal after the* $(m+n-k)^{\text{th}}$ *step.*

PROOF. The proof is almost identical to that for Lemma 3a. ■

LEMMA 4a. *The number of paths from* $(0, 0)$ *to* (n, m) *that are never above the diagonal is the determinant*

$$N_{mn} =$$

$$= \begin{vmatrix} \binom{n-\left\langle \frac{n}{m} \right\rangle+1}{1} & \binom{n-\left\langle \frac{2n}{m} \right\rangle+1}{2} & \binom{n-\left\langle \frac{3n}{m} \right\rangle+1}{3} & \dots & \binom{n-\left\langle \frac{(m-1)n}{m} \right\rangle+1}{m-1} \\ 1 & \binom{n-\left\langle \frac{2n}{m} \right\rangle+1}{1} & \binom{n-\left\langle \frac{3n}{m} \right\rangle+1}{2} & \dots & \binom{n-\left\langle \frac{(m-1)n}{m} \right\rangle+1}{m-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \binom{n-\left\langle \frac{(m-1)n}{m} \right\rangle+1}{1} \end{vmatrix}_{m-1 \times m-1}, \tag{3.1}$$

where $\langle x \rangle$ *is the smallest integer greater than or equal to* x.

PROOF. A path from $(0, 0)$ to (n, m) is never above the diagonal if and only if $R_k - k \geqq nk/m$, $k = 1, 2, \ldots, m$. The lemma follows by taking $c_i - i = n$ and $b_j - j = \langle nj/m \rangle$ in Lemma 1. The determinant there is $m \times m$ but the last column is $(0, 0, \ldots, 0, 1)'$. ■ If $n = sm$, s an integer, we will see later that $N_{mn} = \binom{m+n}{n}/(n+1)$.

LEMMA 4b. *The number of paths from* $(0, 0)$ *to* (n, m) *that are never above the diagonal and, moreover, never touch the diagonal after the* $(m+n-r)^{\text{th}}$ *step is*

$$m - \left\langle \frac{mr}{m+n} \right\rangle \text{ columns with } \langle \cdot \rangle \quad \left\langle \frac{mr}{m+n} \right\rangle - 1 \text{ columns with } \{ \cdot \}$$

$$M_{mn}(r) = \begin{vmatrix} \overbrace{\binom{n - \left\langle \frac{n}{m} \right\rangle + 1}{1}} & \overbrace{\binom{n - \left\langle \frac{2n}{m} \right\rangle + 1}{2}} & \cdots & \overbrace{\binom{n - \left\{ \frac{(m-1)n}{m} \right\} + 1}{m-1}} \\ 1 & \binom{n - \left\langle \frac{2n}{m} \right\rangle + 1}{1} & \cdots & \binom{n - \left\{ \frac{(m-1)n}{m} \right\} + 1}{m-2} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \binom{n - \left\{ \frac{(m-1)n}{m} \right\} + 1}{1} \end{vmatrix}_{m-1 \times m-1}, \tag{3.2}$$

$$= N_{mn} - P(D^+_{mn} = 0, R^+_{mn} < r),$$

where $\langle x \rangle$ *and* $\{x\}$ *are, respectively, the smallest integer* $\geqq x$ *and the smallest integer* $> x$.

PROOF. The line $x + y = m + n - r$ intersects the diagonal at $y_0 = m - mr/(m+n)$. Therefore, the required paths are those for which $R_k - k \geqq nk/m$, $1 \leqq k \leqq [y_0]$; $R_k - k > nk/m$, $[y_0] < k < m$; $R_m - m = r$. But $[y_0] = m - \langle mr/(m+n) \rangle$ and (3. 2) follows by Lemma 1 with $c_i - i = n$, all i; $b_j - j = \langle nj/m \rangle$, $j \leqq [y_0]$; $b_j - j = \{nj/m\}$, $[y_0] < j < m$; and $b_m - m = n$.

Although the second formula can be derived from the first, the process is more trouble than it is worth. It is easier to derive it directly by counting.

$$M_{mn}(r) = N_{mn} - (\text{number of paths never above the diagonal that touch it after the } (m+n-r)^{\text{th}} \text{ step})$$

$$= N_{mn} - \sum_{\substack{(i,j) \text{ on diagonal} \\ i+j > m+n-r \\ i < n,\ j < m}} (\text{number of paths never above the diagonal that touch it for the last time at } (i, j))$$

$$= N_{mn} - \binom{m+n}{n} P(mnD^+_{mn} = 0, R^+_{mn} < r).$$

The last equality follows because when $mnD^+_{mn} = 0$ the paths which are never above the diagonal and last touch the diagonal at step $m+n-i$ are easily seen (by interchanging the parts of the path before and after the last touching) to be in 1—1

correspondence with the paths which are never above the diagonal and first touch it at step i. ■

Since m and n relatively prime implies $\langle nk/m\rangle = \{nk/m\}$ for $k = 1, 2, \ldots, m-1$, it follows in this case that $M_{mn}(r)$ is a constant independent of r. Therefore, the values of R^+_{mn} are equally likely and we have the following corollary to Lemma 4b.

COROLLARY. If $\gcd(m, n)=1$, $M_{mn}(r) = \binom{m+n}{n}/(m+n)$.

Furthermore, when $D^+_{mn}=0$, the possible values of R^+_{mn} are $q, 2q, \ldots, pq \equiv m+n$. Thus, as r increases, $M_{mn}(r)$ changes only when $r \equiv 1 \pmod q$. Therefore, $M_{mn}(hq+i) = M_{mn}(hq+1)$ for $i=1, 2, \ldots, q$ and $h = 0, 1, \ldots, p-1$. Since $P(mnD^+_{mn}=0, R^+_{mn}<r)$ is non-decreasing in r it follows that $M_{mn}(r)$ is non-increasing in r. This proves Lemma 5.

LEMMA 5. (i) $M_{mn}(r+1) \geqq M_{mn}(r)$, $r = 1, 2, \ldots, m+n-1$,
(ii) *the* $M_{mn}(r)$ *are equal in blocks of length* q, *i.e.*,

$$M_{mn}(hq+1) = M_{mn}(hq+2) = \cdots = M_{mn}(hq+q) \qquad h = 0, 1, \ldots, p-1.$$

See also Lemma 10 which evaluates $M_{mn}(r)$ in the case $n=sm$.

4. Computation of $P(mnD^+_{mn}=d, R^+_{mn}=r)$. A path from $(0, 0)$ to (n, m) through (x, y), $x \leqq ny/m$, that attains its maximum distance from the diagonal for the first time at (x, y) corresponds to a path for which $mnD^+_{mn} = ny-mx$, $R^+_{mn} = x+y$. By Lemma 2b the number of such paths is the same as the number of paths from $(0, 0)$ to $(n-x, m-y)$ that are never above the line $y=mx/n$ times the number of paths from $(n-x, m-y)$ to (n, m) that never touch the line $y=mx/n$. Call these numbers $N_1(x, y)$ and $N_2(x, y)$, respectively. $N_1(x, y)$ is given by Lemma 1 with $c_i - i = n-x$ and $b_i - i = \langle n_i/m\rangle$, $i = 1, 2, \ldots, m-y$ for sample sizes $m' = m-y$ and $n' = n-x$. Hence

$$N_1(x, y) =$$

$$= \begin{vmatrix} \binom{n-x-\left\langle \frac{n}{m}\right\rangle+1}{1} & \binom{n-x-\left\langle \frac{2n}{m}\right\rangle+1}{2} & \cdots & \binom{n-x-\left\langle \frac{(m-y)n}{m}\right\rangle+1}{m-y} \\ 1 & \binom{n-x-\left\langle \frac{2n}{m}\right\rangle+1}{1} & \cdots & \binom{n-x-\left\langle \frac{(m-y)n}{m}\right\rangle+1}{m-y-1} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \binom{n-x-\left\langle \frac{(m-y)n}{m}\right\rangle+1}{1} \end{vmatrix}_{m-y \times m-y}. \tag{4.1}$$

Since the point $(n-x, m-y)$ is y units below and x units to the left of (n, m), we can take (n, m) as a new origin and consider $N_2(x, y)$ as the number of paths from $(0, 0)$ to (x, y) that are above the line $y=mx/n$. This number is given by Lemma 1

with sample sizes $m'=y$, $n'=x$, $b_i-i=0$, all i, $c_1=1$, and $c_i-i+1 \equiv z_{i+1} = \min(x+1, \langle ni/m \rangle)$, $i=2, 3, \ldots, y$. Hence, since the first row of the lemma determinant is $(1, 0, 0, \ldots, 0)$, we can write

$$N_2(x, y) = \begin{vmatrix} \binom{\langle n/m \rangle}{1} & \binom{\langle n/m \rangle}{2} & \binom{\langle n/m \rangle}{3} & \cdots & \binom{\langle n/m \rangle}{y-1} \\ 1 & \binom{\langle 2n/m \rangle}{1} & \binom{\langle 2n/m \rangle}{2} & \cdots & \binom{\langle 2n/m \rangle}{y-2} \\ 0 & 1 & \binom{\langle 3n/m \rangle}{1} & \cdots & \binom{\langle 3n/m \rangle}{y-3} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \binom{z_y}{1} \end{vmatrix}_{y-1 \times y-1}. \tag{4.2}$$

Finally, we have

LEMMA 6.

$$\binom{m+n}{m} P(mnD^+_{mn} = d, R^+_{mn} = r) = \begin{cases} N_1(x, y) \cdot N_2(x, y) \\ 0 \end{cases} \tag{4.3}$$

according as there exists an integer solution to the equations $ny-mx=d$, $x+y=r$ *such that* $0 \leqq x \leqq n$, $0 \leqq y \leqq m$ *or not.*

If $n=sm$, s an integer, so that gcd $(m, n)=m$, the determinants can be evaluated explicitly and we have

LEMMA 7. *If* $n=sm$ *then*

$$N_1(x, y) = \begin{vmatrix} \binom{n-x-s+1}{1} & \binom{n-x-2s+1}{2} & \cdots & \binom{n-x-(m-y)s+1}{m-y} \\ 1 & \binom{n-x-2s+1}{1} & \cdots & \binom{n-x-(m-y)s+1}{m-y-1} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \binom{n-x-(m+y)s+1}{1} \end{vmatrix}_{m-y \times m-y} \tag{4.4}$$

$$= \frac{sy-x+1}{m+n-x-y+1} \binom{m+n-x-y+1}{m-y}.$$

PROOF. This result is Lemma 2.1 of STECK (1974). ∎

LEMMA 8. *Let* $n=sm$. *If* $x+1 \geqq (y-1)s$ *let* $v=y$; *otherwise let* v *and* a *be defined by* $x+1=vs-a$. *Then*

$$N_2(x, y) = \sum_{j=0}^{y-v} (-1)^j \frac{s}{(y-j)(s+1)-1} \binom{(y-j)(s+1)-1}{y-j-1} \binom{(y-j)s-x-1}{j}. \tag{4.5}$$

PROOF. If $x+1 \geqq (y-1)s$, then $N_2(x, y)$ is the determinant of a matrix which is of the same form as the transpose of the matrix whose determinant is $N_1(x, y)$. In this simple case $N_2(x, y)$ becomes the $j=0$ term of (4.5). The only difficulty comes when $x+1 < (y-1)s$. Suppose

$$N_2(x,y) = \begin{vmatrix} \binom{s}{1} & \binom{s}{2} & \binom{s}{3} & \dots & \binom{s}{y-2} & \binom{s}{y-1} \\ 1 & \binom{2s}{1} & \binom{2s}{2} & \dots & \binom{2s}{y-3} & \binom{2s}{y-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \binom{(v-1)s}{y-v} & \binom{(v-1)s}{y-v+1} \\ 0 & 0 & 0 & \dots & \binom{x+1}{y-v-1} & \binom{x+1}{y-v} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & \binom{x+1}{1} \end{vmatrix}_{y-1 \times y-1} \begin{matrix} \left.\vphantom{\begin{matrix}a\\b\\c\\d\end{matrix}}\right\} v-1 \text{ rows} \\ \\ \\ \\ \left.\vphantom{\begin{matrix}a\\b\\c\\d\end{matrix}}\right\} y-v \text{ rows} \end{matrix}.$$

Replacing the vth row (the first one with $x+1'$ s) by $\sum_{j=v}^{y-1} \binom{a}{j-v} \cdot (\text{row } j)$ makes the vth row equal to $\left(0, 0, \dots, 0, 1, \binom{vs}{1}, \dots, \binom{vs}{y-v-1}, \binom{vs}{y-v} - \binom{a}{y-v}\right)$ by the Vandermonde convolution formula. Changing the rest of the rows similarly gives $N_2(x, y)$ as the difference of two determinants. One corresponds to $x+1 = (y-1)s$ which has already been evaluated and the second can be expanded by the last row into a linear combination of such determinants. ■

According to Lemma 4b, we need the results of Lemma 7 and 8 for the case corresponding to $D^+_{mn}=0$. This implies $ny=mx$ and $x+y=r$ so that $x = rn/(m+n)$, $y = rm/(m+n)$. If, in addition, $n=sm$ then $x = rs/(s+1)$, $y = r/(s+1)$. For x and y to be integers r must be a multiple of $s+1$, say $h(s+1)$. Then $x=hs$, $y=h$ and $x=sy$. Thus the simplest form $(v=y)$ of (4. 5) applies and we have

LEMMA 9. *If* $n=sm$ *and* $r = h(s+1)$

$$\binom{m+n}{m} P(D^+_{mn} = 0, R^+_{mn} = r) = \frac{1}{m+n-r+1} \binom{m+n-r+1}{m-h} \frac{s}{r-1} \binom{r-1}{h-1}.$$

Finally, in the case $n=sm$, we have

$$N_{n,n} = N_1(0, 1) = \binom{m+n}{m-1} \Big/ m = \binom{m+n}{m} \Big/ (n+1),$$

giving

LEMMA 10.

$$M_{m,sm}(r) = \binom{m+n}{m} \Big/ (n+1) - \sum_{h=1}^{\{r/(s+1)\}} \frac{s}{(h(s+1)-1)(m+n-h(s+1)+1)} \times \times \binom{h(s+1)-1}{h-1} \binom{m+n-h(s+1)+1}{m-h}. \tag{4.6}$$

5. Examples. A. $m=n=8$. $s=1$, $\binom{m+n}{n}/(n+1) = 1430$. The summands in (4.6) are 429, 132, 84, 70, 70, 84, 132 for $h=1, 2, \ldots, 7$, respectively.

Therefore,

r	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$M_{8,8}(r)$	1430	1430	1001	1001	869	869	785	785	715	715	645	645	561	561	429	429

B. $m=4$, $n=12$. $s=3$, $q=4$ and $\binom{m+n}{n}/(n+1) = 140$. The summands in (4. 6) are 22, 12, 15 for $h=1, 2, 3$, respectively. Hence

$$M_{4,12}(r) = \begin{cases} 140 & r = 1, 2, 3, 4 \\ 118 & r = 5, 6, 7, 8 \\ 106 & r = 9, 10, 11, 12 \\ 91 & r = 13, 14, 15, 16. \end{cases}$$

C. $m=6$, $n=9$. Here $\gcd(m, n)=3$ and $q=15/3=5$. From (3.2) we find

$$M_{6,9}(r) = \begin{vmatrix} \binom{8}{1} & \binom{7}{2} & \binom{5}{3} & \binom{4}{4} & 0 \\ 1 & \binom{7}{1} & \binom{5}{2} & \binom{4}{3} & 0 \\ 0 & 1 & \binom{5}{1} & \binom{4}{2} & 0 \\ 0 & 0 & 1 & \binom{4}{1} & \binom{2}{2} \\ 0 & 0 & 0 & 1 & \binom{2}{1} \end{vmatrix} = 377, \quad r = 1, 2, 3, 4, 5$$

$$M_{6,9}(r) = \begin{vmatrix} \binom{8}{1} & \binom{7}{2} & \binom{5}{3} & 0 & 0 \\ 1 & \binom{7}{1} & \binom{5}{2} & \binom{3}{3} & 0 \\ 0 & 1 & \binom{5}{1} & \binom{3}{2} & 0 \\ 0 & 0 & 1 & \binom{3}{1} & \binom{2}{2} \\ 0 & 0 & 0 & 1 & \binom{2}{1} \end{vmatrix} = 331, \quad r = 6, 7, 8, 9, 10$$

$$M_{6,9}(r) = \begin{vmatrix} \binom{8}{1} & \binom{6}{2} & \binom{5}{3} & 0 & 0 \\ 1 & \binom{6}{1} & \binom{5}{2} & \binom{3}{3} & 0 \\ 0 & 1 & \binom{5}{1} & \binom{3}{2} & 0 \\ 0 & 0 & 1 & \binom{3}{1} & \binom{2}{2} \\ 0 & 0 & 0 & 1 & \binom{2}{1} \end{vmatrix} = 293, \quad r = 11, 12, 13, 14, 15.$$

D. $m=9$, $n=6$. Again $\gcd(m, n)=3$ and $q=5$, but now from (3.2) we have

$$M_{9,6}(r) = \begin{vmatrix} \binom{6}{1} & \binom{5}{2} & \binom{5}{3} & \binom{4}{4} & 0 & 0 & 0 & 0 \\ 1 & \binom{5}{1} & \binom{5}{2} & \binom{4}{3} & 0 & 0 & 0 & 0 \\ 0 & 1 & \binom{5}{1} & \binom{4}{2} & \binom{3}{3} & 0 & 0 & 0 \\ 0 & 0 & 1 & \binom{4}{1} & \binom{3}{2} & \binom{3}{3} & 0 & 0 \\ 0 & 0 & 0 & 1 & \binom{3}{1} & \binom{3}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \binom{3}{1} & \binom{2}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \binom{2}{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{vmatrix} = 377, \quad r = 1, 2, 3, 4, 5$$

$$M_{9,6}(r) = \begin{vmatrix} \binom{6}{1} & \binom{5}{2} & \binom{5}{3} & \binom{4}{4} & 0 & 0 & 0 & 0 \\ 1 & \binom{5}{1} & \binom{5}{2} & \binom{4}{3} & 0 & 0 & 0 & 0 \\ 0 & 1 & \binom{5}{1} & \binom{4}{2} & \binom{3}{3} & 0 & 0 & 0 \\ 0 & 0 & 1 & \binom{4}{1} & \binom{3}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \binom{3}{1} & \binom{2}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \binom{2}{1} & \binom{2}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \binom{2}{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{vmatrix} = 331, \quad r = 6, 7, 8, 9, 10$$

$$M_{9,6}(r) = \begin{vmatrix} \binom{6}{1} & \binom{5}{2} & \binom{4}{3} & \binom{4}{4} & 0 & 0 & 0 & 0 \\ 1 & \binom{5}{1} & \binom{4}{2} & \binom{4}{3} & 0 & 0 & 0 & 0 \\ 0 & 1 & \binom{4}{1} & \binom{4}{2} & \binom{3}{3} & 0 & 0 & 0 \\ 0 & 0 & 1 & \binom{4}{1} & \binom{3}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \binom{3}{1} & \binom{2}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \binom{2}{1} & \binom{2}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \binom{2}{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{vmatrix} = 293, \quad r = 11, 12, 13, 14, 15.$$

E. $m = 7$, $n = 8$. Since $\binom{15}{7} \Big/ 15 = 6435/15 = 429$ we have

$$M_{7,8}(r) = 429, \quad r = 1, 2, \ldots, 15.$$

REFERENCES

[1] GELLER, N. L.: Two limiting distributions for two-sample Kolmogorov—Smirnov type statistics. Report No. 5, Center for System Science, Case-Western Reserve University, Cleveland, Ohio, (1971).

[2] SARKADI, K.: On Galton's rank order test. *Publ. Math. Inst. Hungar. Acad. Sci* **6** (1961), 125—128.

[3] STECK, G. P.: The Smirnov two-sample tests as rank tests. *Ann. Math. Statist.* **40** (1969), 1449—1466.

[4] STECK, G. P.: Evaluation of some Steck determinants with applications. *Communications Statist.* **3** (1974), to appear.

[5] VINCZE, I.: Einige zweidimensionale Verteilungs und Grenzverteilungssätze in der Theorie der geordneten Stichproben. *Publ. Math. Ist. Hungar, Acad. Sci.* **2**, (1957), 183—209.

Sandia Laboratories, Albuquerque, New Mexico 87115

(Received January 12, 1972.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 91—93.

ON A CERTAIN AFFINE INVARIANT FUNCTIONAL FOR CONVEX BODIES

by

G. D. CHAKERIAN

1. Introduction. In a recent paper [3] L. FEJES TÓTH introduced the following affine invariant functional $M(K)$, defined for convex bodies K in Euclidean n-space. For each direction u let $F(u)$ be a cross-section of K of maximal area by a hyperplane orthogonal to u, and let $df(u)$ be the surface area element at a boundary point of K with outward normal vector u. Then

$$M(K) = \int \frac{df(u)}{F(u)}, \tag{1}$$

where the integration is over the boundary of K. It is shown in [3] how $M(K)$ arises in connection with the maximum number of non-overlapping translates of λK that can be brought into contact with K for small positive values of λ. Bounds on $M(K)$ are of interest. For example, as shown in [3], any n-dimensional centrally symmetric convex body K satisfies $M(K) \leqq 2n$, with equality holding only if K is a parallelotope.

The purpose of this note is to point out the relationship of $M(K)$ with some other functionals and exhibit bounds on $M(K)$ in special cases.

2. Planar convex bodies. A centrally symmetric plane convex body C centered at the origin gives rise to a norm, making the plane a 2-dimensional Banach space for which C is the unit disk. The *self-circumference* $\sigma(C)$ of C is the length of the boundary of C as measured in this Banach space, and it is readily seen that $\sigma(C) = 2M(C)$. It is proved in [6] that $6 \leqq \sigma(C) \leqq 8$, and equality holds on the lefthand side only if C is the affine image of a regular hexagon and on the righthand side only if C is a parallelogram (see also [2] and the references cited there). The *central symmetrization* K^* of a convex body K is defined by $K^* = \frac{1}{2}(K+(-K))$. As shown in [3], $M(K^*) = M(K)$ if K is planar. From these remarks one obtains the bounds

$$3 \leqq M(K) \leqq 4. \tag{2}$$

valid for any plane convex body K, with equality on the lefthand side only if $K+(-K)$ is an affine regular hexagon and on the righthand side only if $K+(-K)$, and hence K itself, is a parallelogram. The upper bound in (2), with the characterization of those sets for which equality holds, was established in [3] directly, while the lower bound verifies a conjecture made there. It is not difficult to prove that $K+(-K)$ is an affine regular hexagon if and only if K is a truncated triangle as described in [3].

3. Higher dimensional bodies. A functional $\tau(K)$ was defined in [2] as follows. For each $p \in K$ and each direction u let $G(p, u)$ be the area of the cross-section of K by the hyperplane through p orthogonal to u. Let

$$\tau(K, p) = \int \frac{df(u)}{G(p, u)}, \tag{3}$$

where the integration is over the boundary of K. Then $\tau(K)$ is the minimum of $\tau(K, p)$ as p ranges over the interior of K. Since for each $p \in K$ and each direction u one has $G(p, u) \leqq P(u)$, it follows that

$$M(K) \leqq \tau(K). \tag{4}$$

Thus the bounds derived for $\tau(K)$ in [2] imply, in particular, the crude upper bound

$$M(K) \leqq n(n+1), \tag{5}$$

valid for any n-dimensional convex body K. One obtains (5) directly by observing that if $b(u)$ is the distance between parallel supporting hyperplanes orthogonal to direction u, then Cavalieri's principle implies $V \leqq b(u)F(u)$, where V is the volume of K. Hence

$$M(K) \leqq \frac{1}{V}\int b(u)\,df(u), \tag{6}$$

as derived in [3]. It is shown in [2] how properties of mixed volumes imply that the righthand side of (6) is not greater than $n(n+1)$.

The bound in (5) seems of too high an order of magnitude in n. Computation of $M(K)$ for various special bodies leads one to ask the following question (analogous to the conjecture 6. 7 of [2]). *Do there exist positive real numbers A and B, independent of n, such that for every n-dimensional convex body K,* $A\sqrt{n} < M(K) < Bn$?

If C is a centrally symmetric convex body in n-space, then the maximum area $F(u)$ in (1) is obtained by a cross-section through the center of C. In this case $M(C)$ becomes the *Minkowskian self-surface area* of C. Bounds for $M(C)$ are studied in [1], but the following question still remains unanswered for $n \geqq 3$. *What is the infimum of* $M(C)$ *as C ranges over all centrally symmetric convex bodies in n-space?* As proved in [1] and [3], the supremum is $2n$, achieved for parallelotopes.

4. A functional related to $M(K)$. Let $P(u)$ be the *brightness function* of a convex body K, that is, $P(u)$ is the area of the orthogonal projection of K onto a hyperplane normal to the direction u. Define

$$\lambda(K) = \int \frac{df(u)}{P(u)}, \tag{7}$$

integrated over the boundary of K. Like $M(K)$ the functional $\lambda(K)$ is affine invariant. Since $P(u) \geqq F(u)$ for all u, one has

$$\lambda(K) \leqq M(K). \tag{8}$$

If equality holds in (8), then $P(u) = F(u)$ for all u, implying that each cylinder circumacribed about K touches K in a set containing a hyperplane cross-section of the boundary. In other words, each *shadow boundary* of K contains a cross-section of

the boundary by a hyperplane, a property that characterizes ellipsoids in n-space if $n \geqq 3$ (see [7]). But a ball is the only ellipsoid satisfying $P(u) \equiv F(u)$; hence in n-space, $n \geqq 3$, equality holds in (8) if and only if K is a ball. For planar convex bodies, equality holds in (8) precisely for sets of constant width.

By means of the operation of *Blaschke addition* (see [4]) one can associate with any convex body K a centrally symmetric convex body $\tilde{K}$ whose surface area element $d\tilde{f}(u)$ satisfies $2d\tilde{f}(u) = df(u) + df(-u)$, where $df(u)$ is the surface area element of K. It follows that the area of the orthogonal projection of $\tilde{K}$ on any hyperplane is the same as that of K, and

$$\lambda(\tilde{K}) = \lambda(K). \tag{9}$$

Hence in obtaining bounds on $\lambda(K)$ one need consider only centrally symmetric bodies.

K is a body of *constant brightness* if $P(u)$ has the same value for all u. Such K are characterized by the property that $\tilde{K}$ is a ball (see [4]). Non-spherical bodies of constant brightness of all dimensions are exhibited in [5]. If B is an n-dimensional body of constant brightness one has from (9) that $\lambda(B) = \lambda(n\text{-ball})$. Thus, from (8), $M(B) \geqq \lambda(n\text{-ball})$, with equality (if $n \geqq 3$) only when B is already a ball, or (if $n = 2$) when B is a set of constant width. In particular, any 3-dimensional body B of constant brightness satisfies $M(B) \geqq 4$, with equality only for a ball.

Any 3-dimensional simplex T also satisfies $M(T) = 4$. However, this is not the minimum possible value for 3-dimensional convex bodies, as shown in [3] with an example of a convex body of revolution R satisfying $M(R) < 4$. That example also shows that the Minkowskian self-surface area of 3-dimensional centrally symmetric bodies is not minimized by ellipsoids.

REFERENCES

[1] BUSEMANN, H. and PETTY, C. M.: Problems on convex bodies, *Math. Scand.* **4** (1956), 88—94.

[2] CHAKERIAN, G. D. and TALLEY, W. K.: Some properties of the self-circumference of convex sets, *Archiv der Math.* **20** (1969), 431—443.

[3] FEJES TÓTH, L.: Über eine Affininvariante Masszahl bei Eipolyedern, *Stud. Sci. Math. Hungar.* **5** (1970), 173—180.

[4] FIREY, W. J.: *Blaschke sums of convex bodies and mixed bodies,* Proc. Colloq. on Convexity, Copenhagen, 1965, 94-101.

[5] FIREY, W. J.: Convex bodies of constant outer p-measure, *Mathematika* **17** (1970), 21—27.

[6] GOLĄB, S.: Quelques problèmes mètriques de la gèomètrie de Minkowski, *Travaux de l'Academie des Mines* à *Cracovie* **6** (1932). (Polish, French summary.)

[7] MARCHAUD, A.: Un théorème sur les corps convexes, *Ann. Sci, École Norm. Sup.* (3) **76** (1959) 283—304.

University of California, Davis, Calif 95616

(Received January 21, 1972)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 95—102.

A LAW OF THE ITERATED LOGARITHM FOR THE ROBBINS—MONRO METHOD

by

P. MAJOR

Let a real number α and a function $H(x, y)$ be given which is a distribution function for fixed x and measurable for fixed y. Set

$$M(x) = \int_{-\infty}^{\infty} yH(dy, x).$$

Let $a_n \geqq 0$ be a sequence of real numbers for which $\Sigma a_n = \infty$, $\Sigma a_n^2 < \infty$. Let us define the random variables $x_1, x_2, \ldots y_1, y_2, \ldots$ with the following properties: x_1 is an arbitrary constant $x_{n+1} = x_n - a_n(y_n - \alpha)$, and

$$\mathsf{P}(y_n < y | x_1, y_1 \ldots y_{n-1}, x_n) = \mathsf{P}(y_n < y | x_n) = H(y, x_n).$$

This construction of the x_i's and y_i's is the Robbins—Monro method [1]. J. R. BLUM proved in [2] that if $M(x)$ and the root θ of the equation $M(x) = \alpha$ satisfy the following conditions:

(I) $|M(x)| \leqq A|x| + B$ for all x and suitable A, B;

(II) $\inf_{\varepsilon < x - \theta < \frac{1}{\varepsilon}} M(x) > \alpha$ and $\sup_{-\varepsilon > x - \theta > -\frac{1}{\varepsilon}} M(x) < \alpha$ for every $\varepsilon > 0$;

(III) $\int_{-\infty}^{\infty} (y - M(x))^2 H(dy, x) \leqq K < \infty$ for every x

then $x_n \to \theta$ with probability 1.

We shall deal with the special case, when the errors are bounded, that is $H(M(x) - K, x) = 0$, $H(M(x) + K, x) = 1$ for every x, and $a_n = \frac{1}{n}$ are chosen. In this case, using mainly the idea of J. R. BLUM [2] we can strengthen his result in the following way.

THEOREM 1. *Let us suppose that* $M'(\theta) > \frac{1}{2}$, $a_n = \frac{1}{n}$, $M(x)$ *is bounded in a neighbourhood of* θ *and* $x_n \to \theta$ *with probability* 1. *Then there exists a number* L *(depending on* $M'(\theta)$ *and the bound* K *of the error) such that*

$$\mathsf{P}\left(\overline{\lim} \sqrt{\frac{n}{\log \log n}} |x_n - \theta| < L\right) = 1.$$

On the other hand, if we suppose that the errors are identically distributed, that is the distribution-function $H(y-M(x), x)$ does not depend on x then we can state the following

THEOREM 2. *If the errors are bounded having the same distribution and* $M'(\theta)<\infty$, *then there exists an* $L'>0$ *such that*

$$\mathsf{P}\left(\overline{\lim}\sqrt{\frac{n}{\log\log n}}|x_n-\theta|>L'\right)=1.$$

For the sake of simplicity we shall suppose that $\alpha=\theta=0$.

It is easy to see that for $m>n$

$$x_m=x_n-\sum_{i=n}^{m-1}\frac{1}{i}y_i=x_n-\sum_{i=n}^{m-1}\frac{1}{i}M(x_i)-\sum_{i=n}^{m-1}\frac{1}{i}(y_i-M(x_i)). \tag{1}$$

Let $F_i=B(x_1, y_1, \ldots, x_i, y_i)$ be the smallest σ-algebra with respect to which the random variables $x_1, y_1, \ldots, x_i, y_i$ are measurable. Then the system

$$\left(\frac{1}{i}(y_i-M(x_i)), F_i\right)$$

is a martingale difference series. If the errors are identically distributed then the random variables $\frac{1}{i}(y_i-M(x_i))$ are even independent. Using these facts we give a sharper bound than that of BLUM and these boundings will enable us to prove our statement. This way we need the following

LEMMA 1. *Let* $(\xi_1, F_1)\ldots(\xi_n, F_n)\ldots$ *be martingale difference series with* $\mathsf{P}(|\xi_i|\leqq K)=1$. *Then for almost all* ω *and for all* $\varepsilon>0$ *there exists an* $n(\omega)=n(\omega, \varepsilon)$ *such that for every* $m>n\geqq n(\omega)$

$$\left|\sum_{i=n}^{m}\frac{1}{i}\xi_i\right|\leqq 2\sqrt{2}K+\varepsilon)\sqrt{\frac{\log\log n}{n}}.$$

PROOF.

$$\mathsf{P}\left(\sup_{N\leqq m\leqq n}\sum_{i=n}^{m}\frac{\xi_i}{i}>u\sqrt{\frac{\log\log n}{n}}\right)=$$

$$=\mathsf{P}\left(\sup_{N\leqq m\leqq n}\exp f(n)\sum_{i=n}^{m}\frac{\xi_i}{i}>\exp u\cdot f(n)\sqrt{\frac{\log\log n}{n}}\right)$$

where $f(n)\geqq 0$ is arbitrary constant.

But $$\left\{\exp\sum_{i=n}^{m}\frac{f(n)}{i}\xi_i, F_m\right\}\quad m=n, n+1, \ldots N$$

is a semimartingale and by the Kolmogorov inequality for semimartingale (see [4] Chapter VII) we get that

$$\mathsf{P}\left(\sup_{N\geqq m\geqq n}\sum_{i=n}^{m}\frac{\xi_i}{i}\geqq u\sqrt{\frac{\log\log n}{n}}\right)\leqq$$

$$\leqq\frac{\mathsf{E}\left\{\exp\left(f(n)\sum_{i=1}^{N}\frac{\xi_i}{i}\right)\right\}}{\exp\left(u\cdot f(n)\sqrt{\frac{\log\log n}{n}}\right)}.$$

For $|x|<\delta(\varepsilon)$ we have $e^x<1+x+\frac{x^2}{2-\varepsilon}$, and choosing $f(n)=v\cdot\sqrt{n\log\log n}$ we get that if n is big enough and $i\geqq n$, then $\left|\frac{f(n)\xi_i}{i}\right|\leqq\delta(\varepsilon)$.

Hence

$$\mathsf{E}\exp\sum_{i=n}^{N}\frac{f(n)\xi_i}{i}\leqq\mathsf{E}\left[\left(\exp\sum_{i=n}^{N-1}\frac{f(n)\xi_i}{i}\right)\cdot\left(1+\frac{f(n)}{N}\xi_N+\frac{f^2(n)\xi_N^2}{(2-\varepsilon)N^2}\right)\right]\leqq$$

$$\leqq\left(1+\frac{f^2(n)K^2}{(2-\varepsilon)N^2}\right)\cdot\mathsf{E}\left(\exp\sum_{i=n}^{N-1}\frac{f(n)\xi_i}{i}\right)\leqq\cdots$$

$$\leqq\prod_{i=n}^{N}\left(1+\frac{f^2(n)K^2}{(2-\varepsilon)i^2}\right)\leqq\exp\frac{f(n)^2K^2}{(2-\varepsilon)(n-1)}.$$

So for large n

$$\mathsf{P}\left(\sup_{N\geqq m\geqq n}\sum_{i=n}^{m}\frac{\xi_i}{i}\geqq u\sqrt{\frac{\log\log n}{n}}\right)\leqq\exp\left[\frac{f^2(n)K^2}{(2-\varepsilon)(n-1)}-u\cdot f(n)\sqrt{\frac{\log\log n}{n}}\right].$$

Tending with N to infinity, for large n

$$\mathsf{P}\left(\sup_{m\geqq n}\sum_{i=n}^{m}\frac{\xi_i}{i}\geqq u\sqrt{\frac{\log\log n}{n}}\right)\leqq\exp\left\{\left[v^2\left(1+\frac{1}{n-1}\right)\frac{K^2}{2-\varepsilon}-uv\right]\log\log n\right\}\leqq$$

$$\leqq\frac{1}{(\log n)^{1+\varepsilon}}\quad\text{if we choose}\quad u=\frac{2(1+\varepsilon)}{\sqrt{2-\varepsilon}}K,\ v=\frac{\sqrt{2-\varepsilon}}{K}.$$

Using the Borel—Cantelli lemma for arbitrary $\theta>1$

$$\sum_{n=1}^{\infty}\mathsf{P}\left(\sup_{m\geqq\theta^n}\sum_{i=\theta^n}^{m}\frac{\xi_i}{i}\geqq 2\frac{(1+\varepsilon)}{\sqrt{2-\varepsilon}}K\sqrt{\frac{\log\log\theta^n}{\theta^n}}\right)\leqq C+\frac{1}{\log\theta}\sum_{n=1}^{\infty}\frac{1}{n^{1+\varepsilon}}<\infty$$

so for almost all ω there is an $n(\omega)$ such that for every

$$m\geqq\theta^k\geqq n(\omega)\sum_{i=\theta^k}^{m}\frac{\xi_i}{i}\leqq(\sqrt{2}K+\varepsilon)\sqrt{\frac{\log\log\theta^k}{\theta^k}}.$$

In the same way

$$\sum_{i=\theta^k}^{m} \frac{\xi_i}{i} \geqq -(\sqrt{2}K+\varepsilon)\sqrt{\frac{\log\log\theta^k}{\theta^k}}.$$

Let $m>n>n(\omega)$ and $n(\omega)\leqq\theta^k<\theta^{k+1}$. Then

$$\left|\sum_{i=n}^{m}\frac{\xi_i}{i}\right| \leqq \left|\sum_{i=\theta^k}^{n}\frac{\xi_i}{i}\right| + \left|\sum_{i=\theta^k}^{m}\frac{\xi_i}{i}\right| \leqq$$

$$\leqq 2(\sqrt{2}K+\varepsilon)\sqrt{\frac{\log\log\theta^k}{\theta^k}} \leqq \frac{2(\sqrt{2}K+\varepsilon)}{\theta}\sqrt{\frac{\log\log n}{n}}.$$

Since ε and θ can be chosen as near to 0 and to 1, respectively as we want, our lemma is proved.

PROOF of Theorem 1. First we prove that $\mathsf{P}\left(\overline{\lim}\sqrt{\frac{n}{\log\log n}}\,|x_n|<\infty\right)=1$. We shall prove this relation for every ω for which $y_i(\omega)-M(x_i(\omega))$ satisfies the lemma and $x_n(\omega)\to 0$ as $n\to\infty$. Let us choose an $\varepsilon>0$ in such a way that

$$|M(y)| > \beta|y| \quad \text{and} \quad |M(y)| < a \quad \text{if} \quad |y|<\varepsilon \tag{2}$$

where $\beta>\frac{1}{2}$, $a>0$ are constants. Let us choose a number c $0<c<1$ such that $2\beta c>1$. Then it is easy to see that

$$\lim_{\delta\to 0}\left[\frac{\beta\sqrt{1+\delta}\log(1+c\delta)}{\sqrt{1+\delta}-1}-\sqrt{1+\delta}\right]=2\beta c-1>0.$$

Let us choose a $\delta>0$ so, that

$$\frac{\beta\sqrt{1+\delta}\log(1+c\delta)}{\sqrt{1+\delta}-1}-\sqrt{1+\delta}=u>0. \tag{3}$$

Let us choose an $n_0=n_0(\omega)$ so large, that

$$|x_n(\omega)|<\varepsilon \quad \text{for} \quad n\geqq n_0 \tag{4}$$

$$\left|\sum_{i=n}^{m}\frac{1}{i}\,(y_i-M(x_i))\right| \leqq 3K\sqrt{\frac{\log\log n}{n}} \quad \text{for} \quad m\geqq n\geqq n_0 \tag{5}$$

$$\frac{K+a}{\sqrt{n_0}} \leqq 3K\frac{1+\delta}{(\sqrt{1+\delta}-1)n}. \tag{6}$$

Let us define

$$L(\omega)=\max\left\{3K\sqrt{1+\delta}\,\frac{1+\delta+u(\sqrt{1+\delta}-1)}{u(\sqrt{1+\delta}-1)},\ \max_{n_0\leqq n\leqq n_0(1+\delta)}\sqrt{\frac{n}{\log\log n}}\,|x_n|\right\}.$$

We state that

$$|x_n| \leqq L(\omega)\sqrt{\frac{\log\log n}{n}} \quad \text{for every} \quad n \geqq n_0. \tag{7}$$

For $n_0 \leqq n \leqq n_0(1+\delta)$ (7) follows from the definition of $L(\omega)$. As a first step we prove that the validity of (7) for an $n \geqq n_0$ implies its validity for every $n(1+c\delta) \leqq \leqq \tilde{n} \leqq (1+\delta)n$ .Our statement is a straight consequence of this fact.

In order to prove this we introduce the following notation:

$$L^* = \frac{u(\sqrt{1+\delta}-1)}{\sqrt{1+\delta}[1+\delta+u(\sqrt{1+\delta}-1)]} L(\omega) \quad \text{and} \quad d = L^* \frac{1+\delta}{(\sqrt{1+\delta}-1)u}.$$

Obviously $L^* \geqq 3K$ and $d \leqq L(\omega)$. Let us consider two different cases:

First case: $x_j > d\sqrt{\frac{\log\log n}{n}}$ for every $n \leqq j < \tilde{n}$.

Since (7) is true for n by (1), (2) and (5) in this case we have

$$x_{\tilde{n}} = x_n + \sum_{j=n}^{\tilde{n}-1} \frac{1}{j}[y_j - M(x_j)] - \sum_{j=n}^{\tilde{n}-1} \frac{1}{j} M(x_j) \leqq$$

$$\leqq L(\omega)\sqrt{\frac{\log\log n}{n}} + L^*\sqrt{\frac{\log\log n}{n}} - \beta d \sum_{j=n}^{\tilde{n}-1} \frac{1}{j}\sqrt{\frac{\log\log n}{n}}.$$

Thus by the definition of d and by (3) and finally using the definition of L^* we get

$$x_{\tilde{n}} \leqq \sqrt{\frac{\log\log n}{n}}[L(\omega) + L^* - \beta d \log(1+\delta c)] =$$

$$= \sqrt{\frac{\log\log n}{n}}\left(L(\omega) + L^* - L^*\sqrt{1+\delta} - L^* \frac{1+\delta}{n}\right) = \frac{L(\omega)}{\sqrt{1+\delta}}\sqrt{\frac{\log\log n}{n}} \leqq$$

$$\leqq L(\omega)\sqrt{\frac{\log\log \tilde{n}}{\tilde{n}}}.$$

Second case: There is a j, $n \leqq j < \tilde{n}$ for which $x_j < \sqrt{\frac{\log\log n}{n}}\, d$. In this case either $x_{\tilde{n}-1} < 0$ and

$$x_{\tilde{n}} \leqq \frac{K+a}{\tilde{n}-1} \leqq d\frac{1}{\sqrt{n-1}} \leqq L(\omega)\sqrt{\frac{\log\log \tilde{n}}{\tilde{n}}}$$

because of (6) or there is a $n \leqq j^* < \tilde{n}$ for which $0 \leqq x_{j^*} \leqq d\sqrt{\frac{\log\log n}{n}}$ and $x_i \geqq 0$ for every $j^* \leqq i < \tilde{n}$ also because of (6). So $M(x_i) \geqq 0$ for these i-s. Therefore

$$x_{\tilde{n}} \leqq x_{j^*} + \left|\sum_{i=j^*}^{\tilde{n}-1} \frac{1}{i}(y_i - M(x_i))\right| \leqq d\sqrt{\frac{\log\log n}{n}} + L^*\sqrt{\frac{\log\log n}{n}} =$$

$$= \sqrt{\frac{\log\log n}{n}} \frac{L(\omega)}{\sqrt{1+\delta}} \leqq L(\omega) \sqrt{\frac{\log\log \tilde{n}}{\tilde{n}}}.$$

In the same way it can be proved that $x_n > -L(\omega)\sqrt{\frac{\log\log n}{n}}$. By a little modification the original statement can be proved too.

Actually we proved the following: There is a constant L

$$\left(\text{explicitely } L = 3K\sqrt{1+\delta}\left(1+\frac{1+\delta}{u(\sqrt{1+\delta}-1)}\right)\right),$$

such that if $|x_n| \leqq K_n\sqrt{\frac{\log\log n}{n}}$ for an $n \geqq n_0(\omega)$ and an appropriate number K_n then $x_{\tilde{n}} \leqq \max(K_n, L)\sqrt{\frac{\log\log n(1+\delta)}{n(1+\delta)}}$ for $n(1+c\delta) < \tilde{n} \leqq n(1+\delta)$. But choosing a $c < \varrho < 1$ constant we get that $x_{\tilde{n}} \leqq \sqrt{\frac{1+\delta\varrho}{1+\delta}} \max(L, K_n)\sqrt{\frac{\log\log \tilde{n}}{\tilde{n}}}$ for $n(1+c\delta) \leqq \tilde{n} \leqq n(1+\delta\varrho)$; If the relation

$$x_n \leqq L(\omega)\sqrt{\frac{\log\log n}{n}}$$

were true only for $L(\omega) > L$ if $n_0 \leqq n < n_0(1+\delta)$ then the unequality $|x_n| \leqq \leqq \sqrt{\frac{1+\delta d}{1+\delta}} L(\omega)\sqrt{\frac{\log\log n}{n}}$ would be true for $n_0(1+\varrho\delta) \leqq n \leqq n_0(1+\varrho\delta)^2$. Iterating this result we get our statement.

By an affine transformation we get the following corollary.

COROLLARY. Let $a_n = \frac{c}{n}$, $2M'(\theta) > c$ and the other conditions of the theorem be fulfilled. Then there is an L such that

$$\mathsf{P}\left(\overline{\lim}\sqrt{\frac{n}{\log\log n}}\,|x_n - \theta| < L\right) = 1.$$

Remark: If $M'(\theta) = \beta \leqq \frac{1}{2}$, $a_n = \frac{1}{n}$ the same argument gives that

$$\mathsf{P}\left(\lim n^{\beta-\varepsilon}(x_n - \theta) = 0\right) = 1.$$

The simplest case is when $M(x) = \beta \cdot x$ we look for the root of $M(x) = 0$ and there is no error. In this case

$$x_n = (1-\beta)\left(1-\frac{\beta}{2}\right)\cdot\left(1-\frac{\beta}{n-1}\right)x_0 = x_0 \exp\sum_{i=1}^{n-1}\log\left(1-\frac{\beta}{i}\right) \sim C \cdot n^{-\beta}$$

with a constant C that is

$$\mathsf{P}\left(\lim n^{\beta}(x_n - \theta) = c\right) = 1.$$

and this shows that the exponent can not be improved.

MAGYAR TUDOMÁNYOS AKADÉMIA KÖNYVTÁRA

To prove theorem 2 we need the following lemma.

LEMMA 2. *Let* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *be independent identically distributed random variables, so that* $\mathsf{P}(|\xi_i| \leqq K)=1$ *for a constant* K *and* $\mathsf{E}\xi_i=0$.

Then there is a $\theta>1$ *and* $t>0$ *so that*

$$\sum_{i=\theta^n+1}^{\theta^{n+1}} \frac{1}{i}\,\xi_i > t\sqrt{\frac{\log\log\theta^n}{\theta^n}}$$

for infinitely many n *with probability* 1.

PROOF. Let us denote $\mathsf{D}\xi_i$ by D. Clearly

$$\sum_{i=\theta^n+1}^{\theta^{n+1}} \frac{1}{i}\,\xi_i = \sum_{i=\theta^n+1}^{\theta^{n+1}} \frac{1}{\theta^n}\,\xi_i + \sum_{i=\theta^n+1}^{\theta^{n+1}} \left(\frac{1}{i}-\frac{1}{\theta^n}\right)\xi_i.$$

To estimate the first member of the sum we need the inequality $1-\Phi(x) \geqq$ $\geqq \frac{1}{\sqrt{2\pi}}\left(\frac{1}{x}-\frac{1}{x^3}\right)e^{\frac{-x^2}{2}}$ ($\Phi(x)$ is the normal distribution) and the theorem of large deviation for equal components (see [3] p. 517)

$$\mathsf{P}\left(\sum_{\theta^n+1}^{\theta^{n+1}} \frac{1}{\theta^n}\,\xi_i > C_1 D\sqrt{\frac{\log\log\theta^n}{\theta^n}}\right) = \mathsf{P}\left(\frac{\sum_{i=1}^{\theta^n(\theta-1)} \xi_i}{D\sqrt{\theta^n(\theta-1)}} > C_1\sqrt{\frac{\log\log\theta^n}{\theta^n}}\right) \geqq$$

$$\geqq 1-\Phi\left[(C_1+\varepsilon)\sqrt{\frac{\log\log\theta^n}{\theta-1}}\right] \geqq \exp\left[-(C_1+\varepsilon)^2\frac{\log\log\theta^n}{2(\theta-1)}\right] = (n\log\theta)^{-\frac{(C_1+\varepsilon)^2}{2(\theta-1)}}.$$

So by the Borel—Cantelli lemma (the sums for different n's are independent) for arbitrary $\varepsilon>0$

$$\sum_{i=\theta^n+1}^{\theta^{n+1}} \frac{1}{\theta^n}\,\xi_i > \sqrt{(2-\varepsilon)(\theta-1)}\,D\sqrt{\frac{\log\log\theta^n}{\theta^n}} \tag{8}$$

infinitely many times with probability 1.

Estimating the second member of the sum by the same method as in Lemma 1 we get with $f(n) = \gamma\sqrt{\theta^n\log\log\theta^n}$.

$$\mathsf{P}\left(\sum_{i=\theta^n+1}^{\theta^{n+1}} \left(\frac{1}{\theta^n}-\frac{1}{i}\right)\xi_i \geqq C_2\sqrt{\frac{\log\log\theta^n}{\theta^n}}\right) =$$

$$= \mathsf{P}\left(\exp f(n)\sum_{\theta^n+1}^{\theta^{n+1}} \left(\frac{1}{\theta^n}-\frac{1}{i}\right)\xi_i \geqq \exp C_2 f(n)\sqrt{\frac{\log\log\theta^n}{\theta^n}}\right) \leqq$$

$$\leqq \prod_{\theta^n+1}^{\theta^{n+1}} \mathsf{E}\left(\exp f(n)\left(\frac{1}{\theta^n}-\frac{1}{i}\right)\xi_i\right)\mathsf{E}\left[\exp\left(-C_2 f(n)\sqrt{\frac{\log\log\theta^n}{\theta^n}}\right)\right] \leqq$$

$$\leqq \exp\left\{\frac{D^2}{2-\varepsilon}\gamma^2\sum_{i=\theta^n+1}^{\theta^{n+1}}\left(\frac{1}{\theta^n}-\frac{1}{i}\right)^2\log\log\theta^n - C_2\gamma\log\log\theta^n\right\}$$

for sufficiently large n's ($\varepsilon>0$ is fixed).

But

$$\sum_{i=\theta^n+1}^{\theta^{n+1}}\left(\frac{1}{\theta^n}-\frac{1}{i}\right)^2 \leqq \frac{\sum_{i=1}^{\theta^{n+1}-\theta^n} i^2}{\theta^{4n}} \leqq (1+\varepsilon)\frac{(\theta^{n+1}-\theta^n)^3}{3\theta^{4n}} = \frac{(\theta-1)^3}{3\theta^n}(1+\varepsilon).$$

So by a little calculation, using the Borel—Cantelli lemma we get that

$$\sum_{i=\theta^n}^{\theta^{n+1}}\left(\frac{1}{\theta^n}-\frac{1}{i}\right)\xi_i \leqq \left[\frac{\sqrt{6}}{3}D(\theta-1)^{\frac{3}{2}}-\varepsilon\right]\sqrt{\frac{\log\log\theta^n}{\theta^n}} \quad \text{for} \tag{9}$$

sufficiently large n with probability 1.

Since we can choose θ and ε as near to 1, and to 0, respectively as we want, Lemma 2 follows from (8) and (9).

Using lemma 2 theorem 2 can be proved easily.

PROOF of theorem 2. Apply lemma 2 for the sequence $y_i - M(x_i)$ and suppose that the statement of the theorem does not hold. Let $|M(x)| < \beta|x|$ if $|x| < \delta$ where β is a positive constant. Then with probability 1 for large k $|x_k(\omega)| \leqq \delta$ and $|x_k(\omega)| \leqq L'\sqrt{\frac{\log\log k}{k}}$ with a constant L' such that $L'(2+\beta\log\theta) < t$. But then

$$|x_{\theta^{n+1}+1} - x_{\theta^n+1}| \leqq 2L'\sqrt{\frac{\log\log\theta^n}{\theta^n}}$$

$$\left|\sum_{i=\theta^n+1}^{\theta^{n+1}}\frac{1}{i}M(x_i)\right| \leqq L'\sqrt{\frac{\log\log\theta^n}{\theta^n}}\sum_{i=\theta^n+1}^{\theta^{n+1}}\frac{\beta}{i} \leqq L'\beta\sqrt{\frac{\log\log\theta^n}{\theta^n}}\log\theta$$

$$\sum_{i=\theta^n+1}^{\theta^{n+1}}\frac{1}{i}[y_i - M(x_i)] = -x_{\theta^{n+1}+1} + x_{\theta^n+1} - \sum_{i=\theta^n+1}^{\theta^{n+1}}\frac{1}{i}M(x_i) \leqq$$

$$\leqq (2L' + \beta\log\theta)\sqrt{\frac{\log\log\theta^n}{\theta^n}}$$

for sufficiently large n contradicting to lemma 2.

REFERENCES

[1] ROBBINS, H. and MONRO, S.: A stochastic approximation method, *Annals of Math. Stat.* **22** (1951), 400—407.

[2] BLUM, J. R.: Approximation methods which converge with probability one, *Annals of Math. Stat. 25* (1954), 386—388.

[3] FELLER, W.: *An introduction to Probability Theory and its Applications,* Vol. II. John Wiley and Sons, 1959.

[4] DOOB, J. L.: *Stochastic processes,* Wiley, New York, 1953.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received February 10, 1972.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 103—112.

О НЕКОТОРЫХ ЗАДАЧАХ, ОТНОСЯЩИХСЯ К ОДНОМУ СЕМЕЙСТВУ УРАВНЕНИЙ В ЧАСТНЫХ ПРОИЗВОДНЫХ СОСТАВНОГО ТИПА

ЧАН СУАНЪ ХИЕН

В настоящей статье речь идёт об одном семействе уравнений в частных производных третьего порядка для неизвестных функций $u(x, y)$ двух действительных переменных x и y. Эти уравнения мы выводим из вырожденных на границе линейных уравнений в частных производных эллиптического типа второго порядка. Рассматриваемые уравнения третего порядка мы назовём уравнениям „составного типа" потому что во внутренных точках их области определения они владеют и действительными и мнимыми характеристическими направлениями (см. [1] стр. 136.)

В дальнейшем мы будем заниматься одним уравнением частного вида этого семейства а именно уравнением

$$Lu = \frac{\partial}{\partial x}[Eu] = 0 \tag{1}$$

где функция $u(x, y)$ искомое решение, производные которой вплоть до тех, которые в уравнении (1) фигурируют, существуют и непрерывны. Такие решения мы будем называть „регулярными решениями". Выражение в квадратных скобках такое:

$$Eu \equiv y^m u_{xx} + u_{yy} + a(y)u_x + b(y)u_y + c(y)u.$$

Здесь $a(y)$, $b(y)$ и $c(y)$ аналитические функции действиткльно переменного y если $(x, y) \in D$, где D та область плоскости (x, y) в которой мы ищем решение. Об области D мы предполагаем, что она является подобластью верхней полуплоскости $y > 0$, кроме того, что она ограничена отрезком $[-p, p]$ оси x и простой гладкой жордановой другой Γ лежащей в полуплоскости $y > 0$ конечными точками которой являются точки $A(-p, 0)$ и $B(p, 0)$ $(p = \text{const} > 0)$. Предположим, что каждая прямая $y = c$, $0 \leqq c \leqq h$ $(h > 0)$ пересекается с дугой Γ в двух точках, если $0 \leqq c \leqq h$ и если $y = h$ то эта прямая имеет единственную общую точку N с дугой Γ. Если $c > h$, то прямая $y = c$ обших точек с Γ не имеет. Далее предположим, что m действительное положительное число и $c(y) \leqq 0$, если $(x, y) \in D$. Предмет настоящей статьи: формулировка двух задач по решению уравнения (1) и доказательство существования и единственности этих решений в заданной области D описуемого типа.

§ 1. Основная Лемма

Лемма. *Любое регулярное решение $u(x, y)$ уравнения (1) может быть представлено в некоторой окрестности любой точки $(x, y) \in D$ в виде*

$$u(x, y) = v(x, y) + w(y) \tag{2}$$

где функция $v(x, y)$ является решением уравнения

$$Ev(x, y) = 0 \tag{3}$$

а $w(y)$ дважды непрерывно дифференцируемая функция переменного y.

Пользуясь свойствами области D можно показать (путём продолжения решения из окрестности некоторой точки на всю области D) что представление (2) имеет силу во всей области D. (Легко видеть, что и обратное утверждение справедливо, а именно функция $u(x, y)$, представляемая уравнением (2) является решением уравнения (1) так как и $v(x, y)$ и $w(y)$ являются решениями уравнения (1)).

Доказательство Леммы:

Пусть $u(x, y)$ регулярное решение уравнения (1), тогда

$$Lu \equiv \frac{\partial}{\partial x}[Eu] = Eu_x = 0$$

итак функция $u_x = v_1$-решение уравнения (3).

Отсюда следует, что в некоторой окрестности точки $(x_0, y_0) \in D$:

$$u(x, y) = \int_{x_0}^{x} v_1(t, y)\,dt + w_1(y). \tag{4}$$

Эту окрестность мы выбирали так, чтобы уравнение (4) там имело смысл (см. [1]). Заметим, что фупция $w_1(y)$ дважды непрерывно дифференцируема. Это следует из регулярности $u(x, y)$. Теперь мы докажем, что существует дважды непрерывно дифференцируемая функция $w_2(y)$ такая, чтобы в рассматриваемой окрестности функция

$$v(x, y) = \int_{x_0}^{x} v_1(t, y)\,dt + w_2(y) = \int_{x_0}^{x} u_t(t, y)\,dt + w_2(y) \tag{5}$$

была рещением уравнения (3). Такую функцию мы можем найти следующим образом. Подставляя функцию $v(x, y)$ из (5) в уравнение (3) и учитывая, что $Eu(x, y)$ зависит только от y, то есть $Eu(x, y) = \varphi(y)$, мы получим для $w_2(y)$ следующее обыкновенное дифференциальное уравнение:

$$\begin{aligned} w_2''(y) + b(y)w_2'(y) + c(y)w_2(y) - u_{yy}(x_0, y) - \\ - b(y)u_y(x_0, y) - c(y)u(x_0, y) + \varphi(y) = 0. \end{aligned} \tag{6}$$

Любое решение $w_2(y)$ этого уравнения даёт функцию

$$v(x, y) = \int_{x_0}^{x} u_t(t, y)\,dt + w_2(y)$$

удовлетворящую уравнению (3). Если обозначим разность $w_1(y) - w_2(y)$ через $w(y)$, то из уравнений (4) и (5) заключаем справедливость нашего утверждения.

Подчеркнём, что всякое решение $w_2(y)$ уравнения (6) даёт функцию $v(x, y)$ удовлетворящую уравнению (3) и поэтому мы имеем возможность выбирать

$w_2(y)$ согласно потребностями для решения той или иной задачи. Так, например мы можем выбрать $w_2(y)$ так, чтобы было:

$$w_2(0)=w_1(0) \quad \text{и} \quad w_2(h)=w_1(h)$$

или

$$w_2'(0)=w_1'(0) \quad \text{и} \quad w_2(h)=w_1(h)$$

$(h=\text{const}\neq 0)$.

Иными словами мы можем без ограничения общности предполагать, то фигурирующая в разложении (2) функция $w(y)$ удовлетворяет требованиям

$$w(0) = w(h) = 0 \tag{7}$$

или

$$w'(0) = w(h) = 0. \tag{8}$$

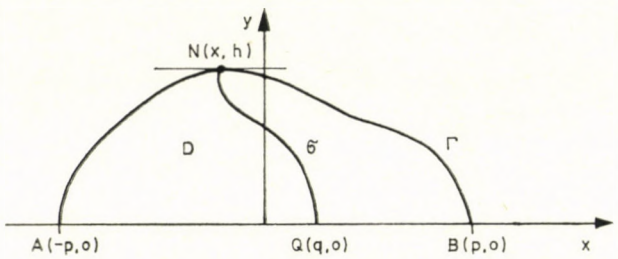
§ 2. Постановка задач

Задача I. Требуется найти регулярное в области D решение $u(x, y)$ уравнения (1), непрерывное в замкнутой области $\bar{D}$ удовлетворяющее следующим условиям:

$$\left.\begin{aligned} u(x, y)|_{\Gamma} &= f_1(s), \quad 0\leqq s\leqq l \\ u(x, y)|_{AB} &= \tau(x), \quad -p\leqq x\leqq p \end{aligned}\right\} \tag{9}$$

$$u(x, y)|_{x=\sigma(y)} = \varphi_1(y), \quad 0\leqq y\leqq h \tag{10}$$

где s означает длину дуги по кривой Γ отсчитываяся от точки $B(p, 0)$, где $s=0$, до точки $A(-p, 0)$, где $s=l$; $f_1(s)$, $\tau(x)$ заданные непрерывные функции переменных s, соответственно x а $\varphi_1(y)$ дважды непрерывно дифференцируемая функция от y. Кривая $x=\sigma(y)$ фигурирующая в условии (10) — простая жорданова дуга с конечными точками $Q(q, 0)$, где $-p<q<p$ и $N(x, h)$ и впрочем, целиком находящаяся в области D (см. Фиг. 1). Кроме того, должны быть



Фиг. 1

$f_1(0)=\tau(p)$; $f_1(l)=\tau(-p)$; $\varphi_1(0)=\tau(q)$; $f_1(l_1)=\varphi_1(h)$, где l_1 длина дуги BN на кривой Γ и $\sigma(0)=q$.

Задача II. Требуется найти регулярное в области D решение уравнения (1) непрерывное в замкнутой области $\bar{D}$ с непрерывной первой производной

$\frac{\partial u}{\partial y}$ в $D+[x: -p<x<p]$ и удовлетворяющее условиям

$$\left.\begin{aligned} u(x, y)|_{\Gamma} &= f_2(s), \quad 0 \leqq s \leqq l \\ \left.\frac{\partial u}{\partial y}\right|_{AB} &= \nu(x), \quad -p<x<p \end{aligned}\right\} \tag{11}$$

$$u(x, y)|_{x=\sigma(y)} = \varphi_2(y), \quad 0 \leqq y \leqq h \tag{12}$$

Здесь $f_2(s)$ и $\nu(x)$ заданные непрерывные функции переменных s, соответственно x, а функция $\varphi_2(y)$ по крайней мере дважды непрерывно дифференцируемая функция переменного y. Здесь должны быть удовлетворены равенства: $f_2(l_1)=$ $=\varphi_2(h)$, $\sigma(0)=q$ (кривая $x=\sigma(y)$ та же самая что и в задаче 1) и елси $\sigma'(0)=0$ то должно быть $\varphi_2'(0)=\nu(q)$.

§ 3. Единственность решений задач I и II

Применяя основную лемму в § 1, довольно легко можно доказать единственность решений вышесформулированных задач I и II. Мы докажем, что решения обоих задач, если они существуют, единственные, если только условие

$$c(y) \leqq 0 \tag{13}$$

выполнено в области D.

В случае первой задачи, чтобы доказать единственность решений, мы должны показать, что если условия (9) и (10) однородны то-есть $f_1(s)\equiv 0$, $\tau(x)\equiv 0$ и $\varphi_1(y)\equiv 0$ единственным возможным решением уравнения (1) в области D является: $u(x, y)\equiv 0$. Учитывая эти условия и представление (2) решения $u(x, y)$ и условия (7) для функции $w(y)$ фигурирующей в (2) мы получим для функции $v(x, y)$ следующие требования:

$$v(x, y)|_{\Gamma} = -w\big(y(s)\big)$$

$$v(x, y)|_{AB} = 0$$

$$v(x, y)|_{x=\sigma(y)} = -w(y); \quad 0 \leqq y \leqq h.$$

Как известно, эта функция $v(x, y)$ будучи решением вырожденного на границе области эллиптического уравнения второго порядка не может принимать своего максимального значения при условии (13) во внутренней точке области (см. [7]), кроме случая $v(x, y)\equiv\text{const}$, и так по вышеприведенным условиям для $v(x, y)$ только этот случай возможен. Однако, тогда необходимым образом $v(x, y)\equiv 0$ а также $w(y)\equiv 0$ так, что $u(x, y) = v(x, y)+w(y)$ тоже исчезает, что мы и хотели доказать.

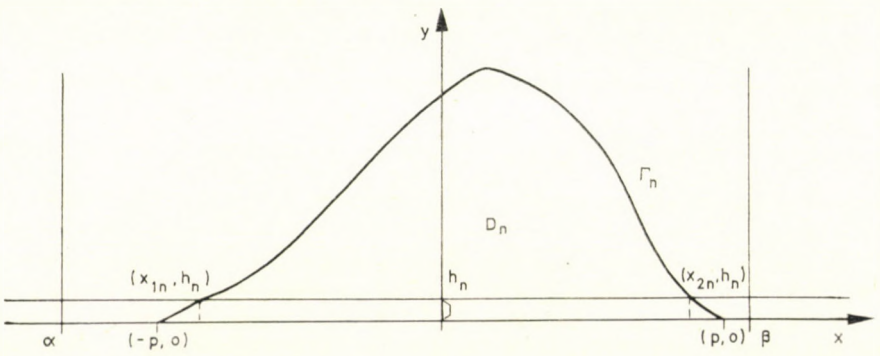
Таким же образом можно доказать единственность решении задачи II для уравнении (1), используя при этом условия (8) для функции $w(y)$.

§ 4. Существование решений задачи I

В этом параграфе мы учитываем, что фигурирующие в уравнении (1) функции (коэффициенты) $a(y)$, $b(y)$ и $c(y)$ — аналитические функции переменного y [поскольку точка $(x, y) \in D$] а в области D имеет силу представление (2) для решений уравнения (1). Сконструируем теперь некоторую последовательность областей D_n в которых уравнение (1) является эллиптического (не вырожденного) типа. Пусть каждая область D_n — подобласть области D а граница S_n от D_n пусть состоит из той дуги Γ_n кривой Γ, которая находится в полуплоскости $y > h_n$, где h_n некоторое положительное число, и из точек $\overline{D}$ находящихся на прямой $y = h_n$. Предположим далее, что последовательность положительных чисел $h_1, h_2, \ldots$ стремится монотонно к нулю, если $n \to \infty$. Согласно уравнениюю (2) регулярное решение уравнения (1), если оно существует в области D_n, представляема в виде

$$u_n(x, y) = v_n(x, y) + w_n(y)$$

где $v_n(x, y)$ регулярное решение уравнения (3) в области D_n а $w_n(y)$ дважди непрерывно дифференцируема по переменному y. Как первый шаг к доказательству существования решения задачи I мы хотим показать, что можно найти такую функцию $w_n(y)$ совместно с функцией $v_n(x, y)$, чтобы предыдущее представление дало решение уравнения (1) и удовлетворяло граничным условиям вытекающим из условии (9) и (10). Эти граничные условия для $u_n(x, y)$ пусть будут такими: на кривой Γ_n граничные значения $u_n(x, y)$ совпадают с граничными значениями $u(x, y)$ предписанными задачей I а на прямолинейной части S_n граничные значения $u_n(x, y)$ определяются некоторой непрерывной функцией $\tau_n(x)$. Эту функцию $\tau_n(x)$ мы задаём следующим образом: Пусть в некоторой полосе $\alpha < x < \beta$ плоскости (x, y) содержащей в себе область $\overline{D}$ в качестве подобласти, мы зададим непрерывную функцию $t(x, y)$, которую мы задали так, чтобы на промежутке $[-p, p]$ было $t(x, 0) = \tau(x)$. (Это возможно получить бесконечно многими способами, например в полосе $\alpha < x < \beta$ пусть будет $t(x, y) = \tau_{\alpha\beta}(x)$, где на променутке $[-p, p]$ $\tau_{\alpha\beta}(x) = \tau(x)$), тогда мы определим $\tau_n(x)$ так: $\tau_n(x) = t(x, h_n)$ поскольку точка (x, h_n) принадлежит к замкнутой области $\overline{D}_n$. Конечные точки прямолинейной части границы S_n обозначим через (x_{1n}, h_n), (x_{2n}, h_n) $(x_{1n} \leqq x \leqq x_{2n}$ см. Фиг. 2).



Фиг. 2

Если предположить, что $u_n(x, y)$ регулярное решение задачи I при вышезаданных граничних условиях в области $\bar{D}_n$ и для которого следовательно имеет силу разножение (2), то на основании условий (9) и (10) задачи I функция $v_n(x, y)$ должна удовлетворять следующим граничным условиям.

$$\left.\begin{aligned} v_n(x, y)|_{\Gamma_n} &= f_1(s) - w_n[y(s)] \\ v_n(x, h_n) &= \tau_n(x) \quad x_{1n} \leqq x \leqq x_{2n} \end{aligned}\right\} \tag{14}$$

где $y(s)$ означает значение y-а на кривой Γ соответствующей длинне s дуги на кривой Γ отсчитываемой от $B(p, 0)$. На кривой $x=\sigma(y)$ должно быть:

$$v_n(x, y)|_{x=\sigma(y)} = \varphi_1(y) - w_n(y) \quad (h_n \leqq y \leqq h). \tag{15}$$

Для этих граничных условий решение уравнения (3) в области D_n известно. Выберем некоторый зафиксированный индекс n_0 и решим граничную задачу для $v_{n_0}(x, y)$ соответствующую задаче I. Обозначим через $G_{n_0}(x, y; \xi, \eta)$ функцию Грина области D_{n_0} относящуюся к задаче Дирихлета для уравнения (3). Тогда решение $v_n(x, y)$ задачи Дирихлета для $n > n_0$ относительно области D_n представляется в области D_{n_0} выражением.

$$v_n(x, y) = \int_{S_{n_0}} v_n(s) \frac{\partial G_{n_0}(x, y;\ \xi, \eta)}{\partial \nu}\, ds, \tag{16}$$

где s означает длину дуги до переменной s по всему контуру S_{n_0} точки отсчитываемой от некоторой фиксированной точки, например $B(p, 0)$, а ν направнелие внутренней нормали по S_{n_0} (где она существует). Итак

$$v_n(s) = v_n(x, y)|_{\Gamma_n}, \quad y > h_n$$

и

$$v_n(s) = \tau_n(x), \quad y = h_n$$

(где значение s принимается) соответственно смыслу этих уравнений. Для того, чтобы было удовлетворено и условие (15), то есть чтобы $v_n(x, y)$ имела и на кривой $x=\sigma(y)$ заданные этим условием значения, необходимо и достаточно, чтобы функция $w_n(y)$ была решением следующего интегрального уравнения [получаемого из условия (15)]:

$$\begin{aligned} w_n(y) + \int_{\Gamma_{n_0}} \{f_1(s) - w_n[\eta(s)]\} \frac{\partial G_{n_0}[\sigma(y), y;\ \xi, \eta]}{\partial \nu}\, ds = \\ = \varphi_1(y) - \int_{x_{1n_0}}^{x_{2n_0}} \tau_n(\xi) \frac{\partial G_{n_0}[\sigma(y), y;\ \xi, h_{n_0}]}{\partial \eta}\, d\xi. \end{aligned} \tag{17}$$

Здесь в первом интеграле $\xi=\xi(s)$; $\eta=\eta(s)$—координаты переменной точки вдоль кривой Γ_{n_0}. Ядро этого интегрального уравнения как функция от ξ, η непрерывно кроме в точке N. (N—особая точка так как там $\xi=\sigma(y): y=\eta$). Применяя теоремы фредгольма на уравнение (17) мы видим, что оно имеет единственное решение $w_n(y)$ — его существование следует из единственности решения. После-

довательности функций $\{w_n\}$ и $\{v_n\}$ и поэтому также $\{u_n\}$ органичены. Из соотношений (16) и (17) следует равностепенная непрерывность функций множеств $\{w_n\}$ и $\{v_n\}$ и поэтому также функций множества $\{u_n\}$, то есть множество $\{u_n\}$ компактно в любой области D_{n_0}. Известным способом заключаем, что из элементов $\{u_n\}$ можно строить подпоследовательность $\{u_{n_k}\}$ равномерно сходящимся в любой подобласти D_{n_0} области D к некоторому решению уравнения (1) удовлетворящему условиям (9) и (10) в области $\bar{D}$.

Примечания

а) подобным же образом можно исследовать единственность и существование решений уравнения $\frac{\partial}{\partial x}[E_1 u]=0$ в рассматриваемой области D, где

$$E_1 u = u_{xx}+y^m u_{yy}+a(y)u_x+b(y)u_y+c(y)u.$$

(Прямая $y=0$ здесь не только параболическая кривая уравнений а также характеристика его). Это исследование выполнимо тогда, если задача Дирихлета относящаяся к уравнению $E_1 u=0$ разрешима, то есть, если число m и функция $b(y)$ удовлетворяют одному из следующих условий

1) $m<1$; 2) $m=1$ и $b(0)<1$

3) $1<m<2$ и $b(0)\leqq 0$; 4) $m\geqq 2$ и $b(0)<0$ (см. [6])

б) вышеполученные результаты остаются в силе также, если коэффициент $a(x, y)$ зависит не только от переменного y но также если он является аналитической функцией двух переменных x и y в области D.

§ 5. Задачи относящиеся к уравнению $L_0 u_x + a u_x + b u_y + c u = f$

В дальнейшем мы исследуем вопрос, связанный с разрешимостью задач, относящихся к уравнению

$$L_0 u_x + a u_x + b u_y + c u \equiv y^m u_{xxx} + u_{xyy} + a(x, y)u_x + b(x, y)u_y + c(x, y)u = f(x, y). \tag{18}$$

Здесь мы предполагаем, что функции двух переменных x и y, $a(x, y)$, $b(x, y)$, $c(x, y)$ и $f(x, y)$ один раз непрерывно дифференцируемы. Применая результат § 3 мы видем, что решения задач I и II для уравнения (18) единственны если коэффициенты $b(x, y)\equiv 0$, $c(x, y)\equiv 0$, и $a(x, y)\leqq 0$ зависит только от y в области D.

Рассмотрим сначала уравнение

$$y^m u_{xxx} + u_{xyy} = f_2(x, y) \tag{19}$$

частным решением которого является функция (см. [4]):

$$u_2(x, y) = \iint_D p_2(\xi, \eta;\ x, y) f_2(\xi, \eta)\, d\xi\, d\eta \tag{20}$$

где

$$p_2(\xi, \eta;\ x, y) = -\int_{\sigma(y)}^{x} q_2(\xi, \eta;\ t, y)\, dt$$

и $q_2(\xi, \eta;\ x, y)$ „второе фундаментальное решение" уравнения

$$y^m u_{xx} + u_{yy} = 0 \tag{21}$$

(см. [3] и [5]). ξ, η-параметры, $\xi \neq x$ и $\eta \neq y$. Согласно [3] (стр. 41) мы имеем:

$$q_2(x, y;\ x_0, y_0) = k_2 \left(\frac{4}{m+2}\right)^{4\beta-2} (r_1^2)^{-\beta} (1-\gamma)^{1-2\beta} F(1-\beta, 1-\beta, 2-2\beta;\ 1-\gamma),$$

здесь

$$\left.\begin{matrix} r^2 \\ r_1^2 \end{matrix}\right\} = (x-x_0)^2 + \frac{4}{(m+2)^2}\left(y^{\frac{m+2}{2}} \mp y_0^{\frac{m+2}{2}}\right)^2$$

$$\gamma = \frac{r^2}{r_1^2}; \quad \beta = \frac{m}{2(m+2)}; \quad k_2 = \frac{1}{4\pi}\left(\frac{4}{m+2}\right)^{2-2\beta} \frac{\Gamma^2(1-\beta)}{\Gamma(2-2\beta)},$$

$F(a, b, c; z)$ гипергеометрическая функция Гаусса в окресности точки $z=0$ а $\Gamma(x)$ функция гамма Эйлера, (x_0, y_0) некоторая точка верхней полуплоскости. Так как уравнение (19) линейно, то пользуясь частным его решением (20) мы ищем его решение в виде

$$u(x, y) = u_2(x, y) + v_2(x, y),$$

здесь $v_2(x, y)$ удовлетворяет однородному уравнению

$$\frac{\partial}{\partial x}\left[y^m \frac{\partial^2 v_2}{\partial x^2} + \frac{\partial^2 v_2}{\partial y^2}\right] = 0. \tag{22}$$

Имея в виду задачу I для уравнения (19), мы получим на основании уравнений (9) и (10) для $v_2(x, y)$ следующие граничные условия:

$$\left.\begin{aligned} v_2(x, y)|_\Gamma &= f_1(s) - u_2(x, y)|_\Gamma; \quad 0 \leqq s \leqq l \\ v_2(x, 0) &= \tau(x); \quad -p \leqq x \leqq p \\ &\text{(так как } u_2(x, 0) \equiv 0\text{)} \\ v_2(x, y)|_{x=\sigma(y)} &= \varphi_1(y) - u_2(x, y)|_{x=\sigma(y)}; \quad 0 \leqq y \leqq h \end{aligned}\right\} \tag{23}$$

$(u_2[\sigma(y), y] \equiv 0)$.

Автор в своей диссертации (2) дал решение задачи выраженной уравнениями

(22) и (23). Для решения $v_2(x, y)$ получается:

$$v_2(x, y) = \int_{-p}^{p} \tau(\xi) \frac{\partial G_2(\xi, 0; x, y)}{\partial \eta} d\xi -$$

$$- \int_{\Gamma} \{f_1(s) - u_2|_{\Gamma} - w[\eta(s)]\} A_s[G_2(\xi, \eta; x, y)] ds + w(y)$$

(во втором интеграле ξ и η-функции от s). Здесь $G_2(\xi, \eta; x, y)$ — функция Грина задачи Дирихлета для области D, относящаяся к уравнению (21), оператор $A_s[\cdot] \equiv \eta^m \frac{d\eta}{ds} \frac{\partial}{\partial \xi} - \frac{d\xi}{ds} \frac{\partial}{\partial \eta}$, а функция $w(y)$ однозначно определена интегральным уравнением

$$w(y) + \int_0^l w[\eta(s)] A_s[G]|_{x=\sigma(y)} ds = F_2(y),$$

где

$$F_2(y) = \varphi_1(y) - \int_{-p}^{p} \tau(\xi) \left. \frac{\partial G_2}{\partial \eta} \right|_{x=\sigma(y)} d\xi + \int_0^l \{f_1(s) - u_2|_{\Gamma}\} A_s[G_2]|_{x=\sigma(y)} ds.$$

Для $v_2(x, y)$ мы можем написать иначе:

$$v_2(x, y) = \iint_D S_2^*(\xi, \eta; x, y) f_2(\xi, \eta) d\xi \, d\eta + P_2(x, y) \tag{24}$$

где функция $P_2(x, y)$ от $f_2(x, y)$ не зависит и непрерывна. С этим решение первой задачи относящейся к уравнению (19) принимает вид

$$u(x, y) = \iint_D S_2(\xi, \eta; x, y) f_2(\xi, \eta) d\xi \, d\eta + P_2(x, y)$$

где

$$S_2(\xi, \eta; x, y) = S_2^*(\xi, \eta; x, y) + p_2(\xi, \eta; x, y). \tag{25}$$

Если мы рассмотрим теперь вопрос о разрешимости первой задачи относительно уравнения (18), то увидим, что этот вопрос равносилен вопросу разрешимости интегрального уравнения

$$f_2(x, y) + \iint_D K(\xi, \eta; x, y) f_2(\xi, \eta) d\xi \, d\eta = f_3(x, y)$$

где

$$K(\xi, \eta; x, y) = a \frac{\partial S_2}{\partial x} + B \frac{\partial S_2}{\partial y} + c S_2 \tag{26}$$

$$f_3(x, y) = f - a \frac{\partial P_2}{\partial x} - b \frac{\partial P_2}{\partial y} - c P_2,$$

и где мы функцию $f_2(\xi, \eta)$, фигурирующую в формуле (25) считаем неизвестной.

В том случае если $b \equiv 0$, $c \equiv 0$, $a(y) \leqq 0$ в области D, однородное интегральное уравнение, соответствующее уравнению (26) имеет только тривиальное

решение; и таким образом применяя к нему теоремы Фредгольма мы можем утверждать, что решение уравнения (26) существует. Таким же способом, каким мы рассматривали существование решения первой задачи мы можем рассмотреть вопрос о существовании решения второй задачи.

ЛИТЕРАТУРА

[1] HADAMARD, J.: Propriétés d'une équation linéarie aux dérivées partielles du quatrième ordre. *The Tôhoku Math. Journal.* **37** (1933), 133—150.
[2] Чан Суань Хиен: *Кандидатская диссертация* Будапешт, 1972.
[3] Смирнов, М. М.: *Вырождающиеся эллитические и гиперболические уравнения.* Изд-во „Наука" Москва, 1966.
[4] Салахитдинов, М. С.: К теории уравнений составного типа, *Изв. АН УССР сер. физмат.* №3, 1965, 13—20.
[5] GELLERSTEDT, S.: Sur un problème aux limites pour l'équation $y^{2s}z_{xx}+z_{yy}=0$. *Arkiv Mat. Astr. och Fysik* **10** (1935) B 25 A.
[6] Келдыш М. В.: О некоторых случях вырождения уравнений эллиптического типа на границе области.
ДАН СССР. **77** №2, 1951, 181—183.
[7] Бицадзе, А. В.: *Уравнения смешанного типа,* Изд-во АН СССР 1959.

Математический Институт Академии Наук Венгрии, Будапешт

(Поступила 22-ого февраля, 1972 г.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 113—122.

ЗАМЕЧАНИЯ О ТРИГОНОМЕТРИЧЕСКОМ ИНТЕРПОЛИРОВАНИИ И СУММАХ ФУРЬЕ

Г. П. НЕВАЙ

§ 1. Введение

Под $f(x)$ будем понимать ограниченную функцию периода 2π. Обозначим через $\tilde{S}_n(x, f)$ $(n=1, 2, \ldots)$ тригонометрический интерполяционный полином порядка не выше n, совпадающий с функцией $f(x)$ в узлах

$$x_k = x_{kn} = \frac{2k\pi}{2n+1} \quad (|k| = 0, 1, 2, \ldots), \tag{1.1}$$

то есть

$$\tilde{S}_n(x, f) = \frac{2}{2n+1} \sum_{k=-n}^{n} f(x_k) D_n(x - x_k),$$

где

$$D_n(t) = \frac{\sin\left(n+\frac{1}{2}\right)t}{2\sin\frac{t}{2}}$$

ядро Дирихле n-го порядка. Пусть далее $S_n(x, f)$ $(n=1, 2, \ldots)$ n-ая частная сумма тригонометрического ряда Фурье функции $f(x)$*. Значит

$$S_n(x, f) = \frac{1}{\pi} \int_0^{2\pi} f(t) D_n(x-t)\, dt.$$

Обозначим через

$$\Delta_\delta^R f(x) = \sum_{k=0}^{R} (-1)^{R-k} \binom{R}{k} f(x + k\delta) \quad (R = 0, 1, 2, \ldots)$$

R-ую разность функции $f(x)$ с шагом δ. Если $\delta = \frac{2\pi}{2n+1} \equiv h$, то будем писать $\Delta^R f(x)$ вместо $\Delta_h^R f(x)$. Далее

$$\omega_R(f; \delta) = \sup_{-\infty < x < \infty,\ |\varepsilon| \leqq \delta} |\Delta_\varepsilon^R f(x)| \quad (R = 0, 1, 2, \ldots)$$

модуль гладкости порядка R функции $f(x)$.

Обозначим через $E_n(f)$ $(n=1, 2, \ldots)$ меру наилучшего приближения функции $f(x)$ — с помощью тригонометрических полиномов $H_n(x)$ порядка не выше n — в равномерной метрике, а $H_n(x, f)$ тот полином, для которого

$$E_n(f) = \max_{|x| \leqq \pi} |f(x) - H_n(x, f)|.$$

---

* Если речь идет о ряде Фурье, то функция $f(x)$ предполагается интегрируемой в смысле Римана.

Будем говорить, что функция $f(x)$ имеет ограниченное (R, p) $(R=1, 2, \ldots; p \geqq 1)$ изменение в узлах (1.1), если

$$V(f, R, p, \{x_k\}) = \sup_n \left(\sum_{k=-n}^{n} |\Delta^R f(x_k)|^p \right)^{\frac{1}{p}} < \infty. \tag{1.2}$$

Отметим, что если выполнено условие (1.2), то и для всех $R_1 > R$ и $1 \leqq p_1 < p$ будет $V(f, R_1, p_1, \{x_k\}) < \infty$ также. Континуальный аналог ограниченности $(R, 1)$ вариации ввел Ф. И. Харшиладзе [3].

§ 2. Об отклонении тригонометрических интерполяционных сумм

В настоящем параграфе мы примем обозначение

$$D(R, f, n) = \max_{\varepsilon=\pm 1} \max_{|j| \leqq n} \left| \sum_{\substack{k=1 \\ k \text{ нечетное}}}^{n} \frac{\Delta^R f(x_{\varepsilon \cdot k+j})}{k} \right| \quad (R, n=1, 2, \ldots), \tag{2.1}$$

и докажем, что справедлива.

Теорема 2.1. *Пусть R фиксированное натуральное число. Для уклонения тригонометрических интерполяционных сумм справедлива оценка*

$$|f(x) - \tilde{S}_n(x, f)| \leqq C_1 2^{-R} D(R, f, n) + C_2 \omega_R\left(f; \frac{1}{n}\right) \quad (n=1, 2, \ldots), \tag{2.2}$$

где C_1 абсолютная, а C_2 от R зависящая постоянные.

Отсюда получаем

Следствие 2.2. *Если при некотором натуральном R выражение $D(R, f, n)$ стремится к нулю при $n \to \infty$ и функция $f(x)$ непрерывна, то $\tilde{S}_n(x, f)$ равномерно на всей оси сходится $f(x)$.*

Отметим, что аналогичный признак для сходимости рядов Фурье в случае $R=1$ был сформуливан Р. Салемом [8].

Было бы интересно выяснить, что Следствие 2.2 обратимо ли, то есть из равномерной сходимости $\tilde{S}_n(x, f)$ к непрерывной функции $f(x)$ вытекает ли, что $D(R, f, n) \to 0$ для некоторого R.

Следствие 2.3. *Если функция $f(x)$ имеет ограниченное (R, p) $(R=1, 2, \ldots; p \geqq 1)$ изменение в узлах (1.1), то*

$$|f(x) - \tilde{S}_n(x, f)| = O\left[\left| \log \omega_R\left(f; \frac{1}{n}\right) \right| \cdot \omega_R\left(f; \frac{1}{n}\right) \right], \tag{2.3}$$

где константа в „O"-члене зависит лишь от R, p и $V(f, R, p\{x_k\})$. Если функция $f(x)$ также непрерывна, то выражение, стоящее в правой части (2.3), стремится к нулю при $n \to \infty$.

Для простоты обозначений предположим, что максимум в (2.1) достигается для $\varepsilon=1, j=0$. Тогда*

$$D(R,f,n) \leqq \left(\sum_{k=1}^{N}+\sum_{k=N+1}^{n}\right)\frac{|\Delta^R f(x_k)|}{k} \leqq$$

$$\leqq C_3 \log N\cdot\omega_R\left(f,\ \frac{1}{n}\right)+\begin{cases} 0, & \text{если} \quad N\geqq n,\\ \sum\limits_{k=N+1}^{n}\dfrac{|\Delta^R f(x_k)|}{k}, & \text{если} \quad N<n,\end{cases}$$

где

$$N=\min\left\{n,\ \left(\frac{V(f,R,p,\{x_k\})}{\omega_R\left(f;\ \frac{1}{n}\right)}\right)^p\right\}.$$

Значит при $p=1$

$$D(R,f,n) \leqq C_3 \log N\cdot\omega_R\left(f;\ \frac{1}{n}\right)+\frac{V(f,R,1,\{x_k\})}{N},$$

если же $p>1$, то в силу неравенства Гельдера

$$D(R,f,n) \leqq C_3 \log N\cdot\omega_R\left(f;\ \frac{1}{n}\right)+\left(\sum_{k=N+1}^{n}|\Delta^R f(x_k)|^p\right)^{\frac{1}{p}}\left(\sum_{k=N+1}^{n}\frac{1}{k^q}\right)^{\frac{1}{q}} \leqq$$

$$\leqq C_3 \log N\cdot\omega_R\left(f;\ \frac{1}{n}\right)+C_4\frac{V(f,R,p,\{x_k\})}{N^{\frac{1}{p}}},$$

где $\dfrac{1}{p}+\dfrac{1}{q}=1$. Таким образом оценка (2.3) доказана.

Относительно (2.3) при $p=1$ см. еще Г. П. Неваи [6] и [7].

Отметим, что Теорема 2.1 очевидно содержит в себе оценку Лебега

$$|f(x)-\tilde{S}_n(x,f)| = O\left[\log n\cdot\omega_R\left(f;\ \frac{1}{n}\right)\right].$$

Доказательство Теоремы 2.1. Исходим из формулы

$$\tilde{S}_n(x,f)-f(x) = (-2)^{-R}\tilde{S}_n(x,\Delta^R f)+O_R\left[\omega_R\left(f;\ \frac{1}{n}\right)\right], \tag{2.4}$$

которая доказана в статье Г. П. Неваи [5]. Если точка x лежит между узлами

* c_3, c_4, c_5, c_7 то R зависящие, а c_6 абсолютная постоянные.

x_α и $x_{\alpha+1}$, то по хорошо известной формуле

$$(2.5)\quad \tilde{S}_n(x, \Delta^R f) = \sin\left(n+\frac{1}{2}\right)x\left(\sum_{k=\alpha-n}^{\alpha-1} + \sum_{k=\alpha,\,\alpha+1}\sum_{k=\alpha+2}^{\alpha+n}\right)\Delta^R f(x_k)\frac{(-1)^{k-1}}{(2n+1)\sin\frac{x_k-x}{2}}.$$

Средняя сумма здесь непревосходит $C_5\omega_R\left(f;\frac{1}{n}\right)$. Далее

$$\sum_{k=\alpha+2}^{\alpha+n}\Delta^R f(x_k)\frac{(-1)^{k+1}}{(2n+1)\sin\frac{x_k-x}{2}} \leqq \left(\left|\sum_{\substack{k=\alpha+2\\ k\text{ четное}}}^{\alpha+n}\right| + \left|\sum_{\substack{k=\alpha+2\\ k\text{ нечетное}}}^{\alpha+n}\right|\right)\frac{\Delta^R f(x_k)}{(2n+1)\sin\frac{x_k-x}{2}} \leqq$$

$$(2.6)\qquad \leqq C_6\left(\left|\sum_{\substack{k=\alpha+2\\ k\text{ четное}}}^{\alpha+n}\right| + \left|\sum_{\substack{k=\alpha+2\\ k\text{ нечетное}}}^{\alpha+n}\right|\right)\frac{\Delta^R f(x_k)}{k-\alpha} + C_7\omega_R\left(f;\frac{1}{n}\right).$$

Левая крайняя сумма в правой части (2.5) в точности также оценивается. Вывод (2.2) из формул (2.4)—(2.6) уже не представляет трудностей.

В заключение отметим, что результаты настоящего параграфа можно перенести в теорию рядов Фурье при надлежащем определении выражения $D(R, f, n)$ и понятия ограниченности (R, p) вариации.

§ 3. О коэффициентах Фурье—Лагранжа

Обозначим через $a_{kn}(f), b_{kn}(f)\,(k=1, 2, \ldots,)$ действительные коэффициенты Фурье—Лагранжа 2π-периодической функции $f(x)$, а через $c_{kn}(f)\,(|k|=0, 1, \ldots, n)$ соответствующие комплексные коэффициенты, то есть

$$(3.1)\qquad \begin{matrix} a_{kn} \\ b_{kn} \end{matrix}(f) = \frac{2}{2n+1}\sum_{j=0}^{2n} f(x_{jn}) \begin{matrix} \cos \\ \sin \end{matrix} kx_{jn} \quad (k=1, 2, \ldots, n),$$

$$(3.2)\qquad c_{kn}(f) = \frac{1}{2n+1}\sum_{j=0}^{2n} f(x_{jn})e^{-ikx_{jn}} \quad (|k|=0, 1, 2, \ldots, n).$$

Лемма 3.1. *Пусть R неотрицательное целое число. Для комплексных коэффициентов Фурье—Лагранжа $c_{kn}(f)$ функции $f(x)$ справедливо следующее представление:*

$$(3.3)\qquad c_{kn}(f) = \frac{(-1)^R e^{-ikRh}}{(e^{-ikh}-1)^R}c_{kn}(\Delta^R f) \qquad (|k|=1, 2, \ldots, n),$$

где $h=\dfrac{2\pi}{2n+1}$ *и* $\Delta^R f(x) = \Delta_h^R f(x)$.

Доказательство. Очевидно, что формула (3.3) верна при $R=0$. Пусть теперь $R\geqq 1$, и предположим, что (3.3) верна для значения $R-1$. Ясно, что

$$c_{kn}(\Delta^{R-1}f) = \frac{1}{2n+1}\sum_{j=0}^{2n}\Delta^{R-1}f(x_j)e^{-ikx_j} =$$

$$= \frac{1}{2n+1}\sum_{j=0}^{2n}[\Delta^{R-1}f(x_j)-\Delta^{R-1}f(x_{j+1})]\sum_{l=0}^{j}e^{-ikx_l} =$$

$$= \frac{-1}{2n+1}\sum_{j=0}^{2n}\Delta^{R}f(x_j)\frac{e^{-ikx_{j+1}}-1}{e^{-ikh}-1} =$$

$$= \frac{-e^{-ikh}}{(e^{-ikh}-1)}c_{kn}(\Delta^{R}f)+\frac{1}{(2n+1)(e^{-ikh}-1)}\sum_{j=0}^{2n}\Delta^{R}f(x_j).$$

Здесь второе слагаемое равняется нулю, так как $R\geqq 1$. Значит в силу индуктивного предположения Лемма 3.1. доказана.

Теорема 3.2. *При всех фиксированных натуральных R имеем следующую оценку для коэффициентов Фурье—Лагранжа $c_{kn}(f)$ функции $f(x)$:*

$$|c_{kn}(f)| \leqq (4k)^{-R}(2n+1)^{R-1}\sum_{j=0}^{2n}|\Delta^{R}f(x_j)| \qquad (|k|=1, 2, \ldots, n), \tag{3.4}$$

в частности, если функция $f(x)$ имеет ограниченное $(R, 1)$ изменение $V(R)$ в узлах (1.1), то

$$|c_{kn}(f)| \leqq V(R)\cdot\frac{n^{R-1}}{k^{R}}. \tag{3.5}$$

Теорема 3.2 есть непосредственно следствие формул (3.3) и (3.2).

Лемма 3.3. *Если R натуральное число, то*

$$\sum_{k=-n}^{n}[\Delta^{R}f(x_k)]^2 = 2^{2R-1}(2n+1)\sum_{k=1}^{n}(\varrho_{kn})^2\sin^{2R}\frac{k\pi}{2n+1} \qquad (n=1, 2, \ldots), \tag{3.6}$$

где $\varrho_{kn} = \sqrt{[a_{kn}(f)]^2+[b_{kn}(f)]^2}$.

Доказательство. Так как $\tilde{S}_n(x, f)$ интерполирует функцию в узлах (1.1), то

$$\sum_{k=-n}^{n}[\Delta^{R}f(x_k)]^2 = \sum_{k=-n}^{n}[\Delta^{R}\tilde{S}_n(x_k, f)]^2. \tag{3.7}$$

Ясно, что $\Delta^{R}\tilde{S}_n(x, f)$ — как и $\tilde{S}_n(x, f)$ — есть тригонометрической полином порядка не выше n, значит в силу равенства Парсеваля

$$\sum_{k=-n}^{n}[\Delta^{R}\tilde{S}_n(x_k, f)]^2 = (2n+1)\sum_{k=-n}^{n}|c_{kn}[\Delta^{R}\tilde{S}_n(x, f)]|^2. \tag{3.8}$$

(см. А. Зигмунд [10], X (2.8)). Отсюда получим, что для доказательства (3.6) нам нужно лишь подсчитать

$|c_{kn}[\Delta^R \tilde{S}_n(x,f)]|^2 \quad (|k| \leqq n)$. Так как

$$\tilde{S}_n(x,f) = \sum_{k=-n}^{n} c_{kn}(f) e^{ikx},$$

то

$$\Delta^R \tilde{S}_n(x,f) = \sum_{k=-n}^{n} c_{kn}(f) \Delta^R e^{ikx} = \sum_{k=-n}^{n} c_{kn}(f)(e^{ikh}-1)^R e^{ikx},$$

значит

$$|c_{kn}[\Delta^R \tilde{S}_n(x,f)]|^2 = |c_{kn}(f)|^2 \cdot 2^{2R} \sin^{2R} \frac{k\pi}{2n+1} =$$

(3.9)

$$= \frac{1}{4}(\varrho_{|k|,n})^2 2^{2R} \sin^{2R} \frac{k\pi}{2n+1}.$$

Теперь (3.6) вытекает из формул (3.7)—(3.9).

Отметим, что для случая $R=1$ эта лемма доказана в работе Д. Л. Берман [1].

Теперь мы уже в состоянии доказать следующую теорему типа Винера, частный случай ($R=1$) которого Читатель может найти в выше цитированной статье Д. Л. Бермана.

Теорема 3.4. *Пусть функция $f(x)$ не имеет разывов второго рода и для всех x значение $f(x)$ заключено между $f(x-0)$ и $f(x+0)$. Если $f(x)$ есть функция ограниченной $(R, 1)$ $(R=1, 2, \ldots)$ вариации в узлах (1.1), то для того чтобы она была непрерывна, необходимо и достаточно выполнение условия*

$$\liminf_{n\to\infty} \frac{1}{n^R} \sum_{k=1}^{n} k^R \varrho_{kn} = 0, \tag{3.10}$$

где $\varrho_{kn} = \sqrt{[a_{kn}(f)]^2 + [b_{kn}(f)]^2}$.

Доказательство. Если функция $f(x)$ непрерывна, то по Лемме 3.3*

$$\frac{1}{n^R} \sum_{k=1}^{n} k^R \varrho_{kn} \leqq C_8 \left(\sum_{k=-n}^{n} [\Delta^R f(x_k)]^2 \right)^{\frac{1}{2}} \leqq$$

$$\leqq C_8 \left[\omega_R \left(f; \frac{2\pi}{2n+1} \right) \cdot V(f, R, 1, \{x_k\}) \right]^{\frac{1}{2}} = o(1).$$

Пусть теперь выполнено (3.10) и $\{n_i\}$ такая строго возрастающая последовательность натуральных чисел, что

$$\lim_{i\to\infty} \frac{1}{n_i^R} \sum_{k=1}^{n_i} k^R \varrho_{kn_i} = 0.$$

*C_8, C_9 и C_{10} зависят от R.

Тогда по Лемме 3.3 и Теореме 3.2

$$\sum_{k=-n_i}^{n_i} [\Delta^R f(x_k)]^2 \leqq C_9 n_i \sum_{k=1}^{n_i} (\varrho_{kn_i})^2 \frac{k^{2R}}{n_i^{2R}} \leqq$$

(3.11)

$$\leqq C_{10} \frac{1}{n_i^R} \sum_{k=1}^{n_i} \varrho_{kn_i} k^R = o(1),$$

Но если функция $f(x)$ разрывна с разрывами первого рода, то левая часть в (3.11) не может стремиться к нулю при $i \to \infty$. Теорема доказана.

§ 4. О суммах Фурье

Пусть $\alpha = \{\alpha_1, \alpha_2, \ldots\}$ последовательность неотрицательных вещественных чисел и обозначим через

$$BR_n(x, f, \alpha) = \frac{1}{2}[S_n(x, f) + S_n(x + \alpha_n, f)] \qquad (n = 1, 2, \ldots)$$

суммы Бернштейна—Рогозинского, индуцированные при помощи α, далее $L_n(\alpha)$ — константы Лебега процесса $BR_n(x, f, \alpha)$. Например, если $\alpha_n = \dfrac{2\pi}{2n+1}$ $(n = 1, 2, \ldots)$, то $L_n(\alpha) \leqq \dfrac{2}{\pi} \int_0^{\pi} \dfrac{\sin x}{x} dx$ (см. С. Н. Бернштейн [2], стр. 525, формула (10)). В статье С. Б. Стечкина [9] можно найти условия, при которых $\sup_n L_n(\alpha) < \infty$.

Лемма 4.1. *Пусть R $(\geqq 0)$ целое, δ вещественное число. Тогда*

$$g(x) \equiv (-2)^{-R} \Delta_\delta^R g(x) + \frac{1}{2}\left[\sum_{i=0}^{R-1} (-2)^{-i} \Delta_\delta^i g(x) + \sum_{i=0}^{R-1} (-2)^{-i} \Delta_\delta^i g(x + \delta)\right] \tag{4.1}$$

для любой ограниченной функции $g(x)$.

Тождество (4.1) легко доказать с помощью индукции по числу R (см. Г. П. Неваи [7] (2.28)).

Теорема 4.2. *Для любой последовательности α и для любого неотрицательного целого числа R имеем*

$$|S_n(x, f) - f(x) - (-2)^{-R} S_n(x, \Delta_{\alpha_n}^R f)| \leqq$$

$$\leqq 2^{-R} \omega_R(f; \alpha_n) + (R \cdot L_n(\alpha) + 2) E_n(f) \tag{4.2}$$

$(n = 1, 2, \ldots)$, *в частности R может стремиться к бесконечности вместе с n.*

Следствие 4.3. *С помощью Леммы 4.1 оценку (4.2) можно перефразировать так, что*

$$\left| BR_n\left(x, \sum_{i=1}^{R-1} (-2)^{-i} \Delta_{\alpha_n}^i f, \alpha\right) - f(x) \right| \leqq 2^{-R} \omega_R(f; \alpha_n) + (R \cdot L_n(\alpha) + 2) E_n(f).$$

Следствие 4.4. *Для всех* α *и* R

$$|f(x)-S_n(x,f)| \leqq 2^{-R}(L_n+1)\omega_R(f;\alpha_n)+ \\ +(R\cdot L_n(\alpha)+2)E_n(f) \qquad (n=1,2,\ldots), \tag{4.3}$$

где $L_n \equiv L_n(\{0\}) \approx \frac{4}{\pi^2}\log n$ *константы Лебега сумм Фурье.*

В. В. Жук в работе [11] другим методом указал аналогичную к (4.3) оценку для случая $\alpha_n = \frac{2\pi}{2n+1}$ $(n=1,2,\ldots)$.

О. Киш [4] показал справедливость неравенства

$$|f(x)-S_n(x,f)| \leqq \left(\frac{1}{2}L_n+3\right)\omega_1\left(f;\frac{2\pi}{2n+1}\right) \qquad (n=1,2,\ldots).$$

Доказательство Теоремы 4.2. В силу Леммы 4.1

$$S_n(x,f)-f(x) = S_n(x,f-H_n(f))-[f(x)-H_n(x,f)] = \\ = (-2)^{-R}S_n(x,\Delta_{\alpha_n}^R f)-(-2)^{-R}S_n(x,\Delta_{\alpha_n}^R H_n(f))+ \\ +BR_n\left(x,\sum_{i=0}^{R-1}(-2)^{-i}\Delta_{\alpha_n}^i[f-H_n(f)],\alpha\right)-[f(x)-H_n(x,f)] = A+B+C+D.$$

Далее, так как $\Delta_{\alpha_n}^R H_n(x,f)$ есть тригониметрический полином порядка не выше n, то

$$|B| = |-(-2)^{-R}\Delta_{\alpha_n}^R H_n(x,f)| \leqq 2^{-R}\omega_R(H_n(f);\alpha_n) \leqq \\ \leqq 2^{-R}\omega_R(f-H_n(f);\alpha_n)+2^{-R}\omega_R(f;\alpha_n) \leqq E_n(f)+2^{-R}\omega_R(f;\alpha_n).$$

Очевидно, что

$$|C| \leqq L_n(\alpha)\max_{|x|\leqq\pi}\left|\sum_{i=0}^{R-1}(-2)^{-i}\Delta_{\alpha_n}^i[f(x)-H_n(x,f)]\right| \leqq \\ \leqq L_n(\alpha)\sum_{i=0}^{R-1}2^{-i}\omega_i(f-H_n(f);\alpha_n) \leqq R\cdot L_n(\alpha)\cdot E_n(f).$$

Таким образом

$$|B+C+D| \leqq 2^{-R}\omega_R(f;\alpha_n)+(R\cdot L_n(\alpha)+2)E_n(f),$$

что и требовалось доказать.

Результаты настоящего пункта — как видно из доказательства — остаются в силе, если $S_n(x,f)$ заменить на любой линейный метод суммирования рядов Фурье типа Валле—Пуссена $U_n(x,f)$, то есть на такой метод, где $U_n(x,f)$ есть линейная комбинация сумм $S_k(x,f)$ с индексами $k \geqq n$.

§ 5. Еще раз о сходимости тригонометрической интерполяции

В этом пункте рассмотрим еще одно применение формулы (2.4).

Пусть A некоторое нормированное пространство, элементами которого являются все финитные последовательности вещественных чисел вида $a=\{\ldots, 0, a_K, \ldots, a_{-1}, a_0, a_1, \ldots, a_l, 0, \ldots\}$. Пусть далее на A задана последовательность линейных функционалов φ_i $(i=1, 2, \ldots)$, действующие по правилу

$$\varphi_i(a) = \sum_{|k|\geqq i} \frac{a_k}{|k|}.$$

Пусть норма в A определена так, что для норм φ_i в сопряженном к A пространстве имеем

$$\|\varphi_i\| \underset{i\to\infty}{\to} 0.$$

Обозначим через $\Delta_n(R, f, \alpha)$ (R и n натуральные, α целое числа) последовательность

$$\{\ldots, 0, \Delta^R f(x_{\alpha-n,n}), \ldots, \Delta^R f(x_{\alpha n}), \ldots, \Delta^R f(x_{\alpha+n,n}), 0, \ldots\}.$$

Теорема 5.1. *Пусть $\alpha=\alpha(n)$ тот индекс, для которого $x_\alpha \leqq x < x_{\alpha+1}$. Если функция $f(x)$ непрерывна и для некоторого натурального R нормы $\Delta_n(R, f, \alpha)$ в пространстве A в совокупности ограничены, то в точке x интерполяционные полиномы $\tilde{S}_n(x, f)$ сходятся к значению $f(x)$ при $n\to\infty$. Если же $\|\Delta_n(R, f, \alpha)\|$ ограничены равномерно по $x\in[-\pi, \pi]$, то и сходимость будет равномерной на оси.*

Доказательство этого утверждения весьма просто, ибо по (2.4)*

$$|f(x)-\tilde{S}_n(x, f)| \leqq C_{11} \sum_{|k|=1}^{n} \frac{|\Delta^R f(x_{k+\alpha})|}{|k|} + o(1) \leqq$$

$$\leqq C_{12} \log N \cdot \omega_R\left(f; \frac{1}{n}\right) + C_{11}\varphi_{N+1}(\Delta_n(R, f, \alpha)) + o(1) \leqq \tag{5.1}$$

$$\leqq C_{12} \log N \cdot \omega_R\left(f; \frac{1}{n}\right) + C_{11}\|\Delta_n(R, f, \alpha)\| \cdot \|\varphi_{N+1}\| + o(1)$$

и если N достаточно медленно стремится к бесконечности вместе с n, то все три выражения в правой части (5.1) будут стремиться к нулю при $n\to\infty$.

Конкретные реализации пространства A предостовляем Читателю.

ЛИТЕРАТУРА

[1] Берман, Д. Л.: Об одном интерполяционном аналоге критерия Винера непрерывности функции ограниченной вариации, *Доклады АН СССР*, **196** (1971), 495—497.

[2] Бернштейн, С. Н.: *Содрание сочинений*, том I, Москва, 1952.

---

* c_{11} и c_{12} абсолютные постоянные.

[3] Харшиладзе, Ф. И.: О функциях с ограниченным вторым изменением, *Доклады АН СССР*, **79** (1951), 201—204.
[4] Киш, О.: Оценка отклонения частных сумм ряда Фурье, *Acta Math. Acad. Sci. Hungar.*, **22** (1971), 173—176.
[5] Неваи, Г. П.: Об отклонении тригонометрических интерполяционных сумм, *Acta Math. Acad. Sci. Hungar.*, **23** (1972), 203—205.
[6] Неваи, Г. П.: Асимптотическая формула для отлонения тригонометрических интерполяционных сумм, *Studia Sci. Math. Hungar.* **7** (1972), 395—398.
[7] Névai, G. P.: Az ekvidisztáns csomópontokon alapuló trigonometrikus interpolációról. (О тригонометрическом интерполировании по равноотстоящим узлам), *Magyar Tud. Akad. Mat. és Fiz. Oszt. Közleményei*, **21** (1973), 449—484.
[8] Salem, R.: Essais sur les séries trigonométriques, *Actual. Sci. et Industr.*, № 862, Paris, 1940.
[9] Стечкин, С. Б.: Методы суммирования С. Н. Бернштейна и В. Рогозинского, в книге *Г. Х. Харди „Расходящиеся ряды"*, Москва, 1951, 479—492.
[10] Зигмунд, А.: *Тригонометрические ряды*, том II, Москва, 1965.
[11] Жук, В. В.: О точности представления 2π-периодической функции при помощи линейных методов суммирования, *Известия ВУЗ-ов*, серия математика, **8** (1972), 46—59.

Математический Институт Венгерской Академии Наук, Будапешт

(Поступила: 26-ого мая 1972. г.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 123—124.

ON THE COEFFICIENTS OF POLYNOMIALS WITH NEGATIVE ZEROS

by

G. J. SZÉKELY

In this paper we shall prove that the distribution of the coefficients of polynomials $P_n(x)$ with real negative zeros $-r_{nj}$ $(j=1, 2, \ldots, n)$ is asymptotically normal with mean value

$$M_n = \sum_{j=1}^{n} \frac{1}{1+r_{nj}} \quad \text{and variance} \quad D_n^2 = \sum_{j=1}^{n} \frac{r_{nj}}{(1+r_{nj})^2} \quad \text{if} \quad \lim_{n\to+\infty} D_n = +\infty.$$

E.g. if $\{P_n(x)\}_{n=1}^{\infty}$ is an arbitrary system of orthogonal polynomials defined on a negative interval then the coefficients of $P_n(x)$ are asymptotically normal.

THEOREM. *If* $r_{nj}(n=1, 2, \ldots; j=1, 2, \ldots, n)$ *are non-negative numbers,*

$$P_n(x) = \prod_{j=1}^{n} (x+r_{nj}) = \sum_{k=0}^{n} a_{nk} x^k \quad \textit{and} \quad \lim_{n\to+\infty} \sum_{j=1}^{n} \frac{r_{nj}}{(1+r_{nj})^2} = +\infty, \quad \textit{then}$$

$$\lim_{n\to\infty} c_n^{-1} \sum_{k=0}^{x_n} a_{nk} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt,$$

where $$x_n = x\sqrt{\sum_{j=1}^{n} \frac{r_{nj}}{(1+r_{nj})^2}} + \sum_{j=1}^{n} \frac{1}{1+r_{nj}} \quad \textit{and} \quad c_n = \prod_{j=1}^{n} (1+r_{nj}).$$

PROOF. It is clear from the definition of $P_n(x)$ that $c_n^{-1} P_n(x)$ is the generator function of a discrete random variable $\eta_n \left(\mathsf{P}(\eta_n=k) = \frac{a_{nk}}{c_n}\right)$ which is the sum of n independent random variables $\xi_{nj} \left(\mathsf{P}(\xi_{nj}=0) = \frac{r_{nj}}{1+r_{nj}}, \ \mathsf{P}(\xi_{nj}=1) = \frac{1}{1+r_{nj}}\right)$. It is obvious that $\mathsf{M}\eta_n = \sum_{j=1}^{n} \frac{1}{1+r_{nj}}$ and $\mathsf{D}^2\eta_n = \sum_{j=1}^{n} \frac{r_{nj}}{(1+r_{nj})^2}$. $\lim_{n\to+\infty} \sum_{j=1}^{n} \frac{r_{nj}}{(1+r_{nj})^2} = +\infty$, therefore — denoting $\frac{\xi_{nj} - (1+r_{nj})^{-1}}{\mathsf{D}\eta_n}$ by ξ_{nj}^* — $|\xi_{nj}^*| < \varepsilon$ a.e. if $n > N(\varepsilon)$ thus $\lim_{n\to+\infty} \sum_{j=1}^{n} \int_{|\xi_{nj}^*|>\varepsilon} x^2 \, dF_{nj} = 0$ where F_{nj} is the distribution function of ξ_{nj}^* i.e. the Lindeberg condition holds and because of ((1), p. 295.) the Theorem is true.

Let us consider an arbitrary real-valued non-decreasing bounded function $\alpha(x)$ defined on an interval $[a, b]$ $(-\infty < a < b < +\infty)$ with the property that it takes infinitely many different values and $\int_a^b x^n d\alpha(x)$ is convergent for $n=0, 1, 2, \ldots$. If

$\{P_n(x)\}_{n=0}^{\infty}$ is the system of orthogonal polynomials with respect to $\alpha(x)$ then all zeros of $P_n(x)$ lie in (a, b) (see (2) Theorem 2. 2).

COROLLARY: *If* $-\infty<a<b<0$ *and* $\{P_n(x)\}_{n=1}^{\infty}$ *is the system of orthogonal polynomials with respect to* $\alpha(x)$ *then the coefficients of* $P_n(x)$ *are asymptotically normal with mean value* M_n *and variance* D_n^2.

PROOF. Let us denote by $-r_{nj}$ the roots of $P_n(x)$ $(j=1, 2, \ldots, n)$. $-r_{nj}\in(a, b)$ thus $r_{nj}>0$ and $\lim\limits_{n\to\infty}\sum\limits_{j=1}^{n}\frac{r_{nj}}{(1+r_{nj})^2}\geqq\lim\limits_{n\to\infty}\frac{-bn}{(1-a)^2}=+\infty$ therefore the conditions of the Theorem are fulfilled.

REFERENCES

[1] LOÈVE, M.: *Probability Theory* (Second Edition), Van Nostrand, Princeton, 1960.
[2] FREUD, G.: *Orthogonal Polynomials,* Akadémia Kiadó, Budapest, 1971.

L. Eötvös University, Budapest

(Received: September 1, 1972.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 125—132.

ASYMPTOTIC BEHAVIOUR OF SOLUTIONS OF UNSTABLE-TYPE FIRST-ORDER DIFFERENTIAL EQUATIONS WITH DELAY

by

I. GYŐRI

1. Introduction

Solving a problem of R. BELLMAN [1], COOKE [3] proved that if $0 \leqq \sigma(x) \to 0$ $(A \leqq x \to +\infty)$ and $\int_A^\infty \sigma(x)dx < \infty$ then, for any solution of the equation

$$y'(x) = Ry(x-\sigma(x)) \tag{1.1}$$

(R is a constant)

which is continuous on the interval $A \leqq x < \infty$, there exists

$$\lim_{x \to +\infty} y(x)e^{-Rx} = c \tag{1.2}$$

with a suitable constant c. Conversely, for every real constant c there exists a solution of (1. 1) such that (1. 2) holds.

As a main result of the present paper, we extend COOKE's result to the unstable-type equation

$$y'(x) = \int_0^\infty y(x-s)dr(x, s) \tag{1.3}$$

where we assume that the kernel $r(x, s)$ satisfies the following conditions (see MÝSKIS [4]):

(i) the function $r(x, s)$ is defined in the domain $A \leqq x < \infty$, $0 \leqq s < \infty$, and is non-decreasing in s for any admissible fixed value of x;

(ii) $r(x, 0) = 0$ $(A \leqq x < \infty)$;

(iii) there are given two nonnegative continuous functions $\sigma(x) \geqq 0$ and $R(x) \geqq 0$ for which

$$r(x, s) = R(x)$$

holds whenever $A \leqq x < \infty$ and $\sigma(x) \leqq s < \infty$;

(iv) for any x with $A \leqq x < \infty$ we have

$$\lim_{\substack{t \to x \\ A \leqq t < \infty}} \int_0^{\sigma(x)+1} |r(t, s) - r(x, s)|\, ds = 0$$

(v) there exists an x_1, $A \leqq x_1 < \infty$, such that in case $x_1 \leqq x < \infty$ the inequality,

$$x - \sigma(x) \geqq A$$

holds.

From condition (v) it is obvious that

$$a = \inf \{x - \sigma(x)\} > -\infty \tag{1.4}$$

Assume that $\varphi(x)$ is a continuous function defined in the interval $[a, A]$. Call a function $y(x)$ a solution of the equation (1.3) associated with $\varphi(x)$ if it coincides with $\varphi(x)$ on $[a, A]$, is continuous and differentiable on the interval $[A, \infty)$, and satisfies equation (1. 3) on $[A, \infty)$. Under the assumptions imposed on $r(x, s)$, MÝSKIS [4] established the existence and uniqueness of the solution of equation (1.3) associated with any given $\varphi(x)$.

Now we are in a position to state our main theorem:

THEOREM 1. *Assume the functions* $r(x, s)$, $\sigma(x)$, *and* $R(x)$ *satisfy* (i) *to* (iv). *Then for any solution* $y(x)$ *of equation* (1.3), *we have*

$$\lim_{x \to +\infty} y(x) e^{-\int_A^x R(t)dt} = c, \tag{1.5}$$

with a suitable constant c.

Here c cannot be always arbitrary. Namely, we have:

THEOREM 2. *Suppose that the assumptions of Theorem 1, hold, and further, that we have*

(a) $$\lim_{x \to +\infty} x - \sigma(x) = +\infty,$$

(b) $$\int_A^\infty R(t)\,dt = +\infty$$

and

(c) $$e^{-\int_A^x R(t)dt} \cdot \int_{x_1}^x \left\{ \int_0^\infty e^{\int_A^{t-s} R(\tau)d\tau} dr(t, s) \right\} dt \leqq d,$$

where x_1 *and* d *are constants,* $x_1 \leqq x < \infty$ *and* $0 < d < 1$.

Then we have

$$\lim_{x \to +\infty} y(x) e^{-\int_A^x R(t)dt} = 0, \tag{1.6}$$

for any solution $y(x)$ *of the equation* (1.3).

Our next result confirms that in many cases c assumes every value:

THEOREM 3. *If in addition to the assumptions of Theorem 1. we have*

$$\int_{x_1}^\infty \left\{ \int_0^\infty \left[1 - e^{-\int_{x-s}^x R(t)\,dt} \right] dr(x, s) \right\} dx < \infty, \tag{1.7}$$

then for any constant c *there exists a solution of the equation* (1.3) *which satisfies* (1.5).

These results will be proved in Section 3.

2. Some corollaries

Assume that $r(x, s)$ is the function given above and

$$\int_{x_1}^{\infty} R(x)dx < \infty. \tag{2.1}$$

Then condition (1.7) holds obviously, and so Theorems 1. and 3. ensure:

COROLLARY 1. *If $r(x, s)$ is the function of Theorem 1. and $R(x)$ satisfies* (2. 1), *then any solution $y(x)$ of the equation* (1. 3) *satisfies*

$$\lim_{x \to +\infty} y(x) = c, \tag{2.2}$$

with a suitable constant c. Conversely, for any constant c there exists a solution of the equation (1.3) *such that* (2.2) *holds.*

Now put

$$r(x, s) = R_0(x)e(s) + \sum_{i=1}^{n} R_i(x)e\big(s - \sigma_i(x)\big), \tag{2.3}$$

where $R_i(x) \geqq 0$ $(i=0, 1, 2, \ldots, n)$ and $\sigma_i(x) \geqq 0$ $(i=1, \ldots, n)$ are continuous functions on the interval $[A, \infty)$, and

$$e(s) = \begin{cases} 0 & \text{if} \quad s \leqq 0 \\ 1 & \text{if} \quad s > 0. \end{cases}$$

Assume that there exists a number x_1 with $A \leqq x_1 < \infty$ such that

$$x - \sigma_i(x) \geqq A \qquad (i = 1, 2, \ldots, n)$$

holds whenever $x_1 \leqq x < \infty$. Then, as is easily seen, the kernel $r(x, s)$ satisfies the assumptions (i) to (v), and the equation (1.3) goes over into the equation

$$y'(x) = R_0(x)y(x) + \sum_{i=1}^{n} R_i(x)y\big(x - \sigma_i(x)\big). \tag{2.4}$$

Furthermore, in order that the condition (1. 7) should be satisfied it is sufficient that

$$\int_{A}^{\infty} R_i(x)dx < \infty \quad (i=1, 2, \ldots, n) \tag{2.5}$$

holds. In fact, we have

$$\int_{x_1}^{\infty} \left\{ \int_{0}^{\infty} \left[1 - e^{-\int_{x-s}^{x} R(t)\,dt}\right] dr(x, s) \right\} dx =$$

$$= \int_{x_1}^{\infty} \left\{ \sum_{i=1}^{n} R_i(x) \left[1 - e^{-\int_{x-\sigma_i(x)}^{x} R(t)\,dt}\right] \right\} dx \leqq \sum_{i=1}^{n} \int_{x_1}^{\infty} R_i(x)dx < \infty,$$

where

$$R(x) = \sum_{j=1}^{n} R_j(x).$$

So, from Theorems 1. and 3. we can deduce:

COROLLARY 2. *If $\sigma_i(x)$ and $R_i(x)$ are the functions described above, then any solution $y(x)$ of the equation* (2. 4) *satisfies*

$$\lim_{x\to+\infty} y(x)e^{-\int_A^x R_0(t)\,dt} = c \tag{2.6}$$

with a suitable constant c. Conversely, for any constant c there exists a solution $y(x)$ of the equation (2. 4) *satisfying* (2. 6).

Finally, set

$$r(x, s) = R(x)\,e(s-\sigma(x)) \tag{2.7}$$

where $R(x)\geqq 0$ and $\sigma(x)\geqq 0$ are continuous functions on the interval $[A, \infty)$. Assume that for some x_1 with $A\leqq x_1<\infty$ we have

$$x-\sigma(x) \geqq A \tag{2.8}$$

whenever $x_1\leqq x<\infty$. Then $r(x, s)$ satisfies assumptions (i) to (v), and so Theorems 1. and 3. imply:

COROLLARY 3. *Assume $R(x)\geqq 0$ and $\sigma(x)\geqq 0$ are continuous functions such that*

$$\int_{x_1}^{\infty} R(x)\left[1-e^{-\int_{x-\sigma(x)}^{x} R(t)\,dt}\right]dx < \infty \tag{2.9}$$

holds. Then any solution $y(x)$ of the equation

$$y'(x) = R(x)y(x-\sigma(x)) \tag{2.10}$$

satisfies

$$\lim_{x\to+\infty} y(x)\cdot e^{-\int_A^x R(t)\,dt} = c \tag{2.11}$$

with a suitable constant c. Moreover, for any constant c there exists a solution $y(x)$ of (2.10) *satisfying* (2.11).

Choose $r(x, s)$ the function given under (2. 7) and assume that

$$\lim_{x\to+\infty} x-\sigma(x) = +\infty \tag{2.12}$$

and

$$\int_A^{\infty} R(x)\,dx = \infty. \tag{2.13}$$

Assume, furthermore, that

$$\int_{x-\sigma(x)}^{x} R(s)\,ds \geqq d_1 \tag{2.14}$$

where d_1 is a constant with $0<d_1<\infty$, and $x_1\leqq x<\infty$.

Then $r(x, s)$ satisfies the assumptions of Theorem 2. Indeed, assumptions (a) and (b) hold obviously, and assumption (c) can be verified as follows:

Supposing $x_1 \leqq x < \infty$, by the definition of $r(x, s)$ we obtain

$$e^{-\int_A^x R(t)\,dt} \cdot \int_{x_1}^{x} \left\{ \int_0^{\infty} e^{\int_A^{t-s} R(\tau)\,d\tau}\, dr(t, s) \right\} dt =$$

$$= e^{-\int_A^x R(t)\,dt} \cdot \int_{x_1}^{x} R(t) e^{\int_A^{-\sigma(t)} R(\tau)\,d\tau}\, dt =$$

$$= e^{-\int_A^x R(t)\,dt} \int_{x_1}^{x} R(t) e^{\int_A^t R(\tau)\,d\tau} \cdot e^{-\int_{t-\sigma(t)}^t R(\tau)\,d\tau}\, dt \leqq e^{-d_1} < 1,$$

which is exactly what we wanted to show.

So we get

COROLLARY 4. *Assume that* $R(x) \geqq 0$ *and* $\sigma(x) \geqq 0$ *are continuous functions in the interval* $[A, \infty)$, *and that* (2.12), (2.13) *and* (2.14) *are satisfied. Then any solution* $y(x)$ *of the equation* (2. 10) *satisfies*

$$\lim_{x \to +\infty} y(x) e^{-\int_A^x R(t)\,dt} = 0.$$

3. Proofs of the theorems

PROOF of Theorem 1. Suppose $y(x)$ is an arbitrary continuous solution of the equation (1. 3). Then there exists a function $\xi(x)$ such that

$$0 \leqq \xi(x) \leqq \sigma(x) \tag{3.1}$$

and

$$\int_0^{\infty} y(x-s)\,dr(x, s) = y\big(x - \xi(x)\big) \cdot R(x), \tag{3.2}$$

indeed, we stipulated that $r(x, s)$ is nondecreasing in s for any admissible fixed x.

So, $y(x)$ satisfies the equation

$$y'(x) = R(x) y\big(x - \xi(x)\big) \tag{3.3}$$

i.e. we have

$$y'(x) - R(x) y(x) = R(x)\big[y\big(x - \xi(x)\big) - y(x)\big] \tag{3.4}$$

From this latter equation we obtain that

$$y(x) e^{-\int_{x_1}^x R(t)\,dt} = y(x_1) - \int_{x_1}^{x} e^{-\int_A^t R(s)\,ds} \cdot R(t)\big[y(t) - y\big(t - \xi(t)\big)\big]\,dt. \tag{3.5}$$

Assume that $\varphi(x) \geqq 0$ is a nonnegative continuous function on the interval $a \leqq x \leqq A$. Then a theorem of MŸSKIS [4] says that the solution $y(x)$ associated with

$\varphi(x)$ is nondecreasing. Making use of this fact, it is clear from (3. 5) that the function

$$y(x)e^{-\int_{x_1}^{x} R(t)dt}$$

is nonincreasing for $x_1 \leqq x < \infty$ and that

$$0 \leqq y(x)e^{-\int_{x_1}^{x} R(t)\,dt} \leqq y(x_1); \tag{3.6}$$

so the limit (1.5) indeed exists for this solution $y(x)$.

Assume now that $\varphi(x)$ is an arbitrary continuous function, and put

$$\varphi_+(x) = \max\{0, \varphi(x)\}, \quad \varphi_-(x) = \max\{0, -\varphi(x)\}$$

then

$$\varphi(x) = \varphi_+(x) - \varphi_-(x)$$

If the solution $y(x)$ associated with $\varphi(x)$ and if solutions $y_+(x)$ and $y_-(x)$ associated with $\varphi_+(x)$ and $\varphi_-(x)$, then obviously that

$$y(x) = y_+(x) - y_-(x).$$

As mentioned above, for $y_+(x)$ and $y_-(x)$ the limit under (1.5) exists so, evidently, the limit under (1.5) exists for $y(x)$. The proof is complete.

PROOF of Theorem 2. Let $y(x)$ be a solution of the equation (1. 3), and choose a positive ε. Then, in view Theorem 1, there exists a constant c such that

$$\lim_{x \to +\infty} y(x)e^{-\int_{A}^{x} R(t)\,dt} = c; \tag{3.7}$$

i.e. with a large enough $x_2 \geqq x_1$ we have

$$c - \varepsilon \leqq y(x)e^{-\int_{A}^{x} R(t)\,dt} \leqq c + \varepsilon \tag{3.8}$$

whenever $x_2 \leqq x < \infty$. Since we assumed

$$\lim_{x \to +\infty} x - \sigma(x) = +\infty,$$

there exist an $x_3 \geqq x_2$ such that

$$c - \varepsilon \leqq y(x-s)e^{-\int_{A}^{x-s} R(t)\,dt} \leqq c + \varepsilon, \tag{3.9}$$

whenever $x_3 \leqq x < \infty$ and $0 \leqq s \leqq \sigma(x)$.

From equation (1. 3) we deduce that

$$y(x) = y(x_3) + \int_{x_3}^{x} \left[\int_{0}^{\infty} y(t-s)\,dr(t, s)\right] dt$$

i.e.

$$y(x)e^{-\int_A^x R(t)\,dt} = y(x_3)e^{-\int_A^x R(t)\,dt} +$$

$$+e^{-\int_A^x R(\tau)\,d\tau}\int_{x_3}^{x}\left[\int_0^{\infty} y(t-s)e^{-\int_A^{t-s} R(\tau)\,d\tau}e^{\int_A^{t-s} R(\tau)\,d\tau}\,dr(t,s)\right]dt.$$

Now put $c>0$, then from this latter equality, (3. 9), and condition (c) of Theorem 2 we obtain

$$0 < y(x)e^{-\int_A^x R(\tau)\,d\tau} \leqq y(x_3)e^{-\int_A^x R(\tau)\,d\tau} + (c+\varepsilon)\cdot d$$

$$(x_3 \leqq x < \infty)$$

Making $x \to +\infty$ and using condition (b) of Theorem 2, we get that

$$0 < c \leqq (c+\varepsilon)\cdot d$$

and so

$$0 < c \leqq \frac{d}{1-d}\cdot\varepsilon.$$

Since $0<d<1$ is fixed and $\varepsilon>0$ is arbitrary, the assertion of Theorem 2 follows. If $c<0$, then the proof is analogous. The proof is complete.

PROOF of Theorem 3. This theorem is an obvious consequence of the following lemma:

LEMMA. *Suppose $r(x, s)$ satisfies assumptions* (i) *to* (v) *and $\varphi(x)\equiv 0$ $(\varphi(x)\not\equiv 0)$ is a continuous function in the interval $[a, A]$. Then the solution $y(x)$ of equation* (1.3) *associated with $\varphi(x)$ satisfies the inequality*

$$y(x)e^{-\int_A^x R(t)\,dt} \geqq k_1 e^{-\int_{x_1}^x \left\{\int_0^{\infty}\left[1-e^{-\int_{t-s}^{t} R(\tau)\,d\tau}\right]dr(t,s)\right\}dt}, \tag{3.10}$$

$$(x_1 \leqq x < \infty)$$

where k_1 is a positive constant.

Before embarking upon the proof of this lemma, we show how to derive Theorem 3 from it. To this end, assume that condition (1. 7) of Theorem 3 is satisfied. Then (3. 10) implies that

$$c = \lim_{x\to+\infty} y(x)e^{-\int_A^x R(t)\,dt} \geqq k_2$$

holds with a suitable constant k_2. If now we take $\varphi(x)\equiv 1$, then for the associated solution $y_1(x)$ we have

$$\lim_{x\to+\infty} y_1(x)e^{-\int_A^x R(t)\,dt} = c_1 > 0;$$

thus, for any constant c there exists a solution $y(x)$ of (1.3) such that

$$\lim_{x\to+\infty} y(x)e^{-\int_A^x R(t)\,dt} = c.$$

Indeed, such a solution is the one associated with the initial function $\varphi(x)=c/c_1$.

PROOF of the Lemma. From equation (1. 3) we obtain that $y(x)$ satisfies the equality

$$y(x)e^{-\int_A^x R(t)\,dt} = y(A)+\int_A^x e^{-\int_A^t R(s)\,ds}\Big[\int_0^\infty y(t-s)\,dr(t,s)-R(t)y(t)\Big]\,dt. \tag{3.11}$$

In the course of the proof of Theorem 1 we saw that, under the given conditions, the function

$$z(x) = y(x)e^{-\int_A^x R(t)dt} \tag{3.12}$$

$$(x_1 \leqq x < \infty)$$

is nonincreasing, i.e.

$$z(x-s) \geqq z(x) \geqq 0 \tag{3.13}$$

holds in the domain $x_1\leqq x<\infty$, $0\leqq s\leqq\sigma(x)$.

Assume now $x_1\leqq\bar{x}<x<\infty$. Then making use of (3.11) and (3.12) we obtain that

$$z(x) = z(\bar{x})-\int_{\bar{x}}^x\Big\{\int_0^\infty\big[z(t)-z(t-s)e^{-\int_{t-s}^t R(\tau)\,d\tau}\big]dr(t,s)\Big\}\,dt.$$

So, applying inequality (3. 13), we can derive that

$$z(x)\geqq z(\bar{x})-\int_{\bar{x}}^x\Big\{z(t)\int_0^\infty\big[1-e^{-\int_{t-s}^t R(\tau)\,d\tau}\big]dr(t,s)\Big\}\,dt,$$

where $\bar{x}$ and x ($x_1\leqq\bar{x}<x<\infty$) are arbitrary. If we put $k=z(x_1)$ and apply the inequality of LANGENHOP [2], (3.10) follows. This accomplishes the proof of the Lemma. So the proof of Theorem 3. is also complete.

REFERENCES

[1] BELLMAN, R.: Research problem: Functional differential equations, *Bull. Amer. Math. Soc.* **71** (1965), 495.

[2] LANGENHOP, C. E.: Bounds on the norm of a solution of general differential equation, *Proc. Amer. Math. Soc.* **11** (1960), 795—799.

[3] COOKE, K. L.: Functional differential equations close to differential equations, *Bull. Amer. Math. Soc.* **72** (1966), 285.

[4] MYSKIS, A. D.: Linear differential equations with delayed argument, *GPL* (1951), 71 (in Russian).

Central Research Laboratory, Medical University, Szeged, Hungary

(Received September 10, 1972.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 133—140.

VARIATIONAL PROBLEMS RELATED TO THOSE WHICH LEAD TO THE SCHRÖDINGER EQUATION OF HELIUM-LIKE ATOMS

by

E. MAKAI

1. Let $x_1, y_1, z_1, x_2, y_2, z_2$ be variables ranging from $-\infty$ to ∞, $r_i^2 = x_i^2 + y_i^2 + z_i^2$ $(i=1, 2)$, $r_{12}^2 = (x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2$,

$$\Delta_i = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \qquad (i = 1, 2)$$

and Z some constant. The Schrödinger equation of helium-like atoms is customarily written in the form

$$-\frac{1}{2}(\Delta_1\psi + \Delta_2\psi) + \left(-\frac{Z}{r_1} - \frac{Z}{r_2} + \frac{1}{r_{12}}\right)\psi = \lambda\psi \tag{1.1}$$

and its eigenfunctions are those solutions which for suitable λ's are quadratic integrable in the entire six dimensional Euclidean space $x_1, y_1, \ldots, z_2$.

In spite of long efforts no progress whatever has been achieved in finding the analytical form of the eigenfunctions of (1.1). We quote the classical results that the eigenfunction corresponding to the lowest eigenvalue is a function of r_1, r_2, r_{12} only, and it is a symmetric function of r_1 and r_2: $\psi|_{r_1, r_2, r_{12}} = \psi|_{r_2, r_1, r_{12}}$. A theorem of T. KATO (cf. [2]) states that this eigenfunction is continuous everywhere including the origin.

On the other hand numerous calculations have been made for the approximate determinations of the first (i.e. lowest) eigenvalue and the corresponding eigenfunction. These can be found by minimizing the quotient

$$\int_{-\infty}^{\infty}\!\!\cdots\!\int \left[\frac{1}{2}(\nabla w)^2 + \left(\frac{1}{r_{12}} - \frac{Z}{r_1} - \frac{Z}{r_2}\right) w^2\right] dx_1 \ldots dz_2 \Big/ \int_{-\infty}^{\infty}\!\!\cdots\!\int w^2 dx_1 \ldots dz_2 \tag{1.2}$$

where

$$(\nabla w)^2 = \left(\frac{\partial w}{\partial x_1}\right)^2 + \left(\frac{\partial w}{\partial y_1}\right)^2 + \cdots + \left(\frac{\partial w}{\partial z_2}\right)^2$$

and w is a square integrable function in the entire six dimensional space having there piecewise continuous first derivatives. To find an upper bound for the minimum of (1.2) one restricts the class of functions admitted to concurrence in the variational problem. Though it is known that the first eigenfunction cannot be expanded into a power series of the form

$$\sum_{l, m, n=0}^{\infty} a_{lmn} r_1^l r_2^m r_{12}^n \tag{1.3}$$

convergent in the whole space or even in a neighbourhood of the origin (see e.g. [2]),

very good results have been achieved when in treating the variational problem the class of admissible functions was of the form

$$e^{-ks}\sum_{l=0}^{l_1}\sum_{m=0}^{m_1}\sum_{n=0}^{n_1} b_{lmn}s^l t^{2m}u^n \tag{1.4}$$

where

$$s = r_1+r_2, \quad t = r_2-r_1, \quad u = r_{12} \tag{1.5}$$

and k, b_{lmn} were appropriate constants.

Considering a broader class of admissible functions, namely

$$e^{-ks}\sum_{(l,m,n)\in S} c_{lmn}s^l p^m q^{2n} \tag{1.6}$$

where S is some finite set of triplets (l, m, n) consisting of nonnegative integers,

$$p = \frac{r_{12}}{r_1+r_2} = \frac{u}{s}, \quad q = \frac{r_2-r_1}{r_{12}} = \frac{t}{u}, \tag{1.7}$$

T. KINOSHITA [2] found in the case of helium ($Z=2$) an upper bound for the lowest eigenvalue which is phenomenally near to the value derived from experiments.* In his calculations the set S contained 79 elements and no one of the integers l, m, n exceeded 10. It is reasonable to suppose that KINOSHITA's minimizing function assumes values which closely approach the values of the first eigenfunction of the Schrödinger equation of helium. So one could test some hypotheses regarding this latter function how closely they are satisfied by KINOSHITA's approximate eigenfunction. We mention two of these hypotheses: if ψ is the first eigenfunction and if it can be written in the form

$$\psi = \sum_{n=0}^{\infty}\psi_n,$$

where ψ_n is a homogeneous function of r_1, r_2 and r_{12} of degree n, then

$$\psi_0 = 1 \tag{1.8a}$$

and

$$\psi_1 = -2r_1-2r_2+\tfrac{1}{2}r_{12} \tag{1.8b}$$

KINOSHITA's approximate solution fulfils the condition that its constituent of degree 0 is a constant, on the other hand its part of degree 1 approaches poorly the linear function ψ_1. Modifying, however, the original variational problem by restricting suitably the class of functions admitted to concurrence the corresponding Euler equations, or, more accurately, systems of Euler equations admit solutions whose constituents of degree 0 and 1 are given by ψ_0 and ψ_1, respectively.

---

* KINOSHITA has shown, too, that there exist formal solutions of (1.1) in the form

$$\sum_{l,m,n=0}^{\infty} d_{lmn}s^l p^m q^{2n}.$$

However, the actual convergence of these formal solutions is not known.

More specifically let us restrict the functions admitted to concurrence in the variational problem (1.2) to those which satisfy the additional conditions that they are of the form

$$W = \sum_{i=1}^{M} \varphi_i(s) p^{m_i} q^{n_i} \tag{1.9}$$

where $\varphi_i(s)$ is twice differentiable,

$$\int_0^\infty \varphi_i^2(s) s^5 ds < \infty \quad \text{and} \quad \lim_{s\to\infty} \varphi_i'(s) s^{5/2} < \infty \tag{1.10}$$

($i=1, 2, \ldots, M$), finally $\lim_{s\to 0} W$ exists and is independent of p and q (i.e. W is continuous at the origin). In (1.9) m_i and $n_i/2$ are nonnegative integers, $m_1=n_1=0$, $m_2=1$, $n_2=0$. Conditions (1.10) ensure that

$$\lim_{s\to\infty} \varphi_i(s) s^{5/2} = 0 \quad \text{and} \quad \int_{-\infty}^{\infty} \ldots \int W^2 dx_1 dy_1 \ldots dz_2 < \infty. \tag{1.11}$$

Let us denote by C the class of functions satisfying these conditions and having piecewise continuous partial derivatives. We shall call modified variational problem the following: to find the minimum, the stationary values and the corresponding stationary functions (if they exist at all) of (1.2) if the functions admitted to concurrence are of the class C.

We are going to show in this paper that if the values of certain determinants given later do not vanish and if there exist stationary functions to the modified variational problem, then *every stationary function is of the form*

$$1 - Zs + \tfrac{1}{2}u + O(s^2) \tag{1.12}$$

in the neighbourhood of $s=0$, apart from a constant factor.

This result supports the likelihood of the hypotheses concerning the ground state of the helium atom expressed in (1.8).

2. Let us consider the following more general variational problem. Let $W(x_1, \ldots, x_N)$ be a quadratic integrable function in the entire Euclidean space $x_1, x_2, \ldots, x_N$, let $\mu_1, \mu_2, \ldots, \mu_N$ be non-negative constants and U a homogeneous function of degree -1 of the variables $x_1, x_2, \ldots, x_N$. The problem is to find the stationary values, if any, of the quotient

$$\int_{-\infty}^{\infty} \ldots \int \left[\sum_{\nu=1}^{N} \mu_\nu \left(\frac{\partial W}{\partial x_\nu} \right)^2 + UW^2 \right] dx_1 \ldots dx_N \Big/ \int_{-\infty}^{\infty} \ldots \int W^2 dx_1 \ldots dx_N. \tag{2.1}$$

Instead of this we shall consider the following restricted variational problem. Let H_n denote the class of homogeneous functions of degree n of the variables $x_1, x_2, \ldots, x_N$, and let $q_1, q_2, \ldots, q_N$ be new coordinates, such that the Jacobian

$$J = \frac{\partial(x_1, x_2, \ldots, x_N)}{\partial(q_1, q_2, \ldots, q_N)} \tag{2.2}$$

does not vanish identically. We suppose that $q_\nu \in H_0$ if $\nu=1, 2, \ldots, N-1$, $q_N \in H_1$

$(0 \leqq q_N < \infty)$. Let Σ be the set of all $(N-1)$-tuplets $q_1, q_2, \ldots, q_{N-1}$ when the x_ν's range from $-\infty$ to ∞. (Since $q_\nu(x_1, \ldots, x_N) = q_\nu(\varkappa x_1, \ldots, \varkappa x_N)$ if $\nu < N$, $\varkappa > 0$, Σ can be defined alternatively as the set of all $(N-1)$-tuples $q_1, \ldots, q_{N-1}$ when the x_ν's range on any hypersurface $q_N = \text{const.}$ of the space $x_1, \ldots, x_N$.) Let $\chi_1, \chi_2, \ldots, \chi_M$ be linearly independent fixed functions of $q_1, q_2, \ldots, q_{N-1}$, defined in each point of Σ (thus $\chi_\nu \in H_0$) and let finally $\varphi_i(q_N)$ $(i = 1, 2, \ldots, M)$ be arbitrary functions twice differentiable in the range $0 \leqq q_N < \infty$ such that the quantities

$$\int_0^\infty [\varphi_i(q_N)]^2 q_N^{N-1} dq_N \quad \text{and} \quad \lim_{q_N \to \infty} \varphi_i'(q_N) q_N^{(N-1)/2} \tag{2.2a}$$

exist and are finite, further

$$\varphi_i(0) = c_i \qquad (i = 1, 2, \ldots, M) \tag{2.2b}$$

where the c_i's are some prescribed constants. The stationary values of (2.1), if they exist, should be found if W is restricted to functions of the form

$$W = \sum_{i=1}^{M} \varphi_i(q_N) \chi_i . \tag{2.3}$$

This leads to the search of the system of solutions of the set of Euler equations of our variational problem. In this section we are going to show that *the coefficients of these solutions satisfy under the above suppositions a three-term vectorial recursion formula.*

We have by virtue of our suppositions

$$\frac{\partial q_\alpha}{\partial x_\nu} \in H_{-1}, \quad \text{if} \quad \alpha < N; \quad \frac{\partial q_N}{\partial x_\nu} \in H_0 \qquad (\nu = 1, 2, \ldots, N) \tag{2.4}$$

thus J is of the form

$$J = q_N^{N-1} g(q_1, \ldots, q_{N-1}), \tag{2.5}$$

hence $g(q_1, \ldots, q_{N-1}) \in H_0$, and

$$h_{\alpha\beta} = h_{\beta\alpha} = \sum_{\nu=1}^{N} \mu_\nu \frac{\partial q_\alpha}{\partial x_\nu} \frac{\partial q_\beta}{\partial x_\nu} \begin{cases} \in H_{-2}, & \text{if} \quad \alpha < N, \beta < N \\ \in H_{-1}, & \text{if} \quad \alpha < N, \beta = N \\ \in H_0, & \text{if} \quad \alpha = N, \beta = N. \end{cases} \tag{2.6}$$

Then we have

$$\sum_{\nu=1}^{N} \mu_\nu \left(\frac{\partial W}{\partial x_\nu} \right)^2 = \sum_{i,j=1}^{M} \left(\alpha_{ij} \varphi_i' \varphi_j' + 2 \frac{\beta_{ij}}{\sigma} \varphi_i \varphi_j' + \frac{\gamma_{ij}}{\sigma^2} \varphi_i \varphi_j \right) \tag{2.7}$$

where dashes mean derivation with respect to the variable

$$\sigma = q_N \tag{2.8}$$

and the functions α_{ij}, β_{ij}, γ_{ij}, all belonging to the class H_0, are defined by

$$\alpha_{ij} = \chi_i \chi_j h_{NN},$$

$$\frac{\beta_{ij}}{\sigma} = \sum_{\alpha=1}^{N-1} \frac{\partial \chi_i}{\partial q_\alpha} \chi_j h_{\alpha N}, \tag{2.9}$$

$$\frac{\gamma_{ij}}{\sigma^2} = \sum_{\alpha=1}^{N-1} \sum_{\beta=1}^{N-1} \frac{\partial \chi_i}{\partial q_\alpha} \frac{\partial \chi_j}{\partial q_\beta} h_{\alpha\beta}.$$

Note that $\alpha_{ij}=\alpha_{ji}$, $\gamma_{ij}=\gamma_{ji}$, but $\beta_{ij}\neq\beta_{ji}$.

If W is of the form (2.3) subject to conditions (2.2a) and (2.2b) and it is a stationary function of the restricted variational problem, then there exists a constant λ such that

$$\delta \int_{-\infty}^{\infty} \dots \int \left[\sum_{\nu=1}^{N} \mu_\nu \left(\frac{\partial W}{\partial x_\nu} \right)^2 + UW^2 + \lambda W^2 \right] dx_1 \dots dx_N = 0 \tag{2.10}$$

for all variations δW of the form $\Sigma \chi_i \delta\varphi_i$. Using the notations $d\Sigma = dq_1, \dots, dq_{N-1}$, $g=g(q_1, \dots, q_{N-1})$

$$\begin{cases} U = \dfrac{1}{\sigma} V, \; A_{ij} = \displaystyle\int_\Sigma \alpha_{ij} g \, d\Sigma, \; B_{ij} = \int_\Sigma \beta_{ij} g d\Sigma, \\ C_{ij} = \displaystyle\int_\Sigma \gamma_{ij} g \, d\Sigma, \; D_{ij} = \int_\Sigma V \chi_i \chi_j g \, d\Sigma, \; E_{ij} = \int_\Sigma \chi_i \chi_j g \, d\Sigma \end{cases} \tag{2.11}$$

(where $V, A_{ij}, B_{ij}, \dots, E_{ij} \in H_0$) we transform equation (2.10) by introducing the coordinates $q_1, q_2, \dots, q_{N-1}$, σ and get

$$\delta \int_0^\infty \Phi(\varphi_1, \dots, \varphi_M, \varphi_1', \dots, \varphi_M', \sigma) d\sigma = 0 \tag{2.12}$$

where

$$\Phi = \sum_{i,j=1}^{M} \left[A_{ij} \varphi_i' \varphi_j' + 2 \frac{B_{ij}}{\sigma} \varphi_i \varphi_j' + \left(\frac{C_{ij}}{\sigma^2} + \frac{D_{ij}}{\sigma} + \lambda E_{ij} \right) \varphi_i \varphi_j \right] \sigma^{N-1}. \tag{2.13}$$

With the help of conditions (2.2) one can derive the Euler equations of this variational problem. They are

$$\sum_{j=1}^{M} \left[A_{ij} \varphi_j'' + \frac{F_{ij}}{\sigma} \varphi_j' + \left(\frac{G_{ij}}{\sigma^2} - \frac{D_{ij}}{\sigma} - \lambda E_{ij} \right) \varphi_j \right] = 0 \tag{2.14}$$

($i=1, 2, \dots, M$) where

$$F_{ij} = (N-1) A_{ij} + B_{ji} - B_{ij}, \quad G_{ij} = (N-2) B_{ji} - C_{ij}. \tag{2.15}$$

Let now $\varphi=\varphi(\sigma)$ be the column vector consisting of the elements $\varphi_1, \varphi_2, \dots, \varphi_M$. If one defines the matrices $A, B, \dots, G$ so that their elements should be $A_{ij}, B_{ij}, \dots$ $\dots, G_{ij}$, respectively, then (2.14) can be rewritten in the form of the single vectorial

differential equation

$$A\varphi'' + \frac{1}{\sigma} F\varphi' + \left(\frac{1}{\sigma^2} G - \frac{1}{\sigma} D - \lambda E\right)\varphi = 0. \tag{2.16}$$

Introducing now the M-vector

$$\omega = e^{\mu\sigma}\varphi = \sum_{l=0}^{\infty} v_l \sigma^l, \quad v_l = \{v_{l1}, \ldots, v_{lM}\}$$

where μ is a constant, and substituting $\varphi = e^{-\mu\sigma}\omega$ into (2.16) one gets a three-term recurrence formula for the numerical vectors v_l:

$$[l(l-1)A + lF + G]v_l = [2\mu(l-1)A + \mu F + D]v_{l-1} + (\lambda E - \mu^2 A)v_{l-2} \tag{2.17}$$

($l = 0, 1, 2, \ldots$). This formula leads to a formal solution of (2.16) provided

$$Gv_0 = 0, \quad v_0 \neq 0 \quad \text{thus} \quad \det G = 0, \tag{2.18}$$

$$\det [l(l-1)A + lF + G] \neq 0 \qquad (l = 1, 2, \ldots). \tag{2.19}$$

(2.18) imposes a restriction on the initial values c_ν (see (2.2b)), since $v_{0\nu} = c_\nu$ ($\nu = 1, 2, \ldots, M$).

In the special case $\chi_1 \equiv 1$ one has by (2.9) and (2.11)

$$\beta_{1j} = \gamma_{1j} = B_{1j} = C_{1j} = 0 \qquad (j = 1, 2, \ldots, M)$$

hence by (2.15) each element in the first column of G is 0. Thus (2.18) can be satisfied by $v_0 = \{1, 0, 0, \ldots, 0\}$ or in the case of the initial conditions $c_1 \neq 0$, $c_\nu = 0$ ($\nu = 2, 3, \ldots, M$). (These initial conditions are the only ones for which $W = \Sigma\varphi_i\chi_i$ is continuous in the neighbourhood of the point $x_1 = x_2 = \cdots = x_N = 0$.)

Now one can show by standard methods that if (2.18) and (2.19) are fulfilled, then the formal solution obtained from (2.17) converges for each finite σ. All one needs to use to this end is the fact that A is a positive definite matrix. We restrict ourselves to the particular case when

$$h_{NN} = h = \text{const.}, \tag{2.20}$$

cf. (2.6). Then by (2.9) and (2.11)

$$A = hE \tag{2.21}$$

and if one chooses $\mu = \sqrt{\lambda/h}$, the coefficient matrix of v_{l-2} in (2.17) vanishes. The recurrence formula reduces now to

$$H_l v_l = \frac{1}{l} K_l v_{l-1} \quad (l > 1)$$

where

$$H_l = A + \frac{1}{l-1} F + \frac{1}{l(l-1)} G, \quad K_l = 2\sqrt{\lambda/h}A + \frac{1}{l-1}(\sqrt{\lambda/h}F + D).$$

E is positive definite since

$$\sum_{ij} E_{ij}\xi_i\xi_j = \int_\Sigma \left(\sum_i \xi_i\chi_i\right)^2 g\,d\Sigma > 0 \tag{2.22}$$

whenever $\Sigma\xi_i^2 > 0$, thus $\det E$ and $\det A = \det hE$ are non-vanishing and so are $\det H_l$

and det K_l if $l>l_0$ where l_0 is sufficiently lange. Thus H_l^{-1} exists ($l>l_0$) and there exists a constant $\varkappa$ such that for each vector v one has $|H_l^{-1}v|<\varkappa|v|$, $|K_l v|<\varkappa|v|$ ($l>l_0$) where $|v|$ means the Euclidean norm of v. Hence

$$|v_{il}| \leqq |v_l| < \frac{\varkappa^2}{l}|v_{l-1}| < \cdots < \frac{\varkappa^{2l}}{l!}\frac{l_0!}{\varkappa^{2(l_0+1)}}|v_{l_0}|$$

and

$$\left|\sum_{l=l_0+1}^{\infty} v_{il}\sigma^l\right| \leqq \frac{l_0!}{\varkappa^{2(l_0+1)}}|v_{l_0}|\sum_{l_0+1}^{\infty}\frac{\varkappa^{2l}}{l!}|\sigma^l| \leqq \text{const}\, e^{\varkappa^2\sigma}.$$

3. Let us now return to the special case of Section **1**, the case of helium-like atoms. We now have $N=6$, $\mu_1=\mu_2=\cdots=\mu_6=1/2$, $U=-Zr_1^{-1}-Zr_2^{-1}+r_{12}^{-1}$ and choose $q_1=p$, $q_2=q$, ..., $q_6=s$ using the notations (1.7) and (1.5).* If $f(s, t, u)$ is an even function of t then one has with the notation (1.5)

$$\int\int\cdots\int_{-\infty}^{\infty} f(s,t,u)\,dx_1 dy_1 \ldots dz_2 = 2\pi^2\int_0^{\infty}\int_0^{s}\int_0^{u} f(s,t,u)(s^2-t^2)u\,dt\,du\,ds =$$

$$= 2\pi^2\int_0^{\infty}\int_0^{1}\int_0^{1} f(s,pqs,ps)s^5(1-p^2q^2)p^2\,dp\,dq\,ds$$

(cf. [1] p. 233). The functions χ_i should now be of the form

$$\chi_i = p^{m_i}q^{n_i} \quad (i=1,2,\ldots,M) \tag{3.1}$$

where the m_i's are non-negative integers and the n_i's non-negative even integers.

Then we have for $1\leqq i, j\leqq M$, $c=(2\pi^2)^{-1}$

$$cA_{ij} = \frac{1}{(m_i+m_j+3)(n_i+n_j+1)} - \frac{1}{(m_i+m_j+5)(n_i+n_j+3)}; \tag{3.2}$$

$$cB_{ij} = \frac{m_i}{(m_i+m_j+5)(n_i+n_j+3)} + \frac{n_i-m_i}{(m_i+m_j+3)(n_i+n_j+3)} - \frac{n_i}{(m_i+n_i+3)(m_j+n_j+1)}; \tag{3.3}$$

$$cC_{ij} = (-m_im_j+n_in_j)\left[\frac{1}{(m_i+m_j+3)(n_i+n_j+1)} - \frac{1}{(m_i+m_j+3)(n_i+n_j+3)} - \frac{1}{(m_i+m_j+1)(n_i+n_j+1)}\right] - \frac{m_im_j}{(m_i+m_j+5)(n_i+n_j+3)} + \frac{n_in_j}{(m_i+m_j+1)(n_i+n_j-1)}; \tag{3.4}$$

$$cD_{ij} = \frac{1}{(m_i+m_j+2)(n_i+n_j+1)} - \frac{1}{(m_i+m_j+4)(n_i+n_j+3)} - \frac{4Z}{(m_i+m_j+3)(n_i+n_j+1)} \tag{3.5}$$

* q_3, q_4, q_5 need not be given explicitly since we shall not use them. They are some quantities fixing the position of a triangle of sides r_1, r_2, r_{12} in the space.

and since

$$\left(\frac{\partial s}{\partial x_1}\right)^2+\left(\frac{\partial s}{\partial y_1}\right)^2+\cdots\left(\frac{\partial s}{\partial z_2}\right)^2=2$$

we have by (2.6) $h=1$ and by (2.21)

$$E_{ij}=A_{ij}. \tag{3.6}$$

Formulae (3.2) through (3.5) are straightforward consequences of the integral

$$\int_0^1\int_0^1 p^\varrho q^\sigma\,dp\,dq=\frac{1}{\varrho+1}\frac{1}{\sigma+1}$$

(cf. formula (32.8) in [1]).

Now if $m_1=n_1=0$, $m_2=1$, $n_2=0$, $M\geqq 2$ then it can be shown by a direct substitution that the recurrence formula (2.17) is satisfied in the cases $\mu=0$, $l=0$, and $\mu=0$, $l=1$ by the vectors

$$v_0=\{1, 0, 0, \ldots, 0\} \quad \text{and} \quad v_1=\{-Z, \tfrac{1}{2}, 0, \ldots, 0\},$$

respectively. If, moreover, the rank of G is $M-1$ and (2.19) holds for $l=1$ then these are the only solution vectors of the recurrence formula. In this case the homogeneous part of degree 0 and 1 of each stationary solution of the variational problem of Section 3 is

$$1+(-Z+\tfrac{1}{2}p)s=1-Zs+\tfrac{1}{2}u=1-Z(r_1+r_2)+\tfrac{1}{2}r_{12}$$

as stated in (1.12).

We remark that the rank of G is $M-1$ and (2.19) holds for each positive l in the case $M=2$, $m_1=0$, $n_1=0$, $m_2=1$, $n_2=0$ so our statement is not void.

For some sets $\chi_1, \chi_2, \ldots, \chi_M$ $(M>2)$ defined by (3.1) and for $Z=2$, $\mu=0$ P. KOSIK calculated with a computer the values of v_1 assuming $v_0=\{1, 0, \ldots, 0\}$. He found for small M's the only solution $v_1=\{-2, \frac{1}{2}, 0, \ldots, 0\}$ of (2. 17) which indicates that (2. 19) may hold for $l=1$ in several cases.

REFERENCES

[1] BETHE, H. A. and SALPETER, E. E.: *Quantum Mechanics of one- and two-electron systems*, Handbuch der Physik, vol. 35, Springer, Berlin—Göttingen—Heidelberg, 1957.

[2] KINOSHITA, T.: Ground state of the helium atom I, II, *Phys. Rev.* **105** (1957), 1490—1502 and **115** (1959), 366-374.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received August 15, 1972.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 141—150.

PERMUTATION GROUPS AND ALMOST REGULAR GRAPHS

by

L. BABAI and L. LOVÁSZ

Introduction. The *automorphism group* $A(X)$ of a graph X consists of the edge-preserving permutations of the vertex set $V(X)$. For any finite group G and $k \geqq 3$, there exists a graph X, regular of degree k, such that $A(X) \cong G$ (FRUCHT [4], SABIDUSSI [7]). The situation is more complicated if a permutation group P is given and we want to construct a graph X with automorphism group P; this is, in general impossible. However, it is possible to embed the underlying set Ω of P into the set of vertices of a graph X as a set invariant under $A(X)$ so that the restriction of the elements of $A(X)$ onto Ω is an isomorphism of $A(X)$ onto P (BOUWER [2], BABAI [1]; for generalizations see PULTR [6]). This graph X can be required to have several properties such as prescribed chromatic number, connectivity, spanning topological subgraph [1]. In [1] it was mentioned that X can be regular of sufficiently large degree and for $P = A_n$ (the alternating group) the maximum degree must be $> \frac{n}{2}$.

In the present paper we give a necessary and sufficient condition for a permutation group to be the restriction of the automorphism group of a graph of maximum degree $\leqq k$. We shall show that the condition is sufficient for the embeddability into the vertex set of a regular graph of degree k. One can require that the vertices in Ω have degree r while the other vertices have degree k ($r \geqq k$), except if k is even and r is odd. In this latter case, obviously $2 \mid r|\Omega|$ is a necessary condition for the embeddability. However, it is insufficient. Our rather complicated counterexample (Proposition 2) suggests that this leads to deep questions concerning permutation groups. A number of sufficient conditions is given in Theorem 4. We call the reader's attention to the special consequences, formulated in the Remark after Theorem 3 and in Corollaries 1—6.

Definitions. A *binary (relational) system* X means a vertex set $V(X)$ and a set $\mathscr{R}(X)$ of binary relations on $V(X)$. If $R \in \mathscr{R}(X)$, $(a, b) \in R$ then (a, b) is called an *edge* of X, having *colour* R. This edge is *directed* if $(b, a) \notin R$, *undirected* otherwise. *The automorphism group* $A(X)$ of X consists of those permutations of $V(X)$ which preserve all relations of $\mathscr{R}(X)$.

A *graph* Y (without loops and multiple edges) is a binary system, containing only one irreflexive relation $E(Y)$ (the set of edges). Y is *undirected* if $E(Y)$ is a symmetric relation. Our constructions will give undirected graphs (cf. the Lemma).

The *out-degree* of a vertex a of a relational system X is defined by

$$\sum_{R \in \mathscr{R}(X)} \sum_{\substack{x \in V(X) \\ (a, x) \in R}} 1$$

We define the *in-degree* analogously. The *degree* is the sum of the out- and in-degree. If talking about *undirected graphs* by degree we mean the out-degree.

A permutation group P acting on a set Ω means a group P and a *not necessarily faithful* representation of P in the symmetric group on Ω. Thus the action of P is not necessarily effective. A permutation group Q, acting on a set Σ is an *extension* of P if $P \cong Q$, $\Sigma \supseteq \Omega$, Ω is invariant under Q and there is an isomorphism φ of P onto Q such that the action of $\pi\varphi$ and π coincide on Ω for any $\pi \in P$. Q is an *effective extension* of P if Q acts effectively on Σ. If P acts effectively then so do its extensions. The automorphism group of a binary system acts effectively.

The *stabilizer* $\mathscr{S}_x$ of $x \in \Omega$ is the subgroup of P, leaving x fixed.

§ 1. We begin with a rather technical lemma; some similar idea has been used in almost all related constructions.

LEMMA. *If W is a binary relational system and $g \geqq 3$ a given integer, then there is an undirected graph Z such that $V(Z) \supseteq V(W)$, the degrees of the vertices of $V(Z) - V(W)$ are g, those of the vertices of $V(W)$ are the same in Z as in W and $A(Z)$ is an (effective) extension of $A(W)$. Moreover, if $R_0 \in \mathscr{R}(W)$ is a symmetric relation* then Z can be choosen so that $x, y \in V(W)$ are joined in Z iff $(x, y) \in R_0$.*

PROOF. Let $m > 2g$ and consider a graph Y_m defined as follows: Let $V(Y_m) = \{x_1, \ldots, x_{2m}\}$; connect x_i to x_j $(j \equiv i+r \pmod{2m}, 1 \leqq r < 2m)$ if $1 \leqq r \leqq \left[\frac{g}{2}\right]$ and in case g is odd also if $r = m$. Let us observe the following properties of Y_m:

(a) Y_m is regular of degree g.

(b) No proper automorphism leaves x_1 and x_{2m} fixed. It is enough to show that if an automorphism φ fixes x_1 and x_{2m} then it also fixes x_2. This is easily seen if $g = 3$. Assume $g > 3$. φ fixes $x_{[g/2]+1}$ since this is the only vertex joined to x_1 by an edge belonging to a triangle, but not to x_{2m}. Similarly, φ fixes $x_{2m-[g/2]}$. Consequently, φ keeps the set $\{x_{2m-1}, \ldots, x_{2m-[g/2]+1}\}$ invariant, since these are exactly those vertices which are joined to x_1, x_{2m} and $x_{2m-[g/2]}$. Now x_2 is the only vertex joined to x_1, x_{2m}, $x_{[g/2]+1}$ and to exactly $\left[\frac{g}{2}\right] - 2$ vertices of $\{x_{2m-1}, \ldots, x_{2m-[g/2]+1}\}$.

Let Y'_m denote the graph obtained by removing the edge (x_1, x_{2m}). x_1, x_{2m} will be called the *terminal vertices* of Y'_m.

(c) Each vertex of Y'_m is contained in a triangle or a quadrangle; the edges of Y'_m contained in a triangle or quadrangle form a connected graph.

Consider now W and associate a number $m_R > g|V(W)|$ with every $R \in \mathscr{R}(W)$, $R \neq R_0$, in such a way that $m_R \neq m_{R'}$ if $R \neq R' \in \mathscr{R}(W)$. For each pair $(x, y) \in R \in \mathscr{R}(W)$, $R \neq R_0$, consider two graphs Y'_{m_R}, Y'_{2m_R}. Join x to a terminal vertex of Y'_{m_R}, y to a terminal vertex of Y'_{2m_R} and the remaining terminal vertices of Y'_{m_R}, Y'_{2m_R} to each other. Moreover, join $x, y \in V(W)$ to each other iff $(x, y) \in R_0$. We show that the graph Z constructed above satisfies the requirements of the Lemma.

The only non-trivial property of Z is that each automorphism of it is the unique extension of an automorphism of W. Let $\varphi \in A(Z)$. Removing those edges of Z which are not contained in any triangle or quadrangle, φ is an automorphism of the obtained graph Z' as well. Z' has components Y'_m by (c) and certain small components lying in $V(W)$; therefore, φ keeps $V(W)$ invariant and maps the graphs

* In this case the two edges (x, y), $(y, x) \in R_0$ should be counted only once in the degree of x.

Y'_m onto each other. Since $m_R \neq m_{R'}$ for $R \neq R'$, a pair $(x, y) \in R$ must be mapped onto a pair in the same relation, i.e. $\varphi | V(W) \in A(W)$. Finally, (b) implies that φ is the unique extension of $\varphi | V(W)$.

§ 2. THEOREM 1. *Let X be a connected graph, $x \in V(X)$ a vertex of degree r and assume the other vertices of X to have degree $\leqq k$ ($r \leqq k$). Let $\mathscr{S}_x$ be the stabilizer of x in $A(X)$. Then there is a chain $\mathscr{S}_x = H_0 \geqq \cdots \geqq H_t = \{1\}$ of subgroups such that $|H_i : H_{i+1}| \leqq \leqq k-1$ $(i = 1, \ldots, t-1)$, $|H_0 : H_1| \leqq r$.*

Remark: It would be enough to state the cases $r = k$ and $r = k-1$, since if $r < k-1$ the conclusion is the same as in case $r = k-1$ (choosing $H_1 = H_0$).

PROOF. Let $x_0 = x, x_1, \ldots, x_n$ be an arrangement of the vertices such that x_i $(i \neq 0)$ is connected to some preceding x_j. Let

$$H_i = \mathscr{S}_{x_0} \cap \mathscr{S}_{x_1} \cap \cdots \cap \mathscr{S}_{x_i}.$$

Then $H_0 = S_x$, $H_n = \{1\}$ and $|H_i : H_{i+1}|$ is the cardinality of the orbit O of x_{i+1} in H_i. If x_j is a vertex connected to x_{i+1} $(j \leqq i)$ then O lies in the neighbourhood N of x_j; and $O = N$ can happen only if $i = 0$. This proves the assertion.

The same argument gives the following

THEOREM 1′. *Let X be a graph, $x \in V(X)$ a vertex of degree r, X_1 a connected subgraph of X such that the stabilizer of $V(X_1)$ consists of the identity only; $x \in V(X_1)$ and the vertices of $X_1 - x$ have degree $\leqq k$ in X. Then there is a chain $\mathscr{S}_x = H_0 \geqq \cdots \geqq \cdots \geqq H_t = \{1\}$ of subgraph of $\mathscr{S}_x$ such that $|H_i : H_{i+1}| \leqq k-1$ $(i = 1, \ldots, t-1)$ and $|H_0 : H_1| \leqq r$.*

§ 3. The following construction shows that Theorem 1 can be converted:

THEOREM 2. *Let $g \geqq 3$, $k \geqq 3$, $r \geqq 1$ be given integers, P a permutation group acting on the set Ω and assume that in the stabilizer $\mathscr{S}_x$ of any point $x \in \Omega$ there is a chain $\mathscr{S}_x = H_0^{(x)} \geqq \cdots \geqq H_t^{(x)} = \{1\}$ of subgraphs such that $|H_0^{(x)} : H_1^{(x)}| \leqq r$ and $|H_i^{(x)} : H_{i+1}^{(x)}| \leqq \leqq k-1$ $(i = 1, \ldots, t-1)$. Then there exists a connected graph X such that $\Omega \subseteq V(X)$, $A(X)$ is an effective extension of P, the vertices of Ω have degree $\leqq r$, there are $\leqq |\Omega| \cdot r \cdot \dfrac{(k-1)^{t+1}-1}{k-2}$ points of degree $\leqq k$ and the other vertices have degree g.*

PROOF. Let $\{c_1, \ldots, c_q\}$ be a system of representatives of the orbits of P. Consider a graph Y, regular of degree g, such that $A(Y) \cong P$ and the stabilizer of any edge of Y is the identity (see [4, 7]). Let $\pi^* \in A(Y)$ correspond to $\pi \in P$. Choose an $(a, b) \in E(Y)$.

Let M_{ij} be the set of right cosets of $H_i^{(c_j)}$ in P $(i = 0, \ldots, t;\ j = 1, \ldots, q)$ and associate with every $K \in M_{ij}$ a vertex x_{ijK}. If $K \in M_{0j}$ then K is a right coset of $\mathscr{S}_{c_j}$ and thus $c_j K$ consists of a single point of Ω, which can be chosen for x_{0jK}. Let Ω^* be the set of these new points x_{ijK}.

Define the binary system W as follows: Let $V(W) = V(Y) \cup \Omega^*$. Define a symmetric relation R' on $V(W)$ as follows: $(x, y) \in R'$ iff $(x, y) \in E(Y) - (a, b) A(Y)$. Let R'' consist of all pairs $(a\pi^*, x_{t1\{\pi\}})$, $(x_{t1\{\pi\}}, x_{t2\{\pi\}})$, $\ldots$, $(x_{tq\{\pi\}}, b\pi^*)$, $\pi \in P$. Finally, let $R_i = \{(x_{i-1, jK}, x_{ij,L}) : K \supseteq L\}$ $(i = 1, \ldots, t)$.

We show that $A(W)$ is isomorphic to P in a natural way. For any $\pi \in P$, define a permutation π^{**} of $V(W)$ as follows:

$$x_{ijK}\pi^{**} = x_{i,j,K\pi}, \quad y\pi^{**} = y\pi^{*} \quad (y \in V(Y)).$$

Then, obviously, π^{**} is an automorphism of W. On the other hand, let α be an automorphism of W. Since the elements of $V(Y)$ are exactly those vertices which occur in pairs in relation R', α keeps $V(Y)$ invariant. Moreover, $a\alpha$, $b\alpha$ are joined by some R''-path, hence there is a $\varrho \in P$ such that $a\alpha = a\varrho^{*}$, $b\alpha = b\varrho^{*}$. It follows also that $\alpha | V(Y)$ is an automorphism of Y, hence $\alpha | V(Y) = \varrho^{*}$. By the definition of R'', we have

$$x_{tj\{\pi\}}\alpha = x_{tj\{\pi\}}\varrho^{**}$$

and hence it follows by induction on v

$$x_{t-v,j,K}\alpha = x_{t-v,j,K}\varrho^{**},$$

i.e. $\alpha = \varrho^{**}$. Thus, $A(W)$ is an extension of P.

It is easily seen that $x_{0j,\mathscr{S}}$ has degree $|\mathscr{S}_{c_j} : H_1^{(c_j)}| \leqq r$ $(\mathscr{S} = \mathscr{S}_{c_j})$, x_{ijK} has degree $1 + |H_i^{(c_j)} : H_{i+1}^{(c_j)}| \leqq k$ $(i = 1, \ldots, t-1)$, $x_{ij\{\pi\}}$ has degree 3 $(\pi \in P)$, the other vertices have degree g.

From W we can construct a graph X with the required properties by the Lemma.

§ 4. We show that in the case $r = k$ the bounds given in Theorem 2 can be considered as prescribed values. For the sake of simplicity we consider the case $g = k$.

THEOREM 3. *If X is an undirected graph with maximum degree $\leqq k$ then there exists a regular graph Y of degree k such that $A(Y)$ is an extension of $A(X)$ and Y has as many connected components as X $(k \geqq 3)$.*

PROOF. Subdivide each edge of X by two points; denote by Z the arising graph. Consider a graph Z' isomorphic to Z, $V(Z) \cap V(Z') = \emptyset$. Let φ be an isomorphism of Z onto Z'. Let W be a binary system defined by $V(W) = V(Z) \cup V(Z')$, $\mathscr{R}(W) = \{E(Z) \cup E(Z'), R_1, \ldots, R_k, S_1, \ldots, S_k\}$ where

$$R_i = \{(x, x\varphi) : x \in V(X) \text{ is of degree less than } i\};$$
$$S_i = \{(x, x\varphi) : x \in V(Z) - V(X) \text{ is of degree less than } i\}.$$

Obviously, $A(W)$ is an extension of $A(X)$. By the Lemma, we can find a graph Y such that $A(Y)$ is an extension of $A(W)$ and Y is regular of degree k.

Remark. The question whether any graph X of maximum degree not exceeding k can be embedded into some graph Y regular of degree k such that $A(Y)$ is an extension of $A(X)$ still remains open.

Remark: From theorems 1, 2, 3 one can easily deduce that

for any permutation group P, there exists a $k_0(P)$ such that there exists a connected regular graph of degree k, the automorphism group of which is an effective extension of P, iff $k \geqq k_0(P)$.

If P is the symmetric or alternating group of degree n $(n \geqq 4)$ then $k_0(P) = n - 1$.

§ 5. The situation becomes more complicated if we want to find a graph X to a given permutation group P, acting on the set Ω such that $\Omega \subseteq V(X)$, $A(X)$ is an effective extension of P and the vertices of Ω are of degree r, the other vertices of degree k (r, k given integers). Obviously, if r is odd and k is even then $|\Omega|$ must be even, but this is not a sufficient condition.

THEOREM 4. *Let P be a permutation group acting on the set Ω and let $r \geqq 1$, $k \geqq 3$. Assume that either k is odd or $r \cdot |\Omega|$ is even, and in the stabilizer $\mathscr{S}_x$ of each $x \in \Omega$ there is a chain $\mathscr{S}_x = H_0^{(x)} \geqq \cdots \geqq H_t^{(x)} = \{1\}$ such that $|\mathscr{S}_x : H_1^{(x)}| \leqq r$ and $|H_i^{(x)} : H_{i+1}^{(x)}| \leqq k-1$ for every $x \in \Omega$, $i = 1, \ldots, t-1$. Then there exists a graph X such that $V(X) \supseteq \Omega$, $A(X)$ is an effective extension of P, the vertices of Ω have degree r, the other vertices have degree k, provided one of the following conditions is satisfied:*

a) *k is odd;*
b) *$|\mathscr{S}_x|$ is odd, $|\mathscr{S}_x : H_1^{(x)}| \leqq k$ for every $x \in \Omega$;*
c) *$|\mathscr{S}_x|$ and r are odd for every $x \in \Omega$;*
d) *$|\mathscr{S}_x|$ and r are even, $|H_1^{(x)}|$ is odd for every $x \in \Omega$;*
e) *r is even, $|\mathscr{S}_x : H_1^{(x)}| \leqq k$ for every $x \in \Omega$;*
f) *Ω has an invariant partition $\Omega = \bigcup_1^m \Omega_i$, $|\Omega_i|$ is even and $\leqq \max(k, r) - 1$, and $|\mathscr{S}_x : H_1^{(x)}| \leqq k-1$ for every $x \in \Omega$;*
g) *for every $x \in \Omega$ there is a subgroup K_x of P, $K_x \supseteq \mathscr{S}_x$, $|K_x : \mathscr{S}_x|$ is even and $\leqq \max(k, r) - 1$, and $|\mathscr{S}_x : H_1^{(x)}| \leqq k-1$;*
h) *$|\mathscr{S}_x|$ is odd and there is a subgroup K_x of P, $K_x \supseteq \mathscr{S}_x$, $|K_x : \mathscr{S}_x|$ is even and $|K_x : \mathscr{S}_x| + |\mathscr{S}_x : H_1^{(x)}| \leqq r+2$ (e.g. $|K_x : \mathscr{S}_x| = 2$) for every $x \in \Omega$.*

From this rather technically formulated theorem, we can deduce several corollaries which are probably easier to survey:

Corollaries: Given a permutation group P, acting on the set Ω, let p be the largest prime divisor of $\prod_{x \in \Omega} |\mathscr{S}_x|$. P is called *k-representable* if there exists a connected graph X such that $V(X) \supseteq \Omega$, $A(X)$ is an effective extension of P, the vertices of Ω have degree $< k$, the other vertices have degree k. By theorem 1, $p \leqq k-1$ is a necessary condition for being k-representable. In certain cases this condition is also sufficient:

1. *$|\mathscr{S}_x|$ is odd ($x \in \Omega$);*
2. *$\mathscr{S}_x$ is supersolvable and $|\mathscr{S}_x|$ is not a power of p ($x \in \Omega$);*
3. *$p = 2$.*

PROOF. If $\mathscr{S}_x$ is solvable, it contains a chain $\mathscr{S}_x = H_0^{(x)} \geqq H_1^{(x)} \geqq \cdots \geqq H_t^{(x)} = \{1\}$ where $|H_i^{(x)} : H_{i+1}^{(x)}| \leqq p$ ($i = 0, \ldots, t-1$). Hence corollary 1. follows from Theorem 4b and the FEIT—THOMPSON theorem [3].

In corollary 2. $\mathscr{S}_x$ has a chain of subgroups such that $|H_{i-1}^{(x)} : H_i^{(x)}| \leqq |H_i^{(x)} : H_{i+1}^{(x)}|$ ($i = 1, \ldots, t-1$) [5, p. 206]. Hence $|\mathscr{S}_x : H_1^{(x)}| \leqq p-1 \leqq k-2$. Putting $r = k-2$, cases a) and e) of Theorem 4 imply Corollary 2.

If $p = 2$, $\mathscr{S}_x$ has a chain of subgroups such that $|H_i^{(x)} : H_{i+1}^{(x)}| \leqq 2$ ($i = 0, 1, \ldots, t-1$). Choosing an $r < k$, r even, the application of Theorem 4e) gives the statement.

If $p=2$, we assert more:

COROLLARY 4. *Let P be a permutation group acting on Ω, and $k \geqq 3$, $r \geqq 1$. Assume that, for any $x \in \Omega$, $|\mathscr{S}_x|$ is a power of 2 and $r \cdot |\Omega|$ is even. Then there exists a connected graph X such that $V(X) \supseteq \Omega$, the vertices of Ω have degree r, the other vertices have degree k and $A(X)$ is an effective extension of P.*

PROOF. $\mathscr{S}_x$ has a chain $\mathscr{S}_x = H_0^{(x)} = H_1^{(x)} \geqq H_2^{(x)} \geqq \ldots \geqq H_t^{(x)} = \{1\}$ of subgroups such that $|H_i^{(x)} : H_{i+1}^{(x)}| \leqq 2$ $(i = 1, 2, \ldots, t-1)$ and $|\mathscr{S}_x : H_1^{(x)}| = 1$. Cases a) and e) of Theorem 4 imply the statement if k is odd or r is even. Assume k is even and r is odd. Then $|\Omega|$ is even.

Let G be a 2-Sylow subgroup of P; $|G| = 2^l$. We shall apply case f) of Theorem 4. Let $\Omega = \bigcup_{i=1}^{q} \Sigma_i$ where Σ_i are the orbits of P. If $|\Sigma_i|$ is even then $\mathscr{S}_x$ $(x \in \Sigma_i)$ is contained in a 2-Sylow subgroup of P as a proper subgroup, hence there is a subgroup K_x of P, containing $\mathscr{S}_x$ such that $|K_x : \mathscr{S}_x| = 2$. Thus, $\Sigma_i = \bigcup_{\pi \in P} xK_x\pi$ is a P-invariant partition of Σ_i into 2-element subsets. The number of those indices i, for which $|\Sigma_i|$ is odd, is even: let $|\Sigma_1|, \ldots, |\Sigma_{2d}|$ be odd. This means $|\mathscr{S}_x| = 2^l$ $(x \in \Sigma_i,\ 1 \leqq i \leqq 2d)$ thus $\mathscr{S}_x$ is a 2-Sylow subgroup of P: $\mathscr{S}_x = \pi G \pi^{-1}$. Hence $\mathscr{S}_{x\pi} = G$, and there is a system $\{c_1, \ldots, c_{2d}\}$ of representatives of the Σ_i's such that $\mathscr{S}_{c_i} = G$. Therefore, $\bigcup_{\pi \in P} \{c_{2i-1}\pi,\ c_{2i}\pi\}$ is an invariant partition of $\Sigma_{2i-1} \cup \Sigma_{2i}$ $(i = 1, \ldots, d)$. Thus we have constructed a P-invariant partition of Ω into 2-element subsets, hence case f) of Theorem 4 can be applied.

COROLLARY 5. *Let G be a group. A necessary and sufficient condition for G to be isomorphic to the stabilizer of an edge in the automorphism group of a connected regular graph of degree 3 is that $|G| = 2^l$ $(l \geqq 0)$.*

PROOF. Let Y be a regular graph of degree 3, $e \in E(Y)$. Let us subdivide e by a new vertex x; in this way we obtain a graph X. $\mathscr{S}_x$ is isomorphic to the stabilizer of e in $A(Y)$ and the degree of x is 2. Theorem 1 implies that there is a chain of subgroups of $\mathscr{S}_x$ such that $|H_i^{(x)} : H_{i+1}^{(x)}| \leqq 2$, hence $|\mathscr{S}_x| = 2^l$ for some $l \geqq 0$.

Conversely, let $\Omega = \{x\}$ and the group G of order 2^l act trivially on Ω. Put $P = G$ and apply Corollary 3. Let the neighbours of x be a and b in the resulting graph X. Put new vertices a_1, b_1 on the edges (x, a) and (x, b), resp., and join a new vertex y to a_1, b_1 and x, by directed edges. Applying the Lemma we obtain a graph Y which is regular of degree 3, and whose automorphism group is isomorphic to the automorphism group of X, and fixes, e.g., the „new" edge incident with x. Therefore, the stabilizer of this edge in Y is isomorphic to G.

A similar argument gives

COROLLARY 6. *A permutation group P is (permutationally) isomorphic to the permutation group induced by the automorphisms of a connected regular graph of degree 3 on an invariant set of edges, if and only if all stabilizer subgroups of P are 2-groups.*

§ 6. PROOF of Theorem 4. First we note that if $k \equiv i \pmod 2$, $i = 1, 2$ then there exists a graph Q_i such that all vertices of Q_i are of degree k except one vertex q_i which is of degree i, and $A(Q_i)$ is the identity.

Put $q=k$ in Theorem 2 and consider the graph X_0 constructed in the proof of Theorem 2.

I. Let k be odd. Assume that the degree of the vertex $y \in V(X_0)$ is less than required. Let us associate a copy Q_1^j of Q_1 with every element y_j of the orbit $yA(X_0)$, and identify the vertices y_j and q_1^j. However, consider the edges of the Q_1^j's as a relation different from the edges of X_0. Thus we obtain a binary system Y_0. Obviously, $A(Y_0)$ is an extension of $A(X_0)$. The application of the Lemma yields a graph X_1 such that $A(X_1)$ is an extension of $A(X_0)$, the vertices of $V(X_1)-V(X_0)$ have degree k and the vertices of $V(X_0)$ have the same degree in X_0 as in X_1, except the vertices y_j the degree of which is increased by one. Repeating this procedure we obtain the required graph X. This proves Theorem 4 in case a).

II. In what follows we assume k is even, hence so is $r \cdot |\Omega|$. By the same method as in part I. (using the binary system Q_2) we can increase the degree of the vertices of $yA(X_0)$ by an arbitrary even number. Thus we only have to prove the existence of a connected binary system X' such that $V(X') \supseteq \Omega$, $A(X')$ is an extension of P and for $x \in \Omega$, $\varrho(x) \leqq r$, $r-\varrho(x)$ is even; and for $x \in V(X')-\Omega$, $\varrho(x) \leqq k$, $k-\varrho(x)$ is even, where $\varrho(x)$ denotes the degree of the vertex x.

Assume we already have a connected binary system X'', such that $A(X'')$ is an extension of Ω, $\varrho(x) \leqq r$ and $r-\varrho(x)$ is even if $x \in \Omega$, $\varrho(x) \leqq k$ and $k-\varrho(x)$ is even if $x \in V(X'')-\Omega-\Sigma$ where Σ is an $A(X'')$ invariant subset of $V(X'')$, such that each vertex of Σ has degree 3 and the stabilizer of any vertex of Σ contains the identity only. Then we can construct the required binary system X' as follows.

As $k, r|\Omega|$ and the number of vertices of odd degree are even, $|\Sigma|$ is even. Σ is invariant under $A(X'')$, hence $\Sigma = \bigcup_{i=1}^{l} \Sigma_i$ where the Σ_i's are the orbits of $A(X'')|\Sigma$; obviously $|\Sigma_i| = |P|$. We prove that Σ has an invariant partition $\Sigma = \cup \{a_j, b_j\}$. If l is even, this is clear. If l is odd, then by $|P| \cdot l = |\Sigma|$, $|P|$ is even, hence there is a subgroup of P of order 2; the system of its left cosets defines a partition of P, invariant under right multiplication. As Σ_i and P are in a natural one-to-one correspondence, we obtain a partition of Σ.

Now, adding the new paths (a_j, c_j, k_j) (whose edges are considered to belong to a new relation and c_j are new vertices), the obtained binary system has the required properties, since k is even.

We now turn to the construction of X''. If $|\mathscr{S}_x|+r$ is even and $|H_1^{(x)}|$ is odd we may simply set $X''=X_0$, and this settles cases c) and d).

If $|\mathscr{S}_x|$ is odd and $|\mathscr{S}_x : H_1^{(x)}| \leqq k$ then $|\mathscr{S}_x : H_1^{(x)}|$ is odd and hence $\leqq k-1$. Let $X_0, X_1, \ldots, X_r$ be disjoint copies of the relational system X_0 constructed in the proof of theorem 2; define X'' on the set $\Omega \cup V(X_1) \cup \cdots \cup V(X_r)$ by adding new edges of a new i^{th} relation to the union of $X_1, \ldots, X_r$, connecting the corresponding vertices of Ω and X_i. The resulting binary system X'' proves the theorem in case b).

Consider the case e). Let X_1 be a binary system, isomorphic to X_0, and Y the disjoint union of X_0 and X_1. Let us connect each vertex of X_0 of odd degree to the corresponding vertex of X_1 by a directed edge. The obtained binary system X' has only vertices of even degree $\leqq k$: the vertices of Ω have degree $\leqq r$. $A(X_0)$ naturally extends to a subgroup of $A(X')$; we have only to show that this is not a proper subgroup. Let $\varphi \in A(X')$. Then, obviously, φ leaves $V(X_0)$ invariant and $\varphi|V(X_0) \in$

$\in A(X_0)$. We may assume $\varphi|V(X_0)=1$. Then φ fixes the vertices of X_1 of degree 3, hence, by the construction of X_0, $\varphi=1$.

In case f) we apply a similar idea. Let Y be the union of the binary systems X_1, X_2, isomorphic to X_0 (X_0, X_1, X_2 are disjoint). By e) we may assume r is odd. Let Ω_i' denote the set, corresponding to Ω in X_i ($i=1, 2$). The vertices of Y have degree $\leqq k$, moreover, those of $\Omega_1' \cup \Omega_2'$ have degree $\leqq k-1$. Let us connect each vertex of $V(X_1)-\Omega_1'$ of odd degree and each vertex of Ω_1' of even degree to the corresponding vertex of X_2 by a directed edge. Let, moreover, Ω' denote a set, $|\Omega'|=|\Omega|$, $\Omega'=\{x': x\in\Omega\}$ disjoint from $V(X_0)$, $V(X_1)$, $V(X_2)$, and add edges of a new relation, connecting each vertex of Ω' to the corresponding vertices of Ω_1' and Ω_2'. The resulting binary system U has even degrees $\leqq k$, the degree of the vertices of Ω' is 2. Clearly, $A(U)$ is an extension of $A(X_1)$. Let us add to U all undirected edges (x', y'); $x, y\in\Omega_i$, $x\neq y$, $i=1, \ldots, m$.

If $k\leqq r$, identify Ω and Ω' and put $X'=U$.

If $k>r$, define X' on the set $\Omega\cup V(U)$ by adding the edges (x, x'), $x\in\Omega$ to U.

An easy calculation shows that that the binary system constructed satisfies the requirements in both cases, i.e. case f) is settled.

Condition g) is a special case of condition f):

Let $c_1, \ldots, c_q$ be a system of representatives of the orbits of P and $\pi_1^{(i)}, \ldots, \pi_{n_i}^{(i)}$ a system of representatives of the right cosets of K_{c_i} in P. Put

$$\Omega_{ij} = c_i K_{c_i} \pi_j^{(i)} \qquad (i = 1, \ldots, q;\ j = 1, \ldots, n_i)$$

Clearly, $\Omega = \cup \Omega_{ij}$ and this is an invariant partition. $|\Omega_{ij}|=|K_{c_i}:\mathscr{S}_{c_i}|$ is even and $\leqq \max(k, r)-1$ by the condition.

Finally, case h) will be a consequence of the other cases. By a) and c) we may assume both k and r are even. As above we define the partition $\Omega = \cup \Omega_{ij}$; now $|\Omega_{ij}|+|\mathscr{S}_{c_i}:H_1^{(c_i)}| \leqq r+1$ (as $\leqq r+2$ is the assumption and $|\Omega_{ij}|$, r are even). Define X' by adding the undirected edges $(x, y): x, y\in\Omega_{ij}$ ($i=1, \ldots, q$; $j=1, \ldots, n_i$) of a new relation to X_0. Obviously, X' satisfies the requirements.

§ 7. Counterexamples. We show that in Theorem 4, cases b) and e) the condition $|\mathscr{S}_x:H_1^{(x)}|\leqq k$, and in case e) the condition "r is even" cannot be avoided, in the first case even if P is Abelian of odd order.

Proposition 1. Let P be a permutation group, isomorphic to the cyclic group of order $p\geqq 5$, prime, acting on the set $\Omega=\{a_1, \ldots, a_p, a_{p+1}\}$ in such a way that its generating element π permutes $a_1, \ldots, a_p$ cyclically and leaves a_{p+1} fixed. (Then $\mathscr{S}_{a_i}=\{1\}$ ($i=1, \ldots, p$) and $\mathscr{S}_{a_{p+1}}=P$.) We assert that there is no 2-edge-connected graph X with the following properties: $V(X)\supseteq\Omega$; $A(X)$ contains a subgroup $G\cong P$ such that $G|\Omega=P$ and the degree of the vertices of Ω is $p+1$, that of the other vertices is $\leqq p$. (Note that a connected graph with even degrees would be necessarily 2-edge-connected.)

PROOF. Assume there is a graph with the properties mentioned. Let N denote the neighbourhood of a_{p+1}; $|N| = p+1$. Since a_{p+1} is fixed by π, N is invariant under G. Therefore, either one or all points of N are fixed by G.

Consider a path $x_0, x_1, \ldots, x_s$ such that x_0 is a point of N fixed by G, x_s is one of $a_1, \ldots, a_p$ and the path has no other point in Ω. If N is fixed by G, then an arbitrary path connecting a_{p+1} and a_1 has such a section. If there is one point $y \in N$ fixed by G, then a_{p+1} cannot separate y from $\{a_1, \ldots, a_p\}$ since there is a $z \in N$ which is not separated from $\{a_1, \ldots, a_p\}$ by a_{p+1}, then neither are its images under G and y cannot be the only neighbour of a_{p+1} separated from $\{a_1, \ldots, a_p\}$ by a_{p+1}, since then (a_{p+1}, y) would be a cutting edge. Therefore, there is a path connecting y to one of $a_1, \ldots, a_p$ with the required properties.

Now x_0 is fixed by G while x_s is not, thus there is an $0 \leqq i < s$ such that x_i is fixed by G but x_{i+1} is not. But π has order p, hence x_i has degree $\geqq p+1$, a contradiction.

A more complicated example shows that in case k even, r odd, $|\Omega|$ even conditions like f), g) cannot be omitted, even if P acts transitively and $|\mathscr{S}_x : H_1^{(x)}| \leqq \leqq k-2$.

The stabilizer $\mathscr{S}_x$ of the group below clearly contains a chain $\mathscr{S}_x = H_0^{(x)} \geqq \cdots \cdots \geqq H_7^{(x)} = \{1\}$ of subgroups such that $|H_i : H_{i+1}| \leqq 6$. Let S_n denote, as usual, the symmetric group of degree n.

Proposition 2. Let $S_8^{(2)}$ be the group induced by the permutations of the vertices of a complete 8-gon on the set of its 28 edges; in other words, the (faithful) representation of S_8 on the right cosets of the subgroup $S_6 \times S_2$. Then there is no graph X with the following properties: $V(X)$ contains a subset Ω of 28 elements of odd degree $\leqq 7$, the other vertices are of even degree $\leqq 8$, $A(X)$ contains a subgroup $G \cong S_8$ such that $G|\Omega$ is permutationally isomorphic to $S_8^{(2)}$.

PROOF. Let X be a counterexample with minimum number of edges. Let $x \in \Omega$, N_x the neighbourhood of x. Then $|N_x|$ is odd and $\leqq 7$. We show that there is an $x' \in N_x$ fixed by $\mathscr{S}_x$, the stabilizer of x in G. The action of $\mathscr{S}_x \cong S_6 \times S_2$ cannot be effective on N_x since $6! \cdot 2! \nmid 7!$. Therefore, a proper factor group of $S_6 \times S_2$ operates on N_x. $S_6 \times S_2$ has proper factor groups S_6, $S_2 \times S_2$, S_2. In the first case, the existence of x' is a consequence of the well known theorem asserting that S_7 has no transitive subgroup of index 7 [8, p. 103]; in the other cases this is trivial as well.

Observe that $S_6 \times S_2$ is a maximal subgroup in S_8, therefore the stabilizer $\mathscr{S}_{x'}$ of x' is either $\mathscr{S}_x$ or G; the latter case is impossible since then x' would have degree $\geqq 28$. Thus $\mathscr{S}_{x'} = \mathscr{S}_x$. If $x' \in \Omega$ then there is a $\varphi \in G$ such that $x\varphi = x'$, $\varphi^{-1}\mathscr{S}_x\varphi = \mathscr{S}_{x'} = \mathscr{S}_x$, φ is in the normalizer of $\mathscr{S}_x$. By $\varphi \notin \mathscr{S}_x$ and the maximality of $\mathscr{S}_x$ we obtain that $\mathscr{S}_x$ is a normal subgroup, which is not the case. Hence $x' \notin \Omega$.

Define the graph X_1 by omitting the edges of $(x, x')G$ from X. Put $\Omega_1 = x'G$. The degrees of vertices of $\Omega \cup \Omega_1$ decrease by one; $\mathscr{S}_x = \mathscr{S}_{x'}$ and hence $G|_{\Omega_1}$ is also permutationally isomorphic to $S_8^{(2)}$. This contradicts the minimality of X.

Remark: Instead of $S_8^{(2)}$ and degrees $\leqq 7$ and $\leqq 8$, we could consider $\mathscr{S}_n^{(2)}$, $\leqq n-1$ and $\leqq n$, resp., where $n \equiv 0 \pmod 4$ and $n \geqq 8$. If $n \equiv 2 \pmod 4$ then the proposition is obvious as $|\Omega| = \binom{n}{2}$ is odd.

REFERENCES

[1] BABAI, L.: Representation of permutation groups by graphs, *Comb. Theory and its Appl.* (P. Erdős et al. eds., North Holland, Budapest, 1970), 55—80.
[2] BOUWER, I. Z.: Section graphs for finite permutation groups, *J. Comb. Theory* **6** (1969) 378—386.
[3] FEIT, W., THOMPSON, J. G.: Solvability of groups of odd order, *Pacific J. Math.* **13** (1963), 775—1029.
[4] FRUCHT, R.: Graphs of degree 3 with given abstract group, *Canad. J. Math.* **1** (1949), 365—378.
[5] KOCHENDÖRFFER, R.: *Lehrbuch der Gruppentheorie unter besonderer Berücksichtigung der endlichenGruppen,* Leipzig, 1966.
[6] PULTR, A.: On full embeddings of concrete categories with respect to forgetful functors, *Comment. Math. Univ. Carolinae* **9** (1968), 281—305.
[7] SABIDUSSI, G.: Graphs with given group and given graph theoretical properties, *Canad. J. Math.* (1957), 515—525.
[8] TSCHEBOTAROW, N.: *Grundzüge der Galois'schen Theorie.* P. Nordhoff N. V., 1950.

Eötvös L. University, Budapest

(Received July 15, 1972.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 151—152.

A NOTE TO A PAPER OF DUDLEY

by

L. LOVÁSZ and P. MAJOR

R.M. DUDLEY proved [1] some results about realization of probability measures on a metric space as distribution of random variables. Among others he showed, generalizing a result of Skorohod, that, given a separable metric space S and a sequence $\{P_n\}$ of probability measures weakly converging to a probability measure P_0 on S, there exists a probability space $(\Omega, \mathscr{A}, P)$ and S-valued random variables X_n such that X_n has distribution P_n $(n=0, 1, \ldots)$ and $X_n \to X_0$ with probability 1.

In the present note we prove the following

THEOREM. *Let S be a compact connected metric space; let the probability measures P_n (on the Borel sets of S) converge weakly to a probability measure P_0 such that every open set has positive P_0 measure. Then there is a probability space $(\Omega, \mathscr{A}, P)$ and random variables X_n, $(a=0, 1, \ldots)$ such that X_n has distribution P_n and $X_n \to X_0$ uniformly.*

It is a surprising fact, occuring already in DUDLEY's paper, that the proof of the theorem is based on a combinatorial lemma, the KÖNIG—EGERVÁRY theorem. However, here we use its "continuous" form:

Let G be a bipartite graph with vertices x_i and y_j $(i, j=1, 2, \ldots, n)$ such that the edges join x_i's with y_j's; associate a non-negative real number $p(u)$ with every vertex u of G. It is possible to associate a non-negative real $v(e)$ with every edge e of G in such a way that the sum of values $v(e)$ associated with edges incident with u be $p(u)$, if and only if (i) $\Sigma p(y_i) = \Sigma p(x_i)$ *and* (ii) *if $I \subseteq \{1, \ldots, n\}$ and J is the set of indices of points y_j connected to points x_i $(i \in I)$ then $\sum_{j \in J} p(y_j) \geqq \sum_{i \in I} p(x_i)$.*

PROOF of the Theorem. Let $k \geqq 1$. Divide S into disjoint sets $F_1^{(k)}, \ldots, F_{r_k}^{(k)}$ with diameters $< \frac{1}{k}$, with boundaries of P_0-probability 0 and having an inner point. Then

$$\lim_{n \to \infty} P_n(F_i^{(k)}) = P_0(F_i^{(k)}) > 0. \tag{1}$$

Let us choose an n_k such that

$$P_m(F_i^{(k)}) > \left(1 - \min_{j \leqq r_k} P_m(F_j^{(k)})\right) P_0(F_i^{(k)}) \tag{2}$$

for $m \geqq n_k$. We may assume $n_1 < n_2 < \cdots$.

For $n_k \leqq m < n_{k+1}$, consider the following bipartite graph G_m: let x_i, y_j, $(i, j=1, 2, \ldots, r_k)$ be its vertices, x_i and y_j connected by an edge iff $\overline{F_i^{(k)}} \cap \overline{F_j^{(k)}} \neq \emptyset$, and let $p(x_i) = P_0(F_i^{(k)})$, $p(y_i) = P_m(F_i^{(k)})$. Then the conditions of the KÖNIG—EGER-

VÁRY theorem are fulfiled:

$$\sum_{i=1}^{r_k} p(x_i) = \sum_{i=1}^{r_k} p(y_i) = 1$$

and if $I \subset \{1, \dots, n\}$ then the points x_i $(i \in I)$ are connected to every point y_i $(i \in I)$ and, because of the connectedness of S, to at least one point y_j, $j \notin I$. Thus (2) implies (ii).

Let $a_{ij}^{(m)}$ denote the value associated with the edge connecting x_i to y_j, and $a_{ij}^{(m)} = 0$ if no such edge exists. Consider the Cartesian product $S \times S \times \dots = \Omega'$ and the product σ-algebra $\mathscr{A}'$ on it. Denote by $X_n(\omega)$ the $(a+1)$-st coordinate of $\omega \in \Omega'$. We define a probability measure P on Ω' as follows: let $P(X_0 \in A) = P_0(A)$ and

$$P(X_m \in A | X_0 = x_0, \dots, X_{m-1} = x_{m-1}) = P(X_m \in A | X_0 = x_0) =$$

$$= \sum_{j=1}^{r_k} \frac{a_{ij}^{(m)}}{P_0(F_i^{(k)})} P_m(A | F_j^{(k)})$$

for $n_k \leqq m < n_{k+1}$ and $x_0 \in F_i^{(k)}$. Then

$$P(X_m \in A) = \sum_i \sum_j a_{ij}^{(m)} \frac{P_m(A \cap F_j^{(k)})}{P_m(F_j^{(k)})} = \sum_j P_m(A \cap F_j^{(k)}) = P_m(A).$$

On the other hand, from the construction it follows that $P\left(d(X_0, X_m) > \frac{2}{k}\right) = 0$ for $m \geqq n_k$ and thus the set $\Omega = \bigcap_{\substack{k,m \\ n_k \leqq m}} \left\{d(X_0, X_m) \leqq \frac{2}{k}\right\}$ has probability 1. Thus, the functions X_n restricted to Ω give the required construction.

Remark: The condition that S be connected and P_0 be positive seems to be somewhat artificial. However, it is easy to see that they are necessary. For assume S is not connected, and let P_0 be a positive probability measure on it. Let $S = S_1 \cup S_2$, where $d(S_1, S_2) = \delta > 0$. Define the probability measure P_n by

$$P_n(A) = \left(1 + \frac{P_0(S_2)}{n}\right) P_0(A \cap S_1) + \left(1 - \frac{P_0(S_1)}{n}\right) P_0(A \cap S_2).$$

Then there are no random variables X_n with the properties required in the theorem; really, for each n there would be an ω such that $X_n(\omega) \in S_1$ but $X_0(\omega) \in S_2$ (because $P_0(S_1) < P_n(S_1)$), i.e. $d(X_n, X_0) \geqq \delta$.

On the other hand, imbed the preceding space S into a compact connected metric space S_1, and define all measures to be 0 on $S_1 - S$. Then, obviously, the preceding argument remains valid.

REFERENCES

[1] DUDLEY, R. M.: Distances of probability measures and random variables. *Ann. Math. Statist.* **39** (1968), 1563—1572.

Eötvös L. University, Budapest
Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received July 15, 1972.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 153—159.

BOUNDS FOR THE EXTREMA OF THE EXPECTED VALUE OF A CONVEX FUNCTION OF INDEPENDENT RANDOM VARIABLES

by

D. STOYAN

1. Introduction

For some areas of applied probability (statistics, queueing, reliability) the following problem is of importance (see, for example, [1] [7]): Let f be a real valued function on R^n and let $\mathfrak{M}_1, \mathfrak{M}_2, \ldots, \mathfrak{M}_n$ be sets of distribution functions. The problem is to determine the least upper and the greatest lower bound of the expected value

$$\int_{R^n} f(x_1, x_2, \ldots, x_n) dF_1(x_1) dF_2(x_2) \ldots dF_n(x_n) \tag{1.1}$$

where $F_i \in \mathfrak{M}_i$.

This problem is solved in some papers for special cases. HOEFFDING [4] has proved that for continuous functions f and sets $\mathfrak{M}_i$ of the type

$$\mathfrak{M}_i = \Big\{F: \int_{-\infty}^{\infty} g_j^{(i)}(x) dF(x) = c_{ij},\ F(x) = 0 \quad (x \leqq a_i),\ F(x) = 1 \quad (x > b_i),$$

$$j = 1, 2, \ldots, k\Big\}$$

the supremum and infimum of the integral (1. 1) are given by distribution functions with $k+1$ steps. (Their positions depend on the special form of the function f and are in general non-independent of i.) In some cases there exist more detailed results, for example, in the papers of HARRIS [2] (when the first k derivatives of f are monotone) and HARTLEY—DAVID [3] ($f(x_1, x_2, \ldots, x_n) = \max\{x_1, x_2, \ldots, x_n\}$). Using the results in [5], one can solve the problem for convex functions f and sets $\mathfrak{M}_i$ of distribution functions with the same (finite) mean possessing a maximum or minimum with respect to the partial ordering relation $\overset{(2)}{\leqq}$ (see [5]). (Such sets are, for example, sets of normal distributions with fixed mean and bounded variance or sets of gamma distributions with fixed mean.)

In the present paper we assume that the function f is convex in all variables and the sets $\mathfrak{M}_i$ are of the type

$$\mathfrak{M}_i = \Big\{F: \int_{-\infty}^{\infty} x\, dF(x) = m_i, \quad \int_{-\infty}^{\infty} (x - m_i)^2\, dF(x) = \sigma_i^2, \quad F(x) = 1(x > b_i)\Big\}. \tag{1.2}$$

In this case we give a method for the determination of upper and lower bounds for M and m, where

$$M = \sup_{F_i \in \mathfrak{M}_i} \int_{R^n} f(x_1, x_2, \ldots, x_n) dF_1(x_1) dF_2(x_2) \ldots dF_n(x_n)$$

and

$$m = \inf_{F_i \in \mathfrak{M}_i} \int_{R^n} f(x_1, x_2, \ldots, x_n) dF_1(x_1) dF_2(x_2) \ldots dF_n(x_n).$$

These bounds are given by suitable "extremal distribution functions", which are independent of the the form of f and will be considered in the next section.

2. Extremal distribution functions

By the symbol $\overset{(2)}{\leqq}$ we shall denote the following partial ordering relation for distribution functions with finite mean [6]:

$$F_1 \overset{(2)}{\leqq} F_2 \leftrightarrow xF_1(x) + \int_x^\infty t\, dF_1(t) \leqq xF_2(x) + \int_x^\infty t\, dF_2(t), \quad \forall x.$$

Then the inequality

$$\int_{R^n} f(x_1, x_2, \ldots, x_n) dF_1(x_1) dF_2(x_2) \ldots dF_n(x_n) \leqq$$

$$\leqq \int_{R^n} f(x_1, x_2, \ldots, x_n) dG_1(x_1) dG_2(x_2) \ldots dG_n(x_n)$$

follows from

$$F_i \overset{(2)}{\leqq} G_i; \qquad i = 1, 2, \ldots, n$$

for every convex function f (for which the integrals exist) and for distribution functions $F_1, F_2, \ldots, F_n$ and $G_1, G_2, \ldots, G_n$ with the (finite) means $m_1, m_2, \ldots, m_n$ (see [6]).

We have, therefore, as a corollary of the results of [6] the following

THEOREM 1. *For every convex function f on the R^n the relations*

$$L_i \overset{(2)}{\leqq} F_i \overset{(2)}{\leqq} U_i \quad \textit{for all} \quad F_i \in \mathfrak{M}_i$$

imply the inequalities

$$M \leqq \int_{R^n} f(x_1, x_2, \ldots, x_n) dU_1(x_1) dU_2(x_2) \ldots dU_n(x_n)$$

$$m \geqq \int_{R^n} f(x_1, x_2, \ldots, x_n) dL_1(x_1) dL_2(x_2) \ldots dL_n(x_n).$$

Applying theorem 1, we obtain good estimations for M and m if U_i and L_i are the suprema and infima of the sets $\mathfrak{M}_i$ with respect to $\overset{(2)}{\leqq}$, respectively. (For example, the supremum of a set $\mathfrak{M}$ with respect to $\overset{(2)}{\leqq}$ is the smallest distribution function with $F \overset{(2)}{\leqq} S$ for all $F \in \mathfrak{M}$.)

These extremal distribution functions for the sets (1.2) are given by the following theorems 2 and 3 which will be proved in section 3.

THEOREM 2. *Let $\mathfrak{M}_m^\sigma$ be the set of all distribution functions F satisfying*

$$\int_{-\infty}^{\infty} x\,dF(x) = m \tag{2.1}$$

and

$$\int_{-\infty}^{\infty} (x-m)^2\,dF(x) = \sigma^2. \tag{2.2}$$

Then the distribution function S_m^σ is the supremum (with respect to $\overset{(2)}{\leqq}$) of $\mathfrak{M}_m^\sigma$,

$$S_m^\sigma(x) = \frac{1}{2}\{1+(x-m)[(x-m)^2+\sigma^2]^{-\frac{1}{2}}\}; \quad -\infty < x < \infty,$$

and the distribution function D_m,

$$D_m(x) = \begin{cases} 0; & x \leqq m \\ 1; & x > m, \end{cases}$$

is the infimum of $\mathfrak{M}_m^\sigma$.

THEOREM 3. *Let $\mathfrak{M}_{m,b}^\sigma$ be the set of all distribution functions F satisfying* (2.1) *and* (2.2) *and*

$$F(x) = 0; \quad x \leqq 0 \quad \text{and} \quad F(x) = 1; \quad x > b \quad (b > 0).$$

Then the distribution function $S_{m,b}^\sigma$,

$$S_{m,b}^\sigma(x) = \begin{cases} 0 & ; \quad x \leqq 0 \\ \sigma^2(m^2+\sigma^2)^{-1}; & 0 < x \leqq \dfrac{1}{2m}(m^2+\sigma^2) \\ S_m^\sigma(x) & ; \quad \dfrac{1}{2m}(m^2+\sigma^2) < x \leqq x(b) \\ S_m^\sigma(x(b)) & ; \quad x(b) < x \leqq b \\ 1 & ; \quad x > b \end{cases}$$

$$x(b) = [m^2+\sigma^2-u(b)^2]\big[2(m-u(b))\big]^{-1}$$

$$u(b) = \{m[(m-b)^2+\sigma^2]-\sigma^2 b\}(m-b)^{-2}$$

is the supremum (with respect to $\overset{(2)}{\leqq}$) of $\mathfrak{M}_{m,b}^\sigma$ for $b<\infty$, the distribution function $S_{m,\infty}^\sigma$,

$$S_{m,\infty}^\sigma(x) = \begin{cases} 0 & ; \quad x \leqq 0 \\ \sigma^2(m^2+\sigma^2)^{-1}; & 0 < x \leqq \dfrac{1}{2m}(m^2+\sigma^2) \\ S_m^\sigma(x) & ; \quad x > \dfrac{1}{2m}(m^2+\sigma^2) \end{cases}$$

is the supremum of $\mathfrak{M}_{m,b}^\sigma$ for $b=\infty$.

The distribution function D_m of theorem 2 is the infimum of $\mathfrak{M}_{m,b}^{\sigma}$ in the case $b=\infty$, for $b<\infty$ the distribution function $I_{m,b}^{\sigma}$,

$$I_{m,b}^{\sigma}(x)=\begin{cases}0 & ; \quad 0 \leqq x < \left[m\left((b-m)^2+\sigma^2\right)-b\sigma^2\right](b-m)^{-2} \\ p_{m,b}^{\sigma}; & \left[m\left((b-m)^2+\sigma^2\right)-b\sigma^2\right](b-m)^{-2} \leqq x < \dfrac{1}{m}(m^2+\sigma^2) \\ 1 & ; \quad x > \dfrac{1}{m}(m^2+\sigma^2)\end{cases}$$

is the infimum (with respect to $\overset{(2)}{\leqq}$) of $\mathfrak{M}_{m,b}^{\sigma}$.

Here $p_{m,b}^{\sigma}$ denotes the function

$$p_{m,b}^{\sigma}=\sigma^2/m\left\{\left[m\left((b-m)^2+\sigma^2\right)-b\sigma^2\right](b-m)^2-(m^2+\sigma^2)m^{-1}\right\}.$$

We remark, in addition, that the rather complicated function $S_{m,b}^{\sigma}$ is dominated by $T_{m,b}^{\sigma}$ with respect to $\overset{(2)}{\leqq}$, i.e.

$$S_{m,b}^{\sigma} \overset{(2)}{\leqq} T_{m,b}^{\sigma}, \tag{2.3}$$

where

$$T_{m,b}^{\sigma}(x)=\begin{cases}0 & ; \quad x \leqq 0 \\ \sigma^2(m^2+\sigma^2)^{-1}; & 0 < x \leqq \dfrac{1}{2m}(m^2+\sigma^2) \\ q_{m,b}^{\sigma} & ; \quad \dfrac{1}{2m}(m^2+\sigma^2) < x \leqq b \\ 1 & ; \quad x > b\end{cases}$$

and

$$q_{m,b}^{\sigma}=\left(b-m-\frac{\sigma^2}{2m}\right)\left[b-\frac{1}{2m}(m^2+\sigma^2)\right]^{-1}.$$

3. Proofs of the theorems.

The proofs of the theorems 2 and 3 are elementary but rather lengthy. Therefore, in the following the proofs are sketched only.

PROOF of Theorem 2: The supremum S_m^{σ} of $\mathfrak{M}_m^{\sigma}$ with respect to $\overset{(2)}{\leqq}$ must have the following property: For every real number x and every $F \in \mathfrak{M}_m^{\sigma}$

$$xF(x)+\int_x^{\infty} t\,dF(t) \leqq xS_m^{\sigma}(x)+\int_x^{\infty} t\,dS_m^{\sigma}(t)=\mathscr{S}_m^{\sigma}(x).$$

holds.
Let

$$xF(x)+\int_x^{\infty} t\,dF(t)=\int_{-\infty}^{\infty} e_x(t)\,dF(t), \tag{3.1}$$

$$e_x(t)=\begin{cases}x; & t \leqq x \\ t; & t > x.\end{cases}$$

It is known from [4] that the supremum of the integral (3. 1) may be determined by considering only the three-point distributions of $\mathfrak{M}_m^\sigma$. In an elementary way one can show that it is sufficient even to consider only the two-point distributions of $\mathfrak{M}_m^\sigma$. $\mathscr{S}_m^\sigma$ is also the least real function satisfying

$$\mathscr{S}_m^\sigma(x) \geqq xZ(x) + \int_x^\infty t\, dZ(t) = \mathscr{Z}(x), \quad \forall x$$

for all two-point distributions Z of $\mathfrak{M}_m^\sigma$. If Z has the lower step a, then the function $\mathscr{Z}(x)$ is given by

$$\mathscr{Z}(x) = \begin{cases} m & ; \quad x \leqq a \\ m + p_a(x-a); & a < x \leqq (m - ap_a)(1-p_a)^{-1} \\ x & ; \quad x > (m-ap_a)(1-p_a)^{-1} \end{cases}$$

where

$$p_a = \sigma^2[(m-a)^2+\sigma^2]^{-1}.$$

Then envelope of the family of straight lines $\{g_a\}_{-\infty < a \leqq m}$,

$$g_a(x) = m + p_a(x-a); \qquad -\infty < x < \infty$$

satisfies the conditions imposed on $\mathscr{S}_m^\sigma$ and is given by

$$\mathscr{S}_m^\sigma(x) = m + \frac{1}{2}\{x - m + [(x-m)^2+\sigma^2]^{\frac{1}{2}}\}$$

Its derivative is the distribution function S_m^σ.[1])

To finish the proof of theorem 2, it remains to show that D_m is the infimum of $\mathfrak{M}_m^\sigma$ with respect to $\overset{(2)}{\leqq}$. This is easily done by considering two-point distributions with a step close to m and a very large second step.

PROOF of Theorem 3: Similarly to the proof of theorem 2 we can prove that $S_{m,b}^\sigma$ is the supremum of $\mathfrak{M}_{m,b}^\sigma$. Now we have to take into account that the parameter a of the family $\{g_a\}$ is bounded

$$0 \leqq a \leqq \{m - b\sigma^2[(b-m)^2+\sigma^2]^{-1}\}\{1-\sigma^2[(b-m)^2+\sigma^2]^{-1}\}^{-1}.$$

From this follows that the graph of the supremum function $\mathscr{S}_{m,b}^\sigma$,

$$\mathscr{S}_{m,b}^\sigma(x) = xS_{m,b}^\sigma(x) + \int_x^\infty t\, dS_{m,b}^\sigma(t)$$

consists of some curve pieces: the straight line $y = m$ $(-\infty < x \leqq 0)$, the straight line $y = m + x\sigma^2(m^2+\sigma^2)^{-1}$ $\left(0 \leqq x \leqq \frac{1}{2m}(\sigma^2+m)\right)$, the curve $y = S_m^\sigma(x)$ $\left(\frac{1}{2m}(m^2+\sigma^2) < x \leqq x(b)\right)$, the straight line $y = S_m^\sigma(x(b)) + (m-b)^2(x-x(b))[(m-b)^2+\sigma^2]^{-1}$

[1]) It holds in general $xF(x) + \int_x^\infty t\, dF(t) = m + \int_{-\infty}^x F(t)dt$.

$(x(b)<x\leqq b)$ and the straight line $y=x$ $(x>b)$. Then, by derivating $\mathscr{S}^{\sigma}_{m,b}$ we get $S^{\sigma}_{m,b}$ as stated in the theorem.

To prove the remainder, we remark first that similarly to the proof of theorem 2 the distribution D_m is the infimum of $\mathfrak{M}^{\sigma}_{m,b}$ in the case $b=\infty$. In the case $b<\infty$ we proceed as follows:

The infimum $I^{\sigma}_{m,b}$ of $\mathfrak{M}^{\sigma}_{m,b}$ with respect to $\overset{(2)}{\leqq}$ must have the following property: For every real number x and all $F\in\mathfrak{M}^{\sigma}_{m,b}$

$$xF(x)+\int_x^{\infty} t\,dF(t) \geqq xI^{\sigma}_{m,b}(x)+\int_x^{\infty} t\,dI^{\sigma}_{m,b}(t)$$

holds.

The infimum of the integral

$$\int_{-\infty}^{\infty} e_x(t)\,dF(t)$$

can be determined considering only the three-point distributions of $\mathfrak{M}^{\sigma}_{m,b}$ [4].

It is easy to verify that for $x\in\left[0,\ \{m[(b-m)^2+\sigma^2]-b\sigma^2\}(b-m)^{-2}\right]$ the infimum of the integral is equal to m. This value is given by the two-point distribution of $\mathfrak{M}^{\sigma}_{m,b}$ with the upper step in b. Furthermore it can be shown that for $x\in[(m^2+\sigma^2)m^{-1},\ b]$ the infimum is equal to x, the value x is given by the two-point distribution of $\mathfrak{M}^{\sigma}_{m,b}$ with the lower step in 0.

For $x\in\left[\{m[(b-m)^2+\sigma^2]-b\sigma^2\}(b-m)^{-2},\ (m^2+\sigma^2)m^{-1}\right]$, the infimum is given by three-point distributions with steps in 0, x and b. The values of the infima are located on the straight line connecting the points P_1 and P_2 of the (x, y)-plane,

$$P_1=\left(\{m[(b-m)^2+\sigma^2]-b\sigma^2\}(b-m)^{-2},\ m\right),\quad P_2=\left((m^2+\sigma^2)m^{-1},\ (m^2+\sigma^2)m^{-1}\right).$$

Therefore, the function

$$\mathscr{I}^{\sigma}_{m,b}(x)=xI^{\sigma}_{m,b}(x)+\int_x^{\infty} t\,dI^{\sigma}_{m,b}(t),$$

is given by

$$\mathscr{I}^{\sigma}_{m,b}(x)=\begin{cases} m;\ x\leqq\{m[(b-m)^2+\sigma^2]-b\sigma^2\}(b-m)^{-2} \\ m+\dfrac{\sigma^2}{m}\left[\{m[(b-m)^2+\sigma^2]-b\sigma^2\}(b-m)^{-2}-(m^2+\sigma^2)m^{-1}\right]; \\ \{m[(b-m)^2+\sigma^2]-b\sigma^2\}(b-m)^{-2}<x\leqq(m^2+\sigma^2)m^{-1} \\ x;\ x\geqq(m^2+\sigma^2)m^{-1}. \end{cases}$$

Derivating $\mathscr{I}^{\sigma}_{m,b}$ we get $I^{\sigma}_{m,b}$.

The graph of the function $\mathscr{T}^{\sigma}_{m,b}$,

$$\mathscr{T}^{\sigma}_{m,b}(x)=xT^{\sigma}_{m,b}(x)+\int_x^{\infty} t\,dT^{\sigma}_{m,b}(t),$$

consists of the following four straight line pieces:

$$y = m \qquad ; \quad -\infty < x \leqq 0$$

$$y = m + x\sigma^2(m^2+\sigma^2)^{-1}; \quad 0 \leqq x < \frac{1}{2m}(m^2+\sigma^2)$$

$$y = x \qquad ; \quad x \geqq b$$

and the straight line between the points $Q_1 = \left(\frac{1}{2m}(m^2+\sigma^2), m+\frac{\sigma^2}{2m}\right)$ and $Q_2 = (b, b)$.

From the convexity of $\mathcal{S}^{\sigma}_{m,b}$ it follows that the function $\mathcal{T}^{\sigma}_{m,b}$ is greater than $\mathcal{S}^{\sigma}_{m,b}$. The derivative of $\mathcal{T}^{\sigma}_{m,b}$ is the distribution function $T^{\sigma}_{m,b}$, and it satisfies relation (2. 3).

REFERENCES

[1] GNEDENKO, B. V.: Preface to the German edition of the book: Gnedenko—Kovalenko: *Introduction to queueing theory,* Berlin 1971.
[2] HARRIS, B.: Determining bounds on expected values of certain functions. *Ann. Math. Statist.* **33,** (1962) 1454—1457.
[5] HARTLEY, H. O., DAVID, H. A.: Universal bounds for mean range and extreme observation. *Ann. Math. Statist.* **25,** (1954) 85—99.
[4] HOEFFDING, W.: The extrema of the expected value of a function of independent random variables. Ann. Math. Statist. **26,** (1955) 268—275.
[5] STOYAN, D.: Monotonieeigenschaften stochastischer Modelle. *Z. Angew. Math. Mech.* **52,** (1972). 23—30.
[6] STOYAN, D.: Über einige Eigenschaften monotoner stochastischer Prozesse. *Math. Nachr.* **52,** (1972) 21—34.
[7] SZÁSZ, D.: On a nonlinear optimization problem (russ.) *Mat. Zam.* (to appear).

92 Freiberg, Brennstoffinstitut, DDR.

(Received July 26, 1972.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 161—164.

NOTES ON REGULAR DUO ELEMENTS, RINGS AND SEMIGROUPS

by

O. STEINFELD

§ 1. An associative ring (a semigroup) A is called a *duo ring (duo semigroup)* if every one-sided ideal of A is a two-sided ideal of A. (See E. H. FELLER [2].)

Let A be an associative ring or a semigroup. For an element a of A, $(a)_l[(a)_r]$ denotes the principal left [right] ideal of A generated by a. We call an element a of A *duo element* if $(a)_l=(a)_r$ holds.

First we characterize the duo rings and semigroups by the aid of the duo elements. Then we give equivalent conditions for an element of a ring (semigroup) to be regular, and regular duo respectively. (See Propositions 2 and 3.)

In the third part of this paper we mention some known characterizations of the regular duo rings. A simple example shows that a part of the analogous characterizations for regular duo, semigroups does not hold.

§ 2. PROPOSITION 1. An associative ring (a semigroup) A is a duo ring (a duo semigroup) if and only if every element of A is a duo element.

PROOF. Let L be a left ideal of A and let x be an arbitrary element of L. It holds

$$xA \subseteq (x)_l A = (x)_r A \subseteq (x)_r = (x)_l \subseteq L,$$

that is L is a two-sided ideal of A.

Similarly one can show that every right ideal of A is a two-sided ideal.

An element a of an associative ring (a semigroup) A is called *regular*, if $a \in aAa$. The ring (semigroup) A is called *regular*, if every element of A is regular.

Now we shall prove a partially known result. First we need the following notion.

A submodule (a non-void subset) Q of a ring (of a semigroup) A is called a *quasi-ideal* if $QA \cap AQ \subseteq Q$. It is evident that the intersection of a right ideal and a left ideal of A is a quasi-ideal of A.

PROPOSITION 2. The following three assertions concerning an element a of a ring (semigroup) A are equivalent[1])

(i) a is regular;
(ii) $(a)_r(a)_l = (a)_r \cap (a)_l$;
(iii) $(a)_r^2=(a)_r$, $(a)_l^2=(a)_l$ and the product $(a)_r(a)_l$ is a quasi-ideal of A.

PROOF. (i)$\Rightarrow$(ii). Our condition $a=axa$ implies $a \in aA$ and $a \in Aa$, whence $(a)_r = aA$ and $(a)_l = Aa$. Let us consider an arbitrary element $ak=la$ of $aA \cap Aa (k, l \in A)$. Hence $la=laxa=akxa \in aA \cdot Aa = (a)_r \cdot (a)_l$. The relation $(a)_r(a)_l \subseteq (a)_r \cap (a)_l$ is clear.

[1] The equivalence of conditions (i) and (ii) was proved for semigroups in S. LAJOS [6]. Proposition 2 is a sharpening of a result of L. KOVÁCS [5], J. CALAIS [1], respectively (see also [4]).

(ii)⇒(i). From $a \in (a)_r \cap (a)_l = (a)_r(a)_l \subseteq (a)_r A = aA$ it follows[2]) $(a)_r = aA$. Similarly $(a)_l = Aa$. These and condition (ii) imply $a \in aA \cap Aa = aA^2a$, that is a is a regular element.

(i)⇒(iii). As $a = axa$ is a regular element, the principal right ideal $(a)_r$ has an idempotent generator $ax = e$, that is $(a)_r = (e)_r$. This implies $e = e^2 \in (e)_r^2 = (a)_r^2$, whence $(a)_r = (e)_r = (e)_r^2 = (a)_r^2$. Similarly $(a)_l = (f)_l$ with $f^2 = f = xa$, whence

$$(a)_l^2 = (f)_l^2 = (f)_l = (a)_l.$$

Since $(a)_r = (e)_r = eA$ and $(a)_l = (f)_l = Af$, the intersection $eA \cap Af$ is a quasi-ideal of A. On the other hand, an arbitrary element $es = tf$ of $eA \cap Af$ $(s, t \in A)$ has the form $tf = tf \cdot f = es \cdot f \in eA \cdot Af$, whence $(a)_r(a)_l = (e)_r(f)_l = eA \cdot Af = eA \cap Af$ is a quasi-ideal of A.

(iii)⇒(ii). The condition (iii) implies $(a)_r = (a)_r^2 \subseteq (a)_r A = aA$. Hence $(a)_r = aA$. Similarly $(a)_l = Aa$. As $(a)_r(a)_l = aA \cdot Aa$ is a quasi-ideal of A, we get

$$aAAa \cdot A \cap A \cdot aAAa \subseteq aAAa.$$

This and the first part of condition (iii) imply

$$(a)_r \cap (a)_l = aA \cap Aa = (aA)^3 \cap (Aa)^3 \subseteq aA \cdot A \cdot aA \cap Aa \cdot A \cdot Aa \subseteq aA \cdot Aa =$$

$$= (a)_r(a)_l.$$

Since the inclusion $(a)_r(a)_l \subseteq (a)_r \cap (a)_l$ always holds, the proof is complete.

Now we want to characterize the regular duo elements of a ring (semigroup).

PROPOSITION 3. The following conditions on an element a of a ring (semigroup) A are equivalent:

(α) a is a regular duo element of A;

(β) it holds $(a)_l(a)_r = (a)_l \cap (a)_r$;

(γ) it holds $(a)_r^2 = (a)_l$ and $(a)_l^2 = (a)_r$.

PROOF. (α)⇒(β). As a is a duo element, $(a)_r = (a)_l$. This and condition (ii) in Proposition 2 imply (β).

(β)⇒(α). From $a \in (a)_l \cap (a)_r = (a)_l(a)_r$ it follows that the principal left ideal $(a)_l$ is contained in the two-sided ideal $(a)_l(a)_r$, whence

$$(a)_l \subseteq (a)_l(a)_r = (a)_l \cap (a)_r \subseteq (a)_r.$$

Similarly $(a)_r \subseteq (a)_l(a)_r = (a)_l \cap (a)_r \subseteq (a)_l$. Thus $(a)_l = (a)_r$ and a is a duo element. This and condition (β) imply the property (ii) of a, that is a is a regular element.

(α)⇒(γ). a being a duo element $(a)_l = (a)_r$ holds. This and the property (iii) of the regular element a imply (γ).

(γ)⇒(α). From $(a)_r = (a)_l^2 \subseteq (a)_l = (a)_r^2 \subseteq (a)_r$ it follows $(a)_r = (a)_l$, that is a is a duo element. The properties $(a)_r = (a)_l$ and (γ) imply (iii) of Proposition 2, i.e. a is regular.

From Propositions 1 and 3 one gets immediately

[2] If A is a semigroup then $(a)_r = a \cup aA$, and if A is a ring then $(a)_r = Ia + aA$, where I denotes the ring of integer numbers. In both cases $(a)_r A = aA$.

THEOREM 1. *A ring (a semigroup) A is a regular duo ring (semigroup) if and only if every element a of A has one of the properties* (α), (β) *and* (γ).

§ 3. In [9] we have proved

THEOREM 2. *The following assertions on a ring (semigroup) A are equivalent:*
(**a**) *A is a regular duo ring (semigroup);*
(**b**) *for any two quasi-ideals* Q_1, Q_2 *of A* $\quad Q_1Q_2 = Q_1 \cap Q_2$.

The following results show that the analogy between the regular duo rings and semigroups is not complete.

L. KOVÁCS [5] proved the interesting

THEOREM 3. *For a ring A the following conditions are equivalent:*
(I) *A is a regular ring without nonzero nilpotent elements;*
(II) *every quasi-ideal of A is idempotent;*
(III) *for every right ideal R and left ideal L of A*

$$RL = R \cap L \subseteq LR$$

holds.

It is known that conditions (I), (II), (III) and (**a**), (**b**) are equivalent for *a ring* (See B. M. ŠAIN [8], S. LAJOS and F. SZÁSZ [7] and A. KERTÉSZ and O. STEINFELD [4]). In the proof of this equivalence the following known result plays a rôle:

PROPOSITION 4. (See A. KERTÉSZ [3] Hilfssatz 3. 26.) *Every idempotent element of a ring A without nonzero nilpotent elements lies in the center of A.*

Remark 1. DR. R. WIEGANDT called my attention to the fact that the analogous result does not hold for semigroups:

Example. The semigroup F has three elements 0, a, b and its Cayley-table is

	0	a	b
0	0	0	0
a	0	a	a
b	0	b	b

Remark 2. The semigroup F defined in Remark 1 shows that the implication (I)$\Rightarrow$(**a**) does not hold for semigroups. Namely F is a regular semigroup without nonzero nilpotent elements, but it is not a duo semigroup.

A part of Theorem 3 has an analogue for semigroups.

THEOREM 3′. *For a semigroup S the following two assertations are equivalent:*
(II′) *every quasi-ideal of S is idempotent;*
(III′) *for every right ideal R and for every left ideal L of S*

$$RL = R \cap L \subseteq LR.$$

PROOF. One can prove the implication (II′) ⇒(III′) in the same way as (II) ⇒(III) is proved in Theorem 2 of L. KOVÁCS [5].

(III′) ⇒(II′). It is easy to show that every quasi-ideal Q of an arbitrary semigroup S is the intersection of the right ideal $R_1 = Q \cup QS$ and the left ideal $L_1 = Q \cup SQ$. Now we assume that the semigroup S has the property (III′). It holds

$$(*) \qquad Q = R_1 \cap L_1 = R_1 L_1 \subseteq L_1 R_1.$$

On the other hand, the first part of condition (III′) implies that S is a regular semigroup. (See L. KOVÁCS [5].) If a is an arbitrary element of Q, then $a = axa = axaxa$ with a suitable $x \in S$. This and condition $(*)$ imply

$$a = ax \cdot a \cdot xa \in QS \cdot Q \cdot SQ \subseteq R_1 \cdot L_1 R_1 \cdot L_1 = R_1 L_1 \cdot R_1 L_1 = Q^2,$$

whence $Q \subseteq Q^2$. Thus the quasi-ideal Q is idempotent.

Remark 3. The semigroup F defined in Remark 1 has four quasi-ideals $Q_0 = (0)$; $Q_1 = \{0, a\}$; $Q_2 = \{0, b\}$; $Q_3 = F$. Thus F has the property (II′), but it has not the property (**b**) of Theorem 2.

REFERENCES

[1] CALAIS, J.: Demi-groups quasi-inversifs, *Comptes Rendus de l'Acad. des Sciences* **252** (1961), 2357—2359.

[2] FELLER, E. H.: Properties of primary noncommutative rings, *Trans. Amer. Math. Soc.* **89** (1958), 79—81.

[3] KERTÉSZ, A.: *Vorlesungen über Artinsche Ringe*, Budapest, 1968.

[4] KERTÉSZ, A. und STEINFELD, O.: Über reguläre C-Ringe, *Beitr. Alg. Geom.* **3** (1973) (to appear).

[5] KOVÁCS, L.: A note on regular rings, Publ. Math. (Debrecen) **4** (1956), 465—468.

[6] LAJOS, S.: A félcsoportok ideálelméletéhez, MTA III. Oszt. Közl. **11** (1961), 57—66.

[7] LAJOS, S. and SZÁSZ, F.: Characterizations of strongly regular rings, II. *Proc. Japan Acad.* **46**; 3 (1970), 287—289.

[8] ŠAIN, B. M.: *o*-Rings and La-Rings. *Izv. Vyss. Učebn. Zaved. Matematika* 1966, no. 2(51), 111—122 (russian); *Amer. Math. Soc. Transl.* (2) Vol. 96, (1970), 137—152 (translated).

[9] STEINFELD, O.: Über die regulären duo-Elemente in Gruppoid-Verbände, Acta Sci. Math. **32** (1971), 327—331.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received August 27, 1971; revised September 10, 1972.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 165—168.

INDEPENDENT SETS IN CRITICAL CHROMATIC GRAPHS

by

L. LOVÁSZ

Let $\alpha_k(n)$ denote the maximum number of independent points in a critical k-chromatic graph on n vertices. W. G. BROWN and J. W. MOON [1] showed $\alpha_k(n) \geqq n - c_1\sqrt{n}$ for any $k \geqq 4$. M. SIMONOVITS [2] proved $\alpha_4(n) \leqq n - c_2 n^{2/5}$. In the present note we are going to show $n - 2kn^{\frac{1}{k-2}} \leqq \alpha_k(n) \leqq n - \frac{k}{6} n^{\frac{1}{k-2}}$ for $k \geqq 4$. The lower bound follows from a construction which is an improvement of that of BROWN's and MOON's. The proof of the upper bound is a modification of SIMONOVITS's proof, and shows an interesting connection to SPERNER's lemma in combinatorial topology.

Let $\tau(G)$ denote the minimum number of points covering every edge in the graph G. Then $|V(G)| - \tau(G)$ is the maximum number of independent points of G. It will be more convenient to formulate our arguments for $\tau(G)$.

A vertex (edge) of a graph G will be called *critical* if its removal decreases the chromatic number of G. A graph is *k-critical* if it has chromatic number k and every edge and vertex is critical.

THEOREM. *A k-critical graph G on n points has*

$$\tau(G) \geqq \frac{1}{6} kn^{\frac{1}{k-2}}.$$

On the other hand, there are infinitely many k-critical graphs G such that, putting $|V(G)| = n$,

$$\tau(G) = 2kn^{\frac{1}{k-2}}.$$

PROOF. Let T be a set of points of G covering all edges, $S = V(G) - T$, $|T| = t$. Following an idea of SIMONOVITS [2], we may assume S has vertices of degree $k-1$ only; for if $x \in S$ has degree $d \geqq k$, we may substitute $\binom{d}{k-1} = r$ points $x_1, \ldots, x_r$ for x, connecting them to $(k-1)$-tuples of the neighborhood of x. The obtained graph G' has chromatic number k, and all vertices but the x_i's are critical. Removing some of the x_i's without decreasing the chromatic number, we get a graph G'' in which all vertices are critical. However, in this graph the edges are also critical; this is trivial for the edges not incident with the x_i's but also follows for the x_i's since these points have degree $k-1$. Thus G'' is k-critical. It is obvious that the number of remaining x_i's is $\geqq 2$. Going on, we can split every vertex of S and construct a k-critical graph G_1 with $|V(G_1)| > |V(G)|$ and $\tau(G_1) \leqq \tau(G)$. If the theorem is valid for G_1 it also holds for G.

Let, now, H be the system of all $(k-1)$-tuples of T connected to the same point of S. Then H has the following property:

($*$) *For any $A \in H$, the vertices of T can be $(k-1)$-colored in such a way that A is the only $(k-1)$-tuple which gets all the $k-1$ colors.*

Really, remove the point corresponding to A and color the remaining graph by $k-1$ colors. Now it follows

($**$) *Whenever $A_1, \ldots, A_s \in H$, there is a $(k-2)$-tuple of points of T which is contained in an odd number of A_i's.*

For assume this is not the case and consider a coloring of the vertices of T by $(k-1)$ colors such that A_1 meets every color but $A_2, \ldots, A_s$ don't. Let us count incidences between A_i's and $(k-2)$-tuples of colors $1, 2, \ldots, k-2$ in two different ways. Since any $(k-2)$-tuple is contained in an even number of A_i's, the resulting number is even. On the other hand, A_1 contains exactly one $(k-2)$-tuple of colors $1, 2, \ldots, k-2$, the other A_i's contain 0 or 2. Hence the sum is odd, a contradiction.

Let $p = \binom{t}{k-2}$, $B_1, \ldots, B_p$ all $(k-2)$-tuples of T elements and define, for any $A \in H$, a vector $v_A = (v_1, \ldots, v_p)$ where

$$v_i = \begin{cases} 1 & \text{if} \quad B_i \subseteq A, \\ 0 & \text{if} \quad B_i \not\subseteq A. \end{cases}$$

Then ($**$) can be re-formulated as

($***$) *The vectors v_A $(A \in H)$ are linearly independent over $GF(2)$.*

Thence it follows

$$n - t = |S| = |H| \leqq \binom{t}{k-2},$$

hence

$$n \leqq \binom{t}{k-2} + t \leqq 2\binom{t}{k-2} \leqq t^{k-2} \cdot \frac{2}{(k-2)!} \leqq t^{k-2} \cdot \frac{3^{k-2}}{(k-2)^{k-2}},$$

whence

$$t \geqq \frac{k-2}{3} \cdot n^{\frac{1}{k-2}} \geqq \frac{k}{6} \cdot n^{\frac{1}{k-2}}.$$

Remarks: I. The vectors v_A don't span the whole p-dimensional space over $GF(2)$ but only a subspace; but this won't give any improvement even in the constant.

II. SIMONOVITS's proof used ($*$) but deduced only a smaller class of excluded partial systems — in fact, some special triangulations of the $(k-2)$-sphere. (The exclusion of all triangulated spheres would not give better result then SIMONOVITS's one; this follows from the results of BROWN, ERDŐS, and V. T. SÓS [3].) The idea of the present proof was that ($*$) follows for triangulated manifolds from SPERNER's lemma, moreover, the proof of SPERNER's lemma is applicable under the more general condition ($**$). It would be interesting to characterize all set-systems with property ($*$), and determine the maximum number of $(k-1)$-tuples in them. For $k=3$, graphs with property ($*$) are forests.

III. Assume we are given a system H of $(k-1)$-tuples on m points such that

(i) no $(k-1)$-coloring of the points yields exactly one $(k-1)$-tuple meeting all colors;

(ii) removing an $E_0 \in H$, the remaining system has property ($*$).

Then we can construct a k-critical graph G as follows: connect any two points of E_0; take a point x_E for any $E \in H$, $E \neq E_0$ and connect it to every point of E. The obtained graph G will have $\tau(G) \leqq m$, $|V(G)| = m + |H| - 1$. It would be interesting to know what is the maximum cardinality of H; may be it gives the best possible construction, but I cannot prove this for $k \geqq 4$. Therefore, another way is chosen below to construct a k-critical graph with many independent points.

Construction. Let G_i be an $(i+1)$-critical graph $(i = 2, \ldots, k-2)$ $V(G_i) = \{b_{i,1}, \ldots, b_{i,m_i}\}$. Let P be a $(k-1)$-critical graph, $V(P) = \{p_1, \ldots, p_{m_1}\}$. Let $B = \{b_{1,1}, \ldots, b_{1,m_1}\}$ be other points, $b_{1,i}$ joined to p_i. This anomaly for $i = 1$ is because there is only one 2-critical graph; the set $\{b_{1,1}, \ldots, b_{1,m_1}\}$, however, behaves like a 2-critical graph in the following sense: using $k-1$ colors, the set $\{b_{1,1} \ldots b_{1,m_1}\}$ must get at least two different colors; on the other hand, giving $b_{1,1}$ the color 1 and the other points color 2, this is a "good" coloring, i.e. it can be extended to a coloring of P.

Let A be the set of all vectors $(v_1, \ldots, v_{k-2})$ such that $1 \leqq v_i \leqq m_i$. Let $(v_1, \ldots, v_{k-2})$ be joined to b_{ij} if and only if $v_i = j$.

Finally, take a point d and connect it to every point of A.

We show the obtained graph G is not $(k-1)$-colorable. Assume indirectly it has a coloration by colors $1, \ldots, k-1$. Let d have color 1, say. One of the points in B, b_{11} say, has color different from 1, color 2, say. A point $b_{2,1}$ of G_2 has a third color, 3 say. And so on, we can select a point $b_{i,1}$ from every G_i which has different color from d and from the colors of $b_{1,1}, \ldots, b_{i-1,1}$. These points d, $b_{1,1}, \ldots, b_{k-2,1}$ have all the $k-1$ colors; but $(1, \ldots, 1) \in A$ is joined to each of them, which yields a contradiction.

As a second step we show that the edge $(d, (1, \ldots, 1))$ is critical. Remove it and consider the following $(k-1)$-coloring f of the remaining graph G'.

Let $f: V(G) \to \{1, \ldots, k-1\}$ be defined on $V(G_i)$ as an $(i+1)$-coloring such that

$$f(b_{i,1}) = i+1,$$

$$f(b_{i,j}) \leqq i \quad \text{for} \quad 2 \leqq j \leqq m_i;$$

let it be defined on $V(P)$ as a $(k-1)$-coloring such that

$$f(p_1) = 1,$$

$$f(p_j) \geqq 2 \quad \text{for} \quad 2 \leqq j \leqq m_1;$$

let

$$f(b_{1,1}) = 2,$$

$$f(b_{1,j}) = 1 \quad \text{for} \quad 2 \leqq j \leqq m_1;$$

$$f(d) = 1,$$

$$f((1, \ldots, 1)) = 1$$

and, finally,

$$f((v_1, \ldots, v_{k-2})) = i+1 \quad \text{where } i \text{ is the last index such that } v_i \neq 1.$$

Then f is a $k-1$-coloring of G'.

Similarly we get that all edges incident with d are critical. Thence, the points of A are critical. Since these points have degree $k-1$, all edges incident with them are critical.

Now we prove that the edges of G_i ($i = 2, \dots, k-2$) are critical. Remove one of them, then what remains from G_i can be colored by colors $1, \dots, i$. Color the graph spanned by $B \cup V(P)$ by $k-1$ colors so that B gets colors 1 and 2 only, and d by color 1. Now any point v of A is connected to $k-i-2$ points of the graphs $G_{i+1}, \dots, G_{k-2}$, and only these points can have colors $\geqq i+1$; therefore, there is a color which does not occur among the neighbors of v. This means that the coloring given above can be extended to A.

We argue similarly if an edge incident with P is removed. This proves G is k-critical.

The number of points of G is

$$n = m_1 \dots m_{k-2} + 2m_1 + m_2 + \dots + m_{k-2} + 1,$$

while

$$\tau(G) \leqq t = 2m_1 + m_2 + \dots + m_{k-2} + 1.$$

Taking $m > k$, $m_i = m$ or $m+1$ (this is possible, if $m_i \equiv i+1 \pmod 2$ $m_i \geqq i+1$ then a circuit of length $m_i - i + 2$ and $i-2$ points, joined to each other and to each point of the circuit, gives an $(i+1)$-critical graph on m_i points), we get

$$|V(G)| = n \geqq m^{k-2}$$

and

$$\tau(G) \leqq k(m+1) \leqq 2km \leqq 2k \cdot n^{\frac{1}{k-2}}.$$

This proves the theorem.

*

Acknowledgment. My thanks are due to M. SIMONOVITS for his valuable remarks.

REFERENCES

[1] BROWN, W. G. and MOON, J. W.: Sur les ensembles de sommets independants dans les graphes chromatiques minimaux, *Can. J. Math.* **21** (1969) 274—278.
[2] SIMONOVITS, M.: On colour-critical graphs, *Studia Sci. Math. Hungar.* **7** (1972), 67—81.
[3] BROWN, W. G., ERDÖS, P. and SÓS, V. T.: On the existence of triangulated spheres in 3-graphs, and related problems, to-appear.

Eötvös L. University, Budapest

(Received October 17, 1972.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 169—176.

ON THE CONSECUTIVE ZEROS OF THE SOLUTIONS OF LINEAR SECOND ORDER DIFFERENTIAL EQUATIONS

by

Á. ELBERT

The subject of this paper is some functionals related with the function $p(t)$ and with the solution $y(t)$ of the differential equation

$$y''+p(t)y = 0 \tag{1}$$

with the boundary conditions $y(0)=y(T)=0$, where $0, T>0$ are its consecutive zeros and $p(t)$ real nonnegative continuous function. Two types of such functionals have been investigated up to now. These are the Lyapunov functional

$$T\int_0^T p(t)dt$$

connected with the famous Lyapunov inequality and

$$\int_0^T \sqrt{p(t)}dt.$$

Assuming suitable conditions concerning $p(t)$ several inequalities are known for this two functionals. To formulate these conditions we give a definition.

Definition. The function $p(t)$ belongs to $C_\nu[a, b]$ if $p(t)\geqq 0$ and in the case $\nu\neq 0$ the function $\operatorname{sgn}\nu\cdot p^\nu(t)$ or in the case $\nu=0$ the function $\log p(t)$ is concave on $[a, b]$.

A. S. GALBRAITH [1] proved that if $0, T$ are two consecutive zeros of a solution of (1) and $p(t)\in C_1[0, T]$ then

$$T\int_0^T p(t)dt \leqq \pi^2. \tag{2}$$

Concerning the second functional E. MAKAI [2] proved

$$\int_0^T \sqrt{p(t)}dt \leqq \pi \tag{3}$$

if $p(t)\in C_{-1/4}[0, T]$. On the other hand D. BANKS [3] showed that the inequality

$$T\int_0^T p(t)dt \geqq \frac{9}{2} j^2_{-2/3} \tag{4}$$

holds if $p \in C_1[0, T]$. Here, and later too, j_μ denotes the first positive root of the Bessel function $J_\mu(z)$ of order μ. Finally, in [4] we can find among others the lower bound

$$\int_0^T \sqrt{p(t)}\,dt \geqq 2j_{\mu-1} \tag{5}$$

if $p(t) \in C_\nu[0, T]$, $0 < \nu < \infty$, and $\mu = \nu/(2\nu+1)$.

Now we shall generalize the form of the functionals.

Let $F:(a, \infty) \to (0, \infty)$ be a strictly increasing convex function, where $a \leqq 0$. Its inverse function is F^{-1}.

Let the new functional defined by

$$I_F(\tau, p) = \tau^2 F\left(\frac{1}{\tau}\int_0^\tau F^{-1}(p(t))\,dt\right) \quad \text{for} \quad \tau > 0. \tag{6}$$

THEOREM 1. *If* $p(t) \geqq F(0)$, $F^{-1}(p(t))$ *is concave, then*

$$I_F(T, p) \leqq \pi^2.$$

The equality holds only if $p(t) = \text{const.}$

PROOF. First of all we mention, that under the conditions of the theorem $p(t)$ is continuous on $(0, T)$ and bounded from above and the integral required in (6) exists.

Here, and later too, we need use the Raleigh principle, therefore it seems to be useful to put it down here. The least positive eigenvalue $\lambda(p^*)$ of the boundary value problem

$$y'' + \lambda p^*(t) y = 0 \tag{1*}$$

with the boundary conditions $y(0) = y(T) = 0$ fulfils the relation

$$\frac{1}{\lambda(p^*)} = \max_{u \in U} \frac{\int_0^T p^*(t) u^2(t)\,dt}{\int_0^T u'^2(t)\,dt}, \tag{7}$$

where u denotes the set of piecewise continuously differentiable functions on $[0, T]$ with the conditions $u(0) = u(T) = 0$ and the equality sign takes place only when $u = y(t)$ is the solution of (1*) with $\lambda = \lambda(p^*)$ satisfying the boundary conditions.

According to our hypothesis the function $F^{-1}(p(t))$ is nonnegative and concave. Therefore it can be written in the form

$$F^{-1}(p(t)) = \int_0^T p_\xi(t)\,d\Gamma_p(\xi) \tag{8}$$

and

$$\int_0^T F^{-1}(p(t))\,dt = \Gamma_p(T+0) - \Gamma_p(-0), \tag{9}$$

where

$$p_\xi(t) = \begin{cases} \dfrac{2}{T}\dfrac{t}{\xi} & \text{for} \quad 0 \leqq t \leqq \xi \\ \dfrac{2}{T}\dfrac{T-t}{T-\xi} & \text{for} \quad \xi \leqq t \leqq T. \end{cases} \tag{10}$$

(9) follows from (8) because of $\int_0^T p_\xi(t)\,dt = 1$, and the function $\Gamma_p(\xi)$ is non-decreasing on $[0, T]$. If $q(t) = F^{-1}(p)$ is twice differentiable then

$$\Gamma_p(\xi) = \begin{cases} 0 & \text{for} \quad \xi < 0 \\ \dfrac{T}{2}q(0) + \int_0^\xi \dfrac{1}{2}s(T-s)[-q''(s)]\,ds & \text{for} \quad 0 \leqq \xi < T \\ \dfrac{T}{2}q(0) + \int_0^T \dfrac{1}{2}s(T-s)[-q''(s)]\,ds + \dfrac{T}{2}q(T) & \text{for} \quad \xi \geqq T, \end{cases}$$

and the representation (8) can be proved by integration by parts. Otherwise $q(t)$ is the limit of a sequence of twice differentiable functions, each of them fulfils (8), hence there is a limit function $\Gamma_p(\xi)$ which represents $q(t)$ by (8).

We take $u(t) = \sin(\pi t/T)$ in (7), then by (8)

$$\pi^2 \geqq 2T \int_0^T p(t) \sin^2 \frac{\pi t}{T}\,dt = T^2 \int_0^T F\Big(\int_0^T p_\xi(t)\,d\Gamma_p(\xi)\Big) \frac{2}{T} \sin^2 \frac{\pi t}{T}\,dt.$$

Hence using the convexity of $F(u)$ and the facts that $\sin^2(\pi t/T) \geqq 0$, and $(2/T) \cdot \int_0^T \sin^2(\pi t/T)\,dt = 1$ we have

$$\pi^2 \geqq T^2 F\left(\int_0^T \left[\int_0^T p_\xi(t)\,d\Gamma_p(\xi)\right] \frac{2}{T} \sin^2 \frac{\pi t}{T}\,dt\right) =$$

$$= T^2 F\left(\int_0^T \left[\int_0^T p_\xi(t)\,\frac{2}{T} \sin^2 \frac{\pi t}{T}\,dt\right] d\Gamma_p(\xi)\right).$$

Now by (10)

$$\int_0^T p_\xi(t)\,\frac{2}{T} \sin^2 \frac{\pi t}{T}\,dt = \frac{1}{T}\left[1 + \frac{\sin^2 \dfrac{\pi\xi}{T}}{\pi^2 \dfrac{\xi(T-\xi)}{T^2}}\right] \geqq \frac{1}{T},$$

F is increasing, thus by (9)

$$\pi^2 \geqq T^2 F\left(\int_0^T \frac{1}{T}\,d\Gamma_p(\xi)\right) = I_F(T, p).$$

The sign of equality can occur only when $y(t) = u(t) = \sin(\pi t/T)$, i.e. when $p(t) \equiv \equiv \pi^2/T^2$. But in this case we have $I_F(T, \pi^2/T^2) = \pi^2$, which completes the proof.

Choosing $F(u)=u$ in Theorem 1 we have (2). The inequality (3) can be proved by choosing $F(u)=u^2$ only for the subset $C_{1/2}[0, T]\subset C_{-1/4}[0, T]$.

In order to obtain lower bound for the functional $I_F(T, p)$ we must specify the function $F(u)$. The following cases will be treated

$$F_m(u) = \left(m^{\frac{1}{\alpha}}+u\right)^{\alpha} \quad m \geqq 0,\ \alpha \geqq 1 \tag{11}$$

$$F_m(u) = me^u \qquad m > 0. \tag{12}$$

Here the conditions on m, α are explained by the requirement that $F_m(u)$ must be convex and increasing.

For the sake of simplicity we shall drop the index m from the side of F when it may not cause confusion.

The significant property of these functions is that there is a function $\varphi(u)=u^{1/\alpha}$ or $\varphi(u)=\log u$, resp., satisfying

$$F_m^{-1}(u) = \varphi(u)-\varphi(m) \quad \text{and} \quad F_m(u) = \varphi^{-1}(u+\varphi(m)). \tag{13}$$

The functional $F_m(\tau, p(t))$ does not depend on m, since by (13)

$$\frac{1}{\tau}\int_0^{\tau} F_m^{-1}(p)\,dt = \frac{1}{\tau}\int_0^{\tau} \varphi(p(t))\,dt-\varphi(m) \tag{14}$$

hence by the second relation in (13)

$$I_{F_m}(\tau, p) = \tau^2\varphi^{-1}\left(\frac{1}{\tau}\int_0^{\tau} \varphi(p(t))\,dt\right).$$

Thus our functionals are of the form

$$\tau^{2-\alpha}\left[\int_0^{\tau} p^{\frac{1}{\alpha}}(t)\,dt\right]^{\alpha} \quad \text{and} \quad \tau^2 e^{\frac{1}{\tau}\int_0^{\tau}\log p(t)\,dt}.$$

Now the functional $I_F(T, p(t))$, where $t=0$ and $t=T$ are consecutive zeros of a solution $y(t)$ of (1), is invariant with respect to any linear transformation of the independent variable. Indeed, if $t=cx$, $Y(x)=y(cx)$, then $Y(x)$ fulfils the differential equation

$$Y''+c^2p(cx)Y = 0 \tag{15}$$

with the boundary conditions $Y(0)=Y(T/c)=0$. Now $F_{mc^2}^{-1}(c^2p)$ is concave, hence by (13), (14)

$$I_{F_{mc^2}}\left(\frac{T}{c},\ c^2p(cx)\right) = \frac{T^2}{c^2}\varphi^{-1}\left(\frac{c}{T}\int_0^{T/c} \varphi(c^2p(cx))\,dx\right) =$$

$$= \frac{T^2}{c^2}\varphi^{-1}\left(\frac{1}{T}\int_0^{T} \varphi(c^2p(t))\,dt\right).$$

By the special form of φ we obtain immediately

$$I_{F_{mc^2}}\left(\frac{T}{c},\ c^2p(cx)\right) = T^2\varphi^{-1}\left(\frac{1}{T}\int_0^T \varphi(p(t))\,dt\right) = I_{F_m}(T, p),$$

as we stated.

Applying Theorem 1 we have for $p(t) \equiv F_m(0) = m$

$$\pi^2 \geqq T^2 F_m\left(\frac{1}{T}\int_0^T F^{-1}(p)\,dt\right) \geqq T^2 m,$$

where the equality holds only if $p(t) \equiv m$, hence in all other cases

$$p(t) > m \quad \text{for} \quad 0 < t < T \tag{16}$$

and

$$\pi^2 > mT^2. \tag{17}$$

Let us consider the differential equation

$$y'' + F_m(\mu p_{T/2}(t))y = 0 \quad (\mu \geqq 0), \tag{18}$$

where $p_{T/2}(t)$ is given by (10). Let $y_\mu(t)$ be a solution of (18) with the initial conditions $y_\mu(0)=0$, $y_\mu'(0)\neq 0$. If $\mu=0$ or small enough then $y_\mu(t)$ has no more zeros on $[0, T]$. If μ tends to $+\infty$ then $F_m(\mu p_{T/2})$ tends to $+\infty$ on a set of positive measure, hence $y_\mu(t)$ has many zeros on $(0, T)$. Therefore there is $\mu_1 = \mu(m, T)$ satisfying $y_{\mu_1}(T)=0$ and $y_{\mu_1}(t)$ has no zeros on $(0, T)$. Let

$$\Lambda(m, T) = I_{F_m}(T, F_m(\mu_1 p_{T/2})). \tag{19}$$

The function $\Lambda(m, T)$ depends only on mT^2. To prove this we use the transformation $t = Tx$ in (18). Then by (15) with $c = T$ and using the above mentioned invariance of the functional I_F we get

$$\Lambda(m, T) = \Lambda(mT^2, 1) = \Lambda(mT^2). \tag{20}$$

Now we are able to formulate our next theorem.

THEOREM 2. *If $F_m(u)$ is given by* (11) *or* (12) *and $F_m^{-1}(p(t))$ is nonnegative and concave then*

$$I_F(T, p) \geqq \Lambda(mT^2).$$

The equality holds if and only if $p(t) = F_m(\mu(m, T)p_{T/2}(t))$.

PROOF. As we have seen in Theorem 1, the case $p(t) \equiv$ const gives the maximum for $I_{F_m}(T, p)$. Now we are interested in finding the minimum of this functional, therefore $p(t)$ fulfils (16) and (17). Define γ by

$$\gamma = \int_0^T F_m^{-1}(p(t))\,dt = \int_0^T d\Gamma_p(\xi) > 0, \tag{21}$$

and so by (8)

$$p(t) = F_m\left(\int_0^T p_\xi(t)\,d\Gamma_p(\xi)\right) = F_m\left(\int_0^T \gamma p_\xi(t)\frac{d\Gamma_p(\xi)}{\gamma}\right) \leqq$$

$$\leqq \int_0^T F_m(\gamma p_\xi(t))\frac{d\Gamma_p(\xi)}{\gamma}.$$

Using (7) with $\lambda(p)=1$

$$1 \leqq \max_{u\in U} \frac{\int_0^T \left[\int_0^T F(\gamma p_\xi(t))\,d\Gamma_p(\xi)\right] u^2(t)\,dt}{\gamma \int_0^T u'^2(t)\,dt} \leqq$$

$$\leqq \frac{1}{\gamma}\int_0^T \max_{u\in U} \frac{\int_0^T F(\gamma p_\xi(t))u^2(t)\,dt}{\int_0^T u'^2(t)\,dt}\,d\Gamma_p(\xi) = \frac{1}{\gamma}\int_0^T \frac{d\Gamma_p(\xi)}{\lambda(F(\gamma p_\xi))}.$$

The functions $F(\gamma p_\xi(t))$ are equimeasurable for $0\leqq\xi\leqq T$, i.e. the Lebesgue measures $m\{t; F(\gamma p_\xi(t))\geqq y\}$ do not depend on ξ. According to a theorem due to BEESACK and SCHWARZ [5] the eigenvalue $\lambda(F(\gamma p_\xi))$ takes on the least value if the function $F(\gamma p_\xi(t))$ is symmetric increasing-decreasing one. This happens only in the case $\xi=T/2$, therefore

$$\lambda_1 = \lambda(F(\gamma p_{T/2})) \leqq 1.$$

We want to show, that $\gamma\geqq\mu_1$. Since λ_1 is the least eigenvalue of the boundary value problem with respect to the differential equation

$$y''+\lambda F_m(\gamma p_{T/2}(t))y = 0,$$

there exists a solution $\tilde{y}(t)$ with $\lambda=\lambda_1$ satisfying the boundary conditions $\tilde{y}(0)=$ $=\tilde{y}(T)=0$. If $\lambda_1=1$ then by the definition of μ_1 $\gamma=\mu_1$. If $\lambda_1<1$ then by the Sturmian comparison theorem comparing $\tilde{y}(t)$ with the solution $y_\gamma(t)$ of (18) with $\mu=\gamma$ has at least two zeros on $[0, T)$, hence $\gamma>\mu_1$.

By (21) and (19), (20) the inequality $\gamma\geqq\mu_1$ implies the required inequality

$$I_F(T,p) = T^2F\left(\frac{1}{T}\int_0^T F^{-1}(p)\,dt\right) = T^2F\left(\frac{\gamma}{T}\right) \geqq$$

$$\geqq T^2F\left(\frac{\mu_1}{T}\right) = I_F(T, F(\mu_1 p_{T/2})) = \Lambda(mT^2).$$

Looking over the proof we can realise that the equality here can occur only in the case $p(t)=F_m(\mu_1 p_{T/2}(t))$, which completes the proof.

Now it becames interesting the behaviour of the function $\Lambda(\xi)$. The next theorem gives some information about this.

THEOREM 3. *The function* $\Lambda(\xi)$ *is strictly increasing on* $[0, \pi^2]$, $\Lambda(\pi^2)=\pi^2$ *and if* $F(u) = (m^{1/\alpha}+u)^\alpha$ *then*

$$\Lambda(\xi) \geqq \Lambda(0) = 2^{-\alpha}(\alpha+2)^2 j^2_{\frac{1}{2+\alpha}-1}, \tag{22}$$

if $F(u)=me^u$, *then*

$$\Lambda(\xi) \geqq \pi\sqrt{\xi}. \tag{23}$$

PROOF. Let m, T be fixed, $m_1<m$. By (19), (20) we must consider the function $F_m(\mu_1 p_{T/2}(t))$, where $\mu_1=\mu(m, T)$ then by (10), (13)

$$\Lambda(mT^2) = T^2 F_m\left(\frac{1}{T}\int_0^T \mu_1 p_{T/2}(t)dt\right) = T^2 F_m\left(\frac{\mu_1}{T}\right) = T^2\varphi^{-1}\left(\varphi(m)+\frac{\mu_1}{T}\right). \tag{24}$$

Using (13) we have

$$F_{m_1}^{-1}(u) - F_m^{-1}(u) = \varphi(m) - \varphi(m_1),$$

hence

$$F_{m_1}^{-1}(F_m(\mu_1 p_{T/2}(t))) = \mu_1 p_{T/2}(t) + \varphi(m) - \varphi(m_1),$$

and this function is positive, (by $\varphi(m)-\varphi(m_1) > 0$) and concave. Theorem 2 gives

$$I_{F_{m_1}}(T, F_m(\mu_1 p_{T/2})) > \Lambda(m_1 T^2),$$

but (14) implies that $I_{F_{m_1}}=I_{F_m}$, hence $\Lambda(mT^2)>\Lambda(m_1T^2)$, i.e. $\Lambda(\xi)$ is strictly increasing.

From (24) we have $\Lambda(mT^2)\geqq mT^2$. Applying this inequality with $mT^2=\pi^2$ and Theorem 1, it follows $\Lambda(\pi^2)=\pi^2$.

We must determine $\Lambda(0)$ in (22). Now $m=0$, $F(u)=u^\alpha$, and $\varphi(u)=u^{1/\alpha}$. To determine $\mu_1=\mu(0, T)$ we must solve the boundary value problem concerning the differential equation

$$y'' + \mu_1^\alpha p_{T/2}^\alpha(t) y = 0$$

with the boundary conditions $y(0)=y(T)=0$. On account of the symmetry of $p_{T/2}(t)$, $y'(T/2)=0$. The solution on $[0, T/2]$ is

$$y(t) = C\sqrt{t} J_{\frac{1}{2+\alpha}}(z)$$

where $z = \int_0^t \sqrt[\alpha]{\mu_1^\alpha p_{T/2}^\alpha(t)}d\tau = \sqrt[\alpha]{\mu_1^\alpha}\left(\frac{2}{T}\right)^{\frac{\alpha}{2}} t^{\frac{\alpha}{2}+1}\Big/\left(\frac{\alpha}{2}+1\right)$, C is arbitrary constant and

$$y'(t) = C\frac{\alpha+2}{2}\frac{z}{\sqrt{t}} J_{\frac{1}{2+\alpha}-1}(z).$$

The requirement $y'(T/2)=0$ is equivalent to

$$\mu_1^\alpha = \left(\frac{\alpha+2}{2}\right)^2 j^2_{\frac{1}{2+\alpha}-1}\left(\frac{2}{T}\right)^{2-\alpha},$$

thus by (24) we have

$$\Lambda(0) = T^2 \frac{\mu_1^\alpha}{T^\alpha} = 2^{-\alpha}(\alpha+2)^2 j^2_{\frac{1}{2+\alpha}-1},$$

in accordance with (22).

Let $F(u)=me^u$, then $\varphi(u)=\log u$. From the boundary value problem concerning the differential equation (18) with $\mu_1=\mu(m, T)$ we have by (7) and (10)

$$1 = \max_{u\in U} \frac{\int_0^T F_m(\mu_1 p_{\frac{T}{2}}(t))u^2(t)\,dt}{\int_0^T u'^2(t)\,dt} \leqq F_m\left(\mu_1 p_{\frac{T}{2}}\left(\frac{T}{2}\right)\right) \max_{u\in U} \frac{\int_0^T u^2(t)\,dt}{\int_0^T u'^2(t)\,dt} =$$

$$= me^{2\frac{\mu_1}{T}}\frac{T^2}{\pi^2} = \frac{\xi}{\pi^2}e^{2\frac{\mu_1}{T}},$$

hence by (24)

$$\Lambda(\xi) = \xi e^{\frac{\mu_1}{T}} \geqq \pi\sqrt{\xi},$$

thus the proof is complete.

In the special case $F(u)=u$ we have $\alpha=1$ and the concavity in common sense, hence by Theorem 2 and 3 we have

$$T\int_0^T p(t)\,dt \geqq \Lambda(0) = \frac{9}{2} j^2_{-\frac{2}{3}},$$

as in (4). Choosing $F(u)=u^2$ then $\alpha=2$ and the condition of the concavity of $F^{-1}(p)$ is equivalent to the condition $p\in C_{1/2}$. Then Theorem 2 and 3 imply

$$\int_0^T \sqrt{p}\,dt \geqq 2j_{-\frac{3}{4}}$$

which is a special case of (5).

REFERENCES

[1] GALBRAITH, A. S.: On the zeros of solutions of ordinary differential equation of the second order. *Proc. Amer. Math. Soc.* **17** (1966), 333—337.

[2] MAKAI, E.: Über die Nullstellen von Funktionen, die Lösungen Sturm-Liouville'scher Differentialgleichungen sind. *Comm. Math. Helv.* **16** (1943—44), 153—199.

[3] BANKS, D.: Bounds for the eigenvalues of some vibrating systems. *Pacific J. Math.* **10** (1960) 439—474.

[4] ELBERT, Á.: On the solutions of the differential equation $y''+q(x)y=0$, where $[q(x)]^\nu$ is concave, II. *Studia Sci. Math. Hung.* **4** (1969), 257—266.

[5] BEESACK, P. R., SCHWARZ, B.: On the zeros of solutions of second order linear differential equation. *Canadian J. Math.* **8** (1956), 504—515.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received November 13, 1972.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 177—185.

DISTRIBUTION OF ZEROS OF SOLUTIONS OF $y'' = q(t)y$ IN RELATION TO THEIR BEHAVIOUR IN LARGE

by

F. NEUMAN

1. In this paper relations between the distribution of zeros of solutions of a differential equation

$$(q) \qquad y'' = q(t)y, \quad q \in C^0[a, b), \quad b \leqq \infty,$$

and the behaviour of the solutions of (q) for $t \to b$ are studied. We give not only a survey of results of this field derived in [1], [11], [12] and [13], but we generalize some of them, prove some results mentioned in footnotes without proofs (in [13]), and, at the same time, simplify some of the previous proofs by a unified method showing the basic ideas of the approach to problems of similar character.

2. Considering

$$(*) \qquad y'' = 0 \quad \text{on} \quad [0, \infty)$$

and

$$(**) \qquad y'' = y \quad \text{on} \quad [0, \infty)$$

we see that every solution of both differential equations has at most one zero on $[0, \infty)$. A t the same time, the behaviour of the solutions of $(*)$ and $(**)$ are different, since e.g. there is no non-trivial solution of $(*)$ tending to zero for $t \to \infty$, however, $(**)$ has such a solution.

This is why in this article we shall suppose differential equations (q) to be oscillatory when studying relations between the distribution of zeros and the behaviour of solutions.

Note 1. It is possible to handle differences between cases similar to $(*)$ and $(**)$, however for it there is better to use a smoother tool than the distribution of zeros of solutions (or dispersion), namely the phases of (q) (introduced in [5]).

3. Basic notions and relations (for proofs see [5] (or [14])).

Let (q) denote a differential equation

$$(q) \qquad y'' = q(t)y, \quad q \in C^0[a, b), \quad b \leqq \infty,$$

$C^n[a, b)$ being the set of all continuous functions having continuous derivatives up to and including the order n ($n \geqq 0$) on $a \leqq t < b$.

For $t_0 \in [a, b)$, let y be a non-trivial solution of (q) vanishing at t_0, and let $\varphi(t_0)$ be the first zero of y lying on the right of t_0. The function φ is called basic central dispersion of the 1st kind (briefly "dispersion") of (q). Every such dispersion satisfies

$$(1) \qquad \varphi(t) > t, \quad \varphi \in C^3, \quad \varphi'(t) > 0$$

on its interval of definition (which may be empty).

Let u_1, u_2 be two linearly independent solutions of (q). A function $\alpha \in C^0[a, b)$ satisfying

$$\tan \alpha(t) = u_1(t)/u_2(t) \quad \text{for all} \quad t \in \{t \in [a, b);\ u_2(t) \neq 0\}$$

is called a (1st) phase of (q) (corresponding to the pair (u_1, u_2)). It was proved that such a function exists, and that every such a function α then satisfies

$$\alpha \in C^3[a, b), \quad \alpha'(t) \neq 0 \quad \text{for all} \quad t \in [a, b) \tag{2}$$

$$q(t) = -\{\alpha, t\} - \alpha'^2(t), \text{ where } \{\alpha, t\} = \frac{1}{2}\alpha'''(t)/\alpha'(t) - \tag{3}$$

$$-\frac{3}{4}(\alpha''(t)/\alpha'(t))^2 \text{ is the Schwartz derivative of } \alpha \text{ at } t,$$

and the general solution of (q) can be written in the form

$$y(t;\, c_1, c_2) = c_1 \frac{\sin[\alpha(t) + c_2]}{\sqrt{|\alpha'(t)|}}, \tag{4}$$

where c_1, c_2 are arbitrary constants.

Between any phase α and the dispersion φ of the same differential equation the following Abel equation holds

$$\alpha(\varphi(t)) = \alpha(t) + \pi \operatorname{sign} \alpha'. \tag{5}$$

A differential equation (q) is called oscillatory on $[a, b)$ iff any (every) non-trivial solution of (q) has infinitely many zeros on $[a, b)$. Define $\varphi_0(t) \equiv t$, $\varphi_n(t) \equiv \equiv \varphi(\varphi_{n-1}(t))$. Then it holds:

A differential equation (q) is oscillatory on $[a, b)$ iff its dispersion φ is defined on the same interval $[a, b)$. Moreover, if (q) is oscillatory, then φ_n is defined on $[a, b)$ for every integer $n \geqq 0$,

$$\begin{gathered} \varphi_n \in C^3[a, b), \ d\varphi_n(t)/dt > 0 \quad \text{on} \quad [a, b), \ \varphi_n(t) > \varphi_{n-1}(t), \text{ and} \\ \lim_{n \to \infty} \varphi_n(t_0) = b \quad \text{for every} \quad t_0 \in [a, b). \end{gathered} \tag{6}$$

For real $p > 0$, denote $L^p[a, b) = \left\{f \in C^0[a, b);\ \int_a^b |f(t)|^p dt < \infty\right\}$.

4. We shall prove the following

THEOREM 1. *Let* (q) *and* ($\bar{\text{q}}$) *be oscillatory on* $[a, b)$, $b \leqq \infty$, *having the same dispersion* φ. *If any of the following assertions holds for* (q), *then the same is true for* ($\bar{q}$):

A1. *Every solution is bounded.*

A2. *The set of the bounded solutions consists of a one-parametric family (of the the form* $c \cdot y$*), every other solution is unbounded.*

A3. *No non-trivial solution is bounded.*

B1. *Every solution tends to zero as* $t \to b$.

B2. *The set of the solutions tending to zero as* $t \to b$ *consists of a one-parametric family of the form* $c.y.$

B3. *No non-trivial solution tends to zero as* $t \to b$

C1. *Every solution belongs to* $L^p[a, b)$,

C2. *The set of solutions belonging to* $L^p[a, b)$ *consists of a one-parametric family,*

C3. *No non-trivial solution belongs to* $L^p[a, b)$,

($p > 0$ *is a given real number).*

Whenever a one-parametric family of solutions is mentioned, the family is formed by dependent solutions, and the family has the same set of zeros of their solutions for both (q) *and* $(\bar{q})$.

Note 2. Similar results can be proved following our futher considerations, if instead of boundedness, etc. such notions as exponential or other growth, or stability, or different types of convergence are introduced.

Note 3. Hence all the types of behaviour in large of solutions of (q) mentioned in Theorem 1 and Note 2 are in the oscillatory case uniquely determined by the distribution of zeros of those solutions.

Let us give also some necessary and/or sufficient conditions on φ ensuring that, for a differential equation (q) having φ as its dispersion, a case listed in Theorem 1 occurs.

THEOREM 2. *Let* (q) *be oscillatory on* $[a, b)$, $b \leqq \infty$, *and* φ *be its dispersion.*

a) *If* $x_0 \in [a, b)$ *and an integer* $k > 0$ *exist such that* $d\varphi_k(t)/dt \leqq 1$ *on* $[x_0, b)$, *then* A1 *holds for* (q).

b) $d\varphi_n(t)/dt$ *is bounded for every integer* $n \geqq 0$ *and every* $t \in [a, \varphi(a)]$ *iff* A1 *holds for* (q).

c) *If the sequence* $\{d\varphi_n(t)/dt\}_{n=1}^{\infty}$ *is unbounded at least for two different values from* $(a, \varphi(a))$, *then* A3 *holds for* (q).

d) *If* $x_0 \in [a, b)$ *and an integer* $k > 0$ *exist such that* $d\varphi_k(t)/dt \geqq c > 1$ *on* $[x_0, b)$, *then* A3 *holds for* (q).

e) *If* $x_0 \in [a, b)$ *and an integer* $k > 0$ *exist such that* $d\varphi_k(t)/dt \leqq c < 1$ *on* $[x_0, b)$, *then* B1 *takes place for* (q).

f) $\lim_{n \to \infty} d\varphi_n(t)/dt = 0$ *uniformly on* $[a, \varphi(a))$, *iff* B1 *occurs for* (q).

g) *If the sequence* $\{d\varphi_n(t)/dt\}_{n=0}^{\infty}$ *does not tend to zero as* $n \to \infty$ *at least for two different values from* $[a, \varphi(a))$, *then* B3 *holds for* (q).

h) $\sum_{n=0}^{\infty} \gamma_n^{(p)} < \infty$, *where* $\gamma_n^{(p)} = \int_a^{\varphi(a)} [d\varphi_n(t)/dt]^{1+p/2} dt$, *iff* C1 *holds for* (q).

i) *There exists an integer* $k > 0$ *such that* $d\varphi_k(t)/dt \equiv 1$ *on* $[a, b)$ *(then* $b = \infty$*), iff every solution of* (q) *is periodic (i.e. there exists* $d > 0$ *such that* $y(t+d) = y(t)$ *for solution* y *and every* $t \in [a, b) = [a, \infty)$*).*

Note 4. The condition $d\varphi_k(t)/dt \leqq 1$ in Theorem 2, case a, can be expressed as

$$\varphi_k(t) - t \quad \text{is a non-increasing function on} \quad [x_0, b),$$

or, in other words,

if for sufficient large t, the difference between every zero t and its k-th successive zero on the right of t, is a non-increasing function of t, then every solution of (q) is bounded.

Similarly, Theorem 2, case i, can be rewritten as

the necessary and sufficient condition for the existence of only periodic solutions of an oscillatory differential equation, is that the differences between any zero t and its k-th successive zero on the right from it, are constant for some integer $k>0$ and every t.

Analogous results hold for cases d and e of Theorem 2.

5. Proofs

Theorem 1 is an immediate consequence of the following basic lemma, first given in [13].

LEMMA. *Let* (q) and $(\bar{q})$ *be oscillatory differential equations on* $[a, b)$, $b \leqq \infty$, *with the same dispersion* φ. *For every solution* y *of* (q) *there always exist a solution* $\bar{y}$ *of* $(\bar{q})$ *and constants* $m>0$, M *such that*

$$m|\bar{y}(t)| \leqq |y(t)| \leqq M|\bar{y}(t)| \quad \text{for every} \quad t \in [a, b). \tag{7}$$

We are going to give the proof of the lemma, since it demonstrates our approach to problems of this kind. It differs in some aspects and is shorter than the previous one in [13].

Let y be a non-trivial solution of (q), and let t_0 denote its smallest zero. Denote by $\bar{y}$ a non-trivial solution of $(\bar{q})$ having a zero at the t_0. The sets of zeros of y and $\bar{y}$ are identical, since the dispersion φ is the same for both (q) and $(\bar{q})$. We set

$$f(t) \overset{\text{def}}{=} \left| \frac{y(t)}{\bar{y}(t)} \right| \quad \text{for} \quad \bar{y}(t) \neq 0, \quad f(t) \overset{\text{def}}{=} \left| \frac{y'(t)}{\bar{y}'(t)} \right| \quad \text{for} \quad \bar{y}(t) = 0$$

(hence $\bar{y}'(t) \neq 0$).

Evidently f is continuous and positive on $[a, b)$. Since $\lim_{n \to \infty} \varphi_n(t_0) = b$, we have

$$[t_0, b) = \bigcup_{n=0}^{\infty} [\varphi_n(t_0), \varphi_{n+1}(t_0)), \tag{8}$$

and for every $t \in [t_0, b)$ there exist $x \in [t_0, \varphi(t_0))$ and an integer n such that $t = \varphi_n(x)$. This t is necessarily in $[\varphi_n(t_0), \varphi_{n+1}(t_0))$. Hence, according to (4) and (5), if α and $\bar{\alpha}$ are phases of (q) and $(\bar{q})$ resp., then

$$f(t) = \left| \frac{c_1 \sin [\alpha(t) + c_2]}{\sqrt{|\alpha'(t)|}} \cdot \frac{\sqrt{|\bar{\alpha}'(t)|}}{\bar{c}_1 \sin [\bar{\alpha}(t) + \bar{c}_2]} \right| =$$

$$= \left| \frac{c_1}{\bar{c}_1} \right| \cdot \sqrt{\left| \frac{\bar{\alpha}'(\varphi_n(x))}{\alpha'(\varphi_n(x))} \right|} \cdot \left| \frac{\sin [\alpha(\varphi_n(x)) + c_2]}{\sin [\bar{\alpha}(\varphi_n(x)) + \bar{c}_2]} \right| =$$

$$= \left| \frac{c_1}{\bar{c}_1} \right| \cdot \sqrt{\left| \frac{\bar{\alpha}'(x) \cdot d\varphi_n(x)/dx}{\alpha'(x) \cdot d\varphi_n(x)/dx} \right|} \cdot \left| \frac{\sin [\alpha(x) \pm n\pi + c_2]}{\sin [\bar{\alpha}(x) \pm n\pi + \bar{c}_2]} \right| = f(x),$$

t being arbitrary from $[t_0, b)$ (for those t for which $\bar{y}(t) = 0$, the last relations follow from the continuity of f).

Putting

$$m = \inf \{f(x);\ x \in [a, \varphi(t_0)]\}$$

and

$$M = \sup \{f(x);\ x \in [a, \varphi(t_0)]\}$$

(the existence follows from the continuity), we get $0 < m \leqq f(t) \leqq M$ for every $t \in [a, b)$, or

$$m\,|\bar{y}(t)| \leqq |y(t)| \leqq M\,|\bar{y}(t)|,$$

since on $\{t \in [a, b);\ y(t)=0\} \equiv \{t \in [a, b);\ \bar{y}(t)=0\}$ the last relation is satisfied trivially. That completes the proof of the lemma.

Note 5. If (7) is satisfied, then evidently y and $\bar{y}$ must have the same set of zeros.

Now, we prove Theorem 2. Let α denote a phase of (q). Since $\lim_{n\to\infty} \varphi_n(a) = b$, then an argument similar to that in (8) now gives

$$[x_0, b) = \bigcup_{n=0}^{\infty} [\varphi_{nk}(x_0),\ \varphi_{(n+1)k}(x_0)), \tag{9}$$

and, if $t \in [x_0, b)$, then there exists an integer $n \geqq 0$ such that $t = \varphi_{nk}(x)$ for $x \in [x_0, \varphi_k(x_0))$, and also this $t \in [\varphi_{nk}(x_0), \varphi_{(n+1)k}(x_0))$. According to (4), every solution of (q) can be written as

$$y(t) = c_1 \frac{\sin [\alpha(t) + c_2]}{\sqrt{|\alpha'(t)|}}, \quad \text{or for} \quad t \in [x_0, b),$$

$$y(t) = y(\varphi_{nk}(x)) = c_1 \frac{\sin [\alpha(\varphi_{nk}(x)) + c_2]}{\sqrt{|\alpha'(\varphi_{nk}(x))|}} =$$

$$= (-1)^{nk} c_1 \frac{\sin [\alpha(x) + c_2]}{\sqrt{|\alpha'(x)|}} \cdot [\varphi'_{nk}(x)]^{1/2},$$

or

$$y(t) = (-1)^{nk} y(x) \cdot [d\varphi_{nk}(x)/dx]^{1/2}, \quad x \in [x_0, \varphi_k(x_0)), \tag{10}$$

$n \to \infty$ as $t \to b$, since from (5) we have

$$\alpha(\varphi_{nk}(x)) = \alpha(x) + nk\pi \operatorname{sign} \alpha', \quad \text{and} \quad \alpha'(\varphi_{nk}(x)) = \alpha'(x)/(\varphi_{nk}(x))'. \tag{11}$$

a) For $k=1$, the proof was given in [12]. Since y is continuous, denote $K = \max \{|y(t)|;\ t \in [a, \varphi_k(x_0)]\}$. If $(\varphi_k(t))' \leqq 1$ for a positive integer k, then from (10) we have for $t \in [a, b)$

$$|y(t)| \leqq \max \{\max \{|y(t)|;\ t \in [a, x_0]\},\ \sup \{|y(t)|;\ t \in [x_0, b)\}\} \leqq$$

$$\leqq \max \{K, \sup \{|y(x)|. \prod_{i=0}^{n-1} [\varphi'_k(\varphi_{ik}(x))]^{1/2};\ x \in [x_0, \varphi_k(x_0)]\}\} \leqq$$

$$\leqq \max \{K, \sup \{|y(x)|;\ x \in [x_0, \varphi_k(x_0)]\}\} = K.$$

b) For $b=\infty$ the proof was given in [11].

b1. If $(\varphi_n(x))' \leqq K$ for every integer $n>0$ and every $x \in [a, \varphi(a)]$, then (10) for $k=1$, $x_0=a$ gives

$$y(t) = (-1)^n y(x) \cdot (\varphi_n'(x))^{1/2}, \quad x \in [a, \varphi(a)),$$

or

$$|y(t)| \leqq \sqrt{K} \cdot \max \{|y(x)|; \ x \in [a, \varphi(a)]\} = \text{const.}$$

for every $t \in [a, b)$.

b2. If $(\varphi_n(x))'$ is not bounded for every integer $n>0$ on $[a, \varphi(a)]$, then it is possible to find a sequence $\{x_{n_i}\}_{n=1}^{\infty}$ from $[a, \varphi(a)]$ such that $\varphi_{n_i}'(x_{n_i}) > i$ for $i=1, 2, \ldots$. If $x_0 \in [a, \varphi(a)]$ is an accumulation point of $\{x_{n_i}\}_{i=1}^{\infty}$, and a subsequence $\{x_{n_{i_j}}\}_{j=1}^{\infty} \to x_0$, we have $\varphi_{n_{i_j}}'(x_{n_{i_j}}) > i_j$ for $j=1, 2, \ldots$. Let y be a solution of (q), $y(x_0) \neq 0$. Denote $t_{n_{i_j}} = \varphi_{n_{i_j}}(x_{n_{i_j}})$. From (10) with $k=1$, $x_0=a$, we get

$$\lim_{j\to\infty} |y(t_{n_{i_j}})| = \lim_{j\to\infty} |y(\varphi_{n_{i_j}}(x_{n_{i_j}}))| = \lim_{j\to\infty} |y(x_{n_{i_j}})| \cdot [\varphi_{n_{i_j}}'(x_{n_{i_j}})]^{1/2} \geqq$$

$$\geqq |y(x_0)| \lim_{j\to\infty} \sqrt{i_j} = \infty.$$

c) The proof is based on the relations (9) and (10), and can be found in [1].

d) For $k=1$, the proof is in [12]. Let k be a positive integer. If $(\varphi_k(t))' \leqq c = \text{const} > 1$, then for an arbitrary non-trivial solution y of (q) the relation (10) gives

$$|y(t) = |y(x)| \cdot \left[\prod_{i=0}^{n-1} \varphi_k'(\varphi_{ik}(x))\right]^{1/2} \geqq |y(x)| \cdot c^{n/2}.$$

Since y is non-trivial, there exists an $x_1 \in [x_0, \varphi(x_0))$ such that $y(x_1) \neq 0$. If $t_j = \varphi_{jk}(x_1)$, we have $\lim_{j\to\infty} |y(t_j)| \geqq \lim_{j\to\infty} |y(x_1)| \cdot c^{j/2} = \infty$, since $c>1$.

e) For $k=1$, see [12]. If y is a solution of (q) and $(\varphi_k(t))' \leqq c < 1$, then (10) yields

$$\lim_{t\to b} |y(t)| = \lim_{n\to\infty} |y(x)| \cdot \left[\prod_{i=0}^{n-1} \varphi_k'(\varphi_{ik}(x))\right]^{1/2} \leqq \lim_{n\to\infty} |y(x)| \cdot c^{n/2} = 0,$$

since $y(x)$ for $x \in (x_0, \varphi(x_0)) \subset [x_0, \varphi(x_0)]$ is continuous and hence bounded, and $c = \text{const} < 1$.

f) and g) For proofs, see [1].

h) The case was mentioned in footnotes in [13], however only sufficiency for $p>0$ and necessity for $p=2$ were proved there.

Let $p>0$ be given. Again we have $[a, b) = \bigcup_{n=0}^{\infty} (\varphi_n(a), \varphi_{n+1}(a))$, $t \in [a, b)$, or $t \in [\varphi_n(a), \varphi_{n+1}(a))$ written in the form $t=\varphi_n(x)$, where $x \in [a, \varphi(a))$ and $n \to \infty$ as $t \to b$.

If α is a phase of (q), then $y(t; c_1, c_2)$ in (4) is the general solution of (q), and

$$u(t) = \frac{\sin \alpha(t)}{\sqrt{|\alpha'(t)|}} \quad \text{and} \quad v(t) = \frac{\cos \alpha(t)}{\sqrt{|\alpha'(t)|}}$$

are two linearly independent solutions of (y). If $\int_a^b |\alpha'(t)|^{-p/2}dt < \infty$, then $\int_a^b |y(t; c_1, c_2)|^p dt \leqq |c_1|^p \int_a^b |\alpha'(t)|^{-p/2}dt < \infty$ and $C1$ holds for (q). Conversely, if $C1$ holds for (q), then also $u \in L^p[a, b)$ and $v \in L^p[a, b)$, or

$$\int_a^b \frac{|\sin \alpha(t)|^p + |\cos \alpha(t)|^p}{|\alpha'(t)|^{p/2}} dt < \infty. \tag{12}$$

Since $f_p : x \mapsto |\sin x|^p + |\cos x|^p$ is continuous, positive and periodic with period π, we have

$$0 < r = \min \{f_p(x);\ x \in [0, \pi]\} = \min \{f_p(x);\ x \in (-\infty, \infty)\} \leqq f_p(x)$$

for every $x \in (-\infty, \infty)$.

From (12)

$$r \int_a^b |\alpha'(t)|^{-p/2}dt \leqq \int_a^b \frac{f_p(\alpha(t))}{|\alpha'(t)|^{p/2}} dt < \infty,$$

$$\text{or} \quad \int_a^b |\alpha'(t)|^{-p/2}dt < \infty.$$

Hence $C1$ holds for (q) iff $\int_a^b |\alpha'(t)|^{-p/2}dt < \infty$.

Further the proof proceeds as that of Lemma 1 in [13]:

$$\int_a^b |\alpha'(t)|^{-p/2}dt = \sum_{n=0}^{\infty} \int_{\varphi_n(a)}^{\varphi_{n+1}(a)} |\alpha'(t)|^{-p/2}dt = \sum_{n=0}^{\infty} \int_a^{\varphi(a)} \frac{d\varphi_n(x)}{|\alpha'(\varphi_n(x))|^{p/2}} =$$

$$= \sum_{n=0}^{\infty} \int_a^{\varphi(a)} \frac{[(\varphi_n(x))']^{1+p/2}}{|\alpha'(x)|^{p/2}} dx.$$

Since $\alpha'(x) \neq 0$ and $\alpha' \in C^2[a, b)$, there exist s and S such that

$$0 < s = \min \{|\alpha'(x)|^{-p/2};\ x \in [a, \varphi(a)]\}$$

and

$$S = \max \{|\alpha'(x)|^{-p/2};\ x \in [a, \varphi(a)]\}.$$

Hence

$$s \sum_{n=0}^{\infty} \gamma_n^{(p)} \leqq \int_a^b |\alpha'(t)|^{-p/2}dt \leqq S \sum_{n=0}^{\infty} \gamma_n^{(p)},$$

where $\gamma_n^{(p)} \overset{\text{def}}{=} \int_a^{\varphi(a)} [(\varphi_n(x))']^{1+p/2}dx$, that completes the proof of case h.

i1. If every solution of (q) is periodic (with period $d>0$, then also $b=\infty$), then necessarily every such a solution vanishing at some t_0 must vanish also at t_0+d, and the number of all zeros of the solution on $t_0 \leqq t < t_0+d$ (say $k>0$) does not depend on t_0 and on the chosen solution, otherwise we get a contradiction with the Separation Theorem.

Hence $\varphi_k(t_0) = t_0+d$ for every $t_0 \in [a, \infty)$, or

$$d\varphi_k(t)/dt \equiv 1 \quad \text{on} \quad [a, \infty). \tag{12}$$

i2. If (12) holds, then $\varphi_k(t) = t+d$ ($d>0$ from (1)). Since $\varphi_{nk}(t) = t+nd$ and $\lim_{n\to\infty} \varphi_{nk}(t)=b$, we have $b=\infty$. According to (4) and (11), if α is a phase of (q), any solution $y(t; c_1, c_2)$ of (q) satisfies

$$y(t+d;\, c_1, c_2) = c_1 \frac{\sin\left[\alpha(\varphi_k(t))+c_2\right]}{\sqrt{|\alpha'(\varphi_k(t))|}} = (-1)^k c_1 \frac{\sin\left[\alpha(t)+c_2\right]}{\sqrt{|\alpha'(t)/(\varphi_k(t))'}} =$$

$$= (-1)^k y(t;\, c_1, c_2), \text{ or}$$

every solution of (q) is periodic with period d (for k even), or $2d$ (for k odd).

6. The results can be directly applied to global properties of plane curves using methods described in detail in [14]. As an example let us introduce the following consequence of Theorem 2, case a).

Corollary. Let $u_1 \in C^2[a, b)$, $u_2 \in C^2[a, b)$, $b \leqq \infty$, be coordinate functions of a plane curve $\mathbf{u} \equiv (u_1, u_2)$. Let the radius vector be not parallel to the tangent vector at any point of the curve. Let the parametrization be normalized such that $u_1'(t)u_2(t) - u_1(t)u_2'(t) = \text{const.}$ (this is always possible). Let $\mathbf{u}(t_0)$ be an arbitrary point of the curve, and $\mathbf{u}(t_k)$ be the k-th intersection of the curve in the direction of increasing parameter with the straight line passing the origin and $\mathbf{u}(t_0)$ ($\mathbf{u}(s)$ denotes the point of the parameter s). Denote $\Delta(t_0) \overset{\text{def}}{=} t_k - t_0$.

Now, *if there exists a straight line going through the origin and having infinitely many intersections with* $\mathbf{u}$ *for* $[a, b)$, *and if* $\Delta(t)$ *is a non-increasing function of* t *for sufficiently large* t (*i.e. on some* $[x_0, b) \subset [a, b)$), *then the curve* $\mathbf{u}$ *is bounded for* $t \in [a, b)$.

7. Problems. With respect to the form of the theorems 1 and 2, and for their possible applications, the following problems arise.

Problem 1. What are the sufficient and/or necessary conditions on functions q and $\bar{q}$ in order that (q) and ($\bar{\text{q}}$) have the same dispersion?

Problem 2. What are the sufficient and/or necessary conditions on q in order that the dispersion of (q) is of a certain type listed in Theorem 2?

A necessary condition on (q) and ($\bar{\text{q}}$) having the same dispersion φ is that the difference $q(t)-\bar{q}(t)$ vanishes at least four times on every interval $t_0 \leqq t < \varphi(t_0)$, and it was given in [4] (see also [5]). It was shown in [14, p. 589] that the estimation cannot be improved.

In [2] it was established that every function φ satisfying (1) can be considered as the dispersion of a differential equation (q). It was proved in [3] and [8] that the cardinality of all such differential equations (q) is the same as the cardinality of all real numbers.

Several papers [8], [9], [14] were devoted to construction of all differential equations having the dispersion $\varphi(t) = t+d$. The necessary condition $\int_0^d q(t)dt \geqq -\pi^2/d$, the equality just for $q(t) = -\pi^2/d^2$, is derived for those equations.

The explicit form of all differential equations having only periodic solutions (or $(\varphi_k(t))' \equiv 1$) was given in [10] or [14]. Those results are more convenient when dealing with the whole set of those differential equations or if we want to construct such a differential equation, rather then for examining a given particular equation.

In several papers [6], [7], [15], [16] there are derived monotonicity properties of dispersions of (q) when some types of monotonicity of q is required. An extension of these results, perhaps with some kind of integral conditions, would be very useful.

REFERENCES

[1] BARTUŠEK, M.: Connection between asymptotic properties and zeros of solutions of $y''=q(t)y$, *Archivum Mathematicum* **3** (1972), 113—124.

[2] BARVÍNEK, E.: On distribution of zeros of solutions of a differential equation $y''=Q(t)y$ and their derivatives, *Acta Fac. Nat. Univ. Comenian* V, 8—10 Math. (1961), 465—474.

[3] BORŮVKA, O.: Sur l'ensemble des équations différentielles linéaires ordinaires du deuxième ordre qui ont la même dispersion fondamentale, *Bull. Inst. Polyt. Iasi* (*9*) **13** (1963), 11—20.

[4] BORŮVKA, O.: Sur quelques applications des dispersions centrales dans la thèorie des équations différentielles linéaires du deuxième ordre, *Arch. Math.* (Brno) **1** (1965), 1—20.

[5] BORŮVKA, O.: *Linear Differential Transformations of the Second Order*, The Eng. Univ. Press, London 1971.

[6] HARTMAN, P.: On differential equations and the functions $J_\mu^2+Y_\mu^2$, *Amer. J. Math.* **83** (1961), 154—187.

[7] LORCH, L. and SZEGŐ, P.: Higher monotonicity properties of certain Sturm-Liouville functions I: *Acta Math.* **109** (1963), 55—73. *II*: *Bull. Acad. Pol., Ser. Math. Astr. Phys.* **11** (1963), 455—457, together with MULDOON; *Can. Math. J.* **22** (1970), 1238—1265.

[8] NEUMAN, F.: Sur les équations différentielles linéaires oscillatoires du deuxième ordre avec la dispersion fondamentale $\varphi(t) = t+\pi$, *Bull. Inst. Pol. Jassy* (Romania) **10** (*14*), (1964), 37—42.

[9] NEUMAN, F.: Extremal property of the equation $y'' = -k^2y$, *Archivum Mathematicum* **3** (1967), 161—164.

[10] NEUMAN, F.: An explicit form of the differential equation $y''=q(t)y$ with periodic solutions, *Ann. Mat. Pura Appl.* **85** (1970), 295—300.

[11] NEUMAN, F.: Relation between the distribution of the zeros of the solution of a 2nd order linear differential equation and the boundedness of these solutions, *Acta Math. Hung.* **19** (1968), 1—6.

[12] NEUMAN, F.: A role of Abel's equation in the stability theory of differential equations, *Aequat. Math.* **6** (1971), 66—70.

[13] NEUMAN, F.: L^2-solutions of $y''=q(t)y$ and a functional equation, *Aequat. Math.* **6** (1971), 162—169.

[14] NEUMAN, F. Linear differential equations of the second order and their applications, *Rendiconti di Mat.* **4** (1971), 559—617.

[15] VOSMANSKÝ, J.: The monotonicity of extremants of integrals of the differential equation $y''+$ $+q(t)y = 0$, *Arch, Math.* **2** (1966), 105—111.

[16] VOSMANSKÝ, J.: Monotonic properties of zeros and extremants of the differential equation $y''+q(t)y = 0$, *Arch. Math.* (Brno) **6** (1970), 37—73.

Brno, Czechoslovakia

(Received July 30, 1972.)

MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

INDEX

Huq, S. A.: Some remarks on multicategories 1
Варшамов, Р.Р.: Об одном линейном операторе в поле Галуа и его приложении 5
Mogyoródi, J.: On the rarefaction of renewal processes, III. 21
Mogyoródi, J.: On the rarefaction of renewal processes, IV. 29
Scozzafava, R.: The radical of a locally compact semialgebra 39
Chartrand, G. and *Wall, C. E.:* On the hamiltonian index of a graph 43
Rejtő, L.: A more general Farkas-type theorem 49
Jucovič, E.: On the face vector of a 4-valent 3-polytope 53
Крамли, А.: Геодезические потоки на компактнах римановых поверхностях без фокалных точек 59
Steck, G. P. and *Simmons, G. I.:* On the distributions of R_{mn}^{+} and (D_{mn}^{+}, R_{mn}^{+}) 79
Chakerian, G. D.: On a certain affine invariant functional for convex bodies 91
Major, P.: A law of the iterated logarithm for the Robbins—Monro method 95
Чан Суань Хиен: О некоторых задачах относящихся к одному семейству уравнений в частных производных составного типа 103
Неваи, Г. П.: Замечания о тригонометрическом интерполировании и суммах Фурье 113
Székely, G. J.: On the coefficients of polynomials with negative zeros 123
Győri, I.: Asymptotic behaviour of solutions of unstable-type first order differential equations with delay 125
Makai, E.: Variational problems related to those which lead to the Schrödinger equation of helium-like atoms 133
Babai, L. and *Lovász, L.:* Permutation groups and almost regular graphs 141
Lovász, L. and *Major, P.:* A note to a paper of Dudley 151
Stoyan, D.: Bounds for the extrema of the expected value of a convex function of independent random walk 153
Steinfeld, O.: Notes on regular duo elements, rings and semigroups 161
Lovász, L.: Independent sets in critical chromatic graphs 165
Elbert, A.: On the consecutive zeros of the solutions of linear second order differential equations 169
Neuman, F.: Distribution of zeros of solutions of $y''=q(t)y$ in relation to their behaviour in large 177

MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

Printed in Hungary

A kiadásért felel az Akadémia Kiadó igazgatója – Műszaki szerkesztő: Botyánszky Pál
A kézirat a nyomdába érkezett: 1973. III. 10. – Terjedelem: 16,75 (A/5) ív, 15 ábra

---

73-1176 Szegedi Nyomda

Die *Studia Scientiarum Mathematicarum Hungarica* ist eine Halbjahrsschrift der Ungarischen Akademia der Wissenschaften. Sie veröffentlicht Originalbeiträge aus dem Bereiche der Mathematik in deutscher, englischer, französischer oder russischer Sprache. Es erscheint jährlich ein Band.

Adresse der Redaktion: 1061 Budapest V., Reáltanoda u. 13—15, Ungarn.
Technischer Redaktor: Gy. Petruska

Abonnementspreis pro Band (pro Jahr): $ 16.00. Bestellbar bei Buch- und Zeitungs-Aussenhandelsunternehmen *Kultúra* (Budapest 62, P. O. B. 149), oder bei den Vertretungen im Ausland.
Austauschabmachungen können mit der Bibliothek des Matematischen Instituts (1061 Budapest V., Reáltanoda u. 13—15) getroffen werden.
Die zur Veröffentlichung bestimmten Manuskripte sind in zwei Examplaren an die Redaktion zu schicken.

Studia Scientiarum Mathematicarum Hungarica est une revue biannuelle de l'Académie Hongroise des Sciences publiant des essais originaux, en francais, anglais, allemand ou russe, du domaine des mathématiques.

Rédaction: 1061 Budapest V., Reáltanoda u. 13—15, Hongrie.
Rédacteur téchnique: Gy. Petruska

Le prix de l'abonnement: $ 16.00 par an (volume). On s'abonne chez *Kultúra*, Société pour le Commerce de Livres et Journaux (Budapest 62, P. O. B. 149) ou chez ses représentants à l'étranger.
Pour établir des relations d'échange on est prié de s'adresser à la Bibliothèque de l'Institut de Matématique (1061 Budapest V., Reáltanoda u. 13—15).
On est prié 'envoyer les articles destinés à la publication en deux exemplaires à l'adresse de la rédaction.

Studia Scientiarum Mathematicarum Hungarica — выходит два раза в год в Издании Академии Наук Венгрии. Журнал публикует оригинальные исследования в области математики на немецком, английском, французском и русском языках. Отдельвные вышуски составляют ежегодно одни том.

Адрес редакции: 1061 Budapest V., Reáltanoda u. 13—15, Венгрия.
Технический редактор: Gy. Petruska.

Подписная цена на год (за однии том): $ 16.00. Подписка на журнал принимается Внешнеторговым предприятием „Култура" (Budapest 62, P. O. B. 149) или его представительствами за границей.
По поводу отношения обмена просим обращаться к Библиотеке Института Математики (1061 Budapest V. Reáltanoda u. 13—15).
Работы, предназначенные для опубликования в журнале следует направлять по адресу редакции в двух экземплярах.

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.

CHECHOSLOVAKIA

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

31. I. 1974. **Index: 26 748**

Studia Scientiarum Mathematicarum Hungarica

AUXILIO
CONSILII INSTITUTI MATHEMATICI
ACADEMIAE SCIENTIARUM HUNGARICAE

REDIGIT
L. FEJES TÓTH

ADIUVANTIBUS
M. ARATÓ, I. CSISZÁR, T. FREY, G. FREUD,
B. GYIRES, A. HEPPES, L. KALMÁR, E. MAKAI,
A. PRÉKOPA, P. RÉVÉSZ, K. TANDORI

TOMUS VIII.
FASC. 3–4.
1973

AKADÉMIAI KIADÓ, BUDAPEST

Studia Scientiarum Mathematicarum Hungarica

A Magyar Tudományos Akadémia matematikai folyóirata

Szerkesztőség: 1053 Budapest V., Reáltanoda u. 13—15.

Technikai szerkesztő: Petruska Gy.

Kiadja az Akadémiai Kiadó, 1054 Budapest V., Alkotmány u. 21.

A *Studia Scientiarum Mathematicarum Hungarica* angol, német, francia vagy orosz nyelven közöl eredeti értekezéseket a matematika tárgyköréből. Félévenként jelenik meg, évi egy kötetben.

Előfizetési ára belföldre 120,—Ft, külföldre 165,— Ft. Megrendelhető a belföld számára az Akadémiai Kiadónál, a külföld számára pedig a Kultúra Könyv és Hirlap Külkereskedelmi Vállalatnál (1011 Budapest II., Fő u. 32.).

Cserekapcsolatok felvétele ügyében kérjük a MTA Matematikai Kutató Intézete Könyvtárához (1053 Budapest V., Reáltanoda u. 13—15.) fordulni.

A közlésre szánt dolgozatokat kérjük két példányban a szerkesztőség címére küldeni.

Studia Scientiarum Mathematicarum Hungarica is a journal of the Hungarian Academy of Sciences publishing original papers on mathematics, in English, German, French or Russian.

It is published semiannually, making up one volume per year.

Editorial Office: 1053 Budapest V., Reáltanoda u. 13—15, Hungary.

Technical Editor: Gy. Petruska

Subscription rate: $ 16.00 per volume. Orders may be placed with *Kultúra* Trading Co. for Books an Newspapers, Budapest 62, P. O. B. 149 or with its representatives abroad.

For establishing exchange relations please write to the Library of the Mathematical Institute (1053 Budapest V., Reáltanoda u. 13—15.)

Papers intended for publication should be sent to the Editor in 2 copies.

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 187—192.

ON REPRESENTATION OF MODULE-HOMOMORPHISMS (MULTIPLIERS)

by

L. MÁTÉ

§ 1.

Let A be a Banach algebra and W be a (left) Banach A-module. The *essential part* W_1 of W is the closure of the linear hull of $\{aw\,;\, a\in A,\, w\in W\}$. If $W_1=W$ then W is called *essential*. If W is an essential (left) A-module then the strong dual W^*, with the adjoint action, is a (right) Banach A-module such that $w^*a=0$ *for every* $a\in A$ *implies* $w^*=0$.

In § 2 and § 3 it will be supposed that W is essential.

The operators T from A into W^* satisfying

$$Tca = (Tc)a \quad c,\, a\in A$$

are called (A, W^*)-multipliers. These operators, with the operator norm, form a Banach space denoted by $\underset{A}{\mathrm{Hom}}\,(A, W^*)$.

For the characterization of $\underset{A}{\mathrm{Hom}}\,(A, W^*)$ the most general result known to the author is the following ([11] Theorem 8.9).

"Let A be a Banach algebra with bounded approximate identity, and let W be an essential left A-module (or the module operation in W^* be weak*-continuous), then there is a natural isometric right A-module isomorphism

$$\underset{A}{\mathrm{Hom}}\,(A, W_1^*) \cong W^*$$

where W_1^* is the essential part of W^*."

Few attention has been turned to cases where A contains only an *unbounded* approximate identity. Let us consider some examples.

1. $A=L^p$, where L^p, $1<p<\infty$, is the usual Lebesgue spaces of functions on a *compact* group. Here for any approximate identity $\{e_\alpha\}$ of L^1 consisting of continuous functions we have

$$e_\alpha * f \to f \quad \text{for every} \quad f\in L^p \quad \text{and} \quad \|e_\alpha * f\|_p \leqq \|e_\alpha\|_1 \|f\|_p$$

where $*$ is the convolution and $\|\cdot\|_p$ is the norm in L^p.

2. Let G be a locally compact Abelian group, $\hat{G}$ the dual group and let the Fourier transform be denoted by $\hat{}$. Then L_p^1, described in [4] sec. 6.1, as the set

$$\{f: f\in L^1(G),\, \hat{f}\in L^p(\hat{G}) \text{ with norm } \|f\| = \|f\|_1 + \|\hat{f}\|_p\}$$

MAGYAR TUDOMÁNYOS AKADÉMIA KÖNYVTÁRA

is a Banach algebra with the convolution. There is no bounded approximate identity L_p^1, however if $\{e_\alpha\}$ is a bounded approximate identity in $L^1(G)$ such that $\hat{e}_\alpha$ has compact support then we have

$$e_\alpha * f \to f \quad \text{for every} \quad f \in L_p^1 \quad \text{and} \quad \|e_\alpha * f\| = \|e_\alpha\|_1 \|f\|.$$

3. If A is the Banach algebra of Hilbert—Schmidt operators of a Hilbert space, then there is an approximate identity $\{e_\alpha\}$ in A consisting of finite rank operators, unbounded in the original norm, however

$$\{\sup_\alpha \|e_\alpha a\|;\ \|a\| = 1, \quad a \in A\}$$

is bounded.

A common feature of the above examples is that the approximate identity is not bounded in A, however, the set

$$(*) \qquad \{\sup \|e_\alpha a\|;\ \|a\| = 1, \quad a \in A\}$$

is bounded. I.e. considering the regular representation of A, the approximate identity $\{e_\alpha\}$ is bounded in the operator norm.

In the following, an approximate identity in A, satisfying $(*)$, will be called *quasi-bounded.*

The purpose of this Note is to extend RIEFFEL's result ([11], Theorem 8.9) to the case when there exists only a quasi-bounded approximate identity in A.

§ 2.

Let $A \hat{\otimes} W$ be the projective (greatest) tensor product ([2] p. 92—95) of A and W. Then there is a continuous homomorphism Φ from $A \hat{\otimes} W$ into W defined by $\Phi(a \hat{\otimes} w) = aw$, $a \in A$, $w \in W$.

DEFINITION. $A \circ W = \Phi(A \hat{\otimes} W)$ with the quotient norm (mod. ker Φ).

For each $w^* \in W^*$ we define $F_{w^*} \in (A \hat{\otimes} W)^*$ by

$$(1) \qquad \langle F_{w^*}, a \otimes w \rangle = \langle w^*, \Phi(a \otimes w) \rangle \qquad a \in A, \quad w \in W;$$

thus there is a natural monomorphism from W^* into $(A \hat{\otimes} W)^*$. Identifying W^* with its copy $\{F_{w^*};\ w^* \in W^*\}$ we have

$$W^* \subseteqq (A \circ W)^* \subseteqq (A \hat{\otimes} W)^*.$$

Moreover, if $h \in A \hat{\otimes} W$ then $\langle F_{w^*}, h \rangle = 0$ for every $w^* \in W^*$ iff $\Phi(h) = 0$ and $(A \circ W)^*$ consists of exactly those elements F of $(A \hat{\otimes} W)^*$ for which $\Phi(h) = 0$ implies $\langle F, h \rangle = 0$. So, we conclude

PROPOSITION 1. *$(A \circ W)^*$ is the weak*-closure of W^* in $(A \hat{\otimes} W)^*$.*

Now let $B(A, W^*)$ be the usual Banach space of all bounded linear operators from A into W^*; then there is a natural isometric isomorphism

$$(2) \qquad (A \otimes W)^* = B(A, W^*) \quad \text{defined by} \quad F \cong T \quad \text{if} \quad \langle F, a \hat{\otimes} w \rangle = \langle Ta, w \rangle.$$

Particularly if

$$T_{w^*}: T_{w^*}a = w^*a \tag{3}$$

then $T_{w^*} \in \underset{A}{\mathrm{Hom}}\,(A, W^*)$ for every $w^* \in W^*$ and (2) yields an isometry from $\{F_{w^*};\ w^* \in W^*\}$ onto $\{T_{w^*};\ w^* \in W^*\}$. Thus from Proposition 1 we obtain

PROPOSITION 2. *If $F \in (A \circ W)^*$, and T is the operator corresponding to F by* (2), *then $T \in \underset{A}{\mathrm{Hom}}\,(A, W^*)$. If $(A \hat{\otimes} W)^*$ is supplied with the weak* topology then the topology of the homeomorphic image of $(A \hat{\otimes} W)^*$ on $B(A, W^*)$ is called ultraweak topology.*

§ 3.

$\underset{A}{\mathrm{Hom}}\,(A, W^*)$ can be characterized by $(A \circ W)^*$ as follows:

THEOREM 1. *If there is a quasi-bounded approximate identity in A then*

$$\underset{A}{\mathrm{Hom}}\,(A, W^*) = (A \circ W)^*$$

by (2).

PROOF I. In view of Proposition 2, we have only to show that if $T \in \underset{A}{\mathrm{Hom}}\,(A, W^*)$, and $T \cong F$ by (1), then $F \in (A \circ W)^*$. We can follow the line of the proof of [11] Theorem 8.9.

If $\{e_\alpha\}$ is a quasi-bounded approximate identity and T_{e_α} is defined by $T_{e_\alpha}a = e_\alpha a$ $a \in A$, then for every $T \in \underset{A}{\mathrm{Hom}}\,(A, W^*)$ the net TT_{e_α} is bounded. Consequently, if $F_\alpha \cong TT_{e_\alpha}$ by (2), then there is a weak* convergent subnet, still denoted by F_α, which converges to a certain $F \in (A \circ W)^*$ by Proposition 1. If T_F is the operator corresponding to F, then it is enough to show that $T_F = T$.

Indeed, $TT_{e_\alpha} \to T_F$ in the ultraweak topology and hence also in the weak operator topology. On the other hand,

$$Ta = \lim_\alpha T(e_\alpha a) = \lim_\alpha (Te_\alpha)a$$

i.e. $TT_{e_\alpha} \to T$ in the strong operator topology and hence also in the weak operator topology which implies that $T_F = T$.

PROOF II. Let A^* be the Banach space dual of A and I the linear functional on $A \hat{\otimes} A^*$ defined by

$$\langle I, a \hat{\otimes} a^* \rangle = \langle a^*, a \rangle \qquad a \in A, \quad a^* \in A^*;$$

then it follows from the Banach—Steinhaus theorem that if $F_\alpha \in (A \hat{\otimes} A^*)^*$ is defined by

$$\langle F_\alpha, a \hat{\otimes} a^* \rangle = \langle a^*, e_\alpha a \rangle \qquad a \in A, \quad a^* \in A^*,$$

where e_α is a quasi-bounded approximate identity in A, then $F_\alpha \to I$ in the weak* topology.

It is clear that if E is the identity operator in $B(A, A^{**})$ then

$$I \cong E \tag{4}$$

by (2). Considering A^* as a two-sided A-module with the adjoint action and applied the considerations, especially Proposition 1, in § 2 to $W=A^*$ we obtain

$$I \in (A \circ A^*)^*. \tag{5}$$

The second proof is based on (4) and (5).

Let T be a linear operator from A into W^*, T^* the dual operator and let $E \otimes T^*$ be defined by

$$(E \otimes T^*)(a \hat{\otimes} w) = a \hat{\otimes} (T^* w) \qquad a \in A, \quad w \in W.$$

Then $(E \otimes T^*)^*$ is a linear operator from $(A \hat{\otimes} A^*)^*$ into $(A \hat{\otimes} W)^*$ and

$$\langle (E \otimes T^*)^* I, a \hat{\otimes} w \rangle = \langle I, a \hat{\otimes} T^* w \rangle = \langle a, T^* w \rangle = \langle Ta, w \rangle = \langle F_T, a \hat{\otimes} w \rangle$$

$$a \in A, \quad w \in W$$

where F_T is the functional corresponding to T by (2). Hence

$$(E \otimes T^*)^* I = F_T. \tag{6}$$

If we suppose that $T \in \underset{A}{\mathrm{Hom}}\,(A, W^*)$ then from

$$\langle c, a(T^* w) \rangle = \langle ca, T^* w \rangle = \langle Tca, w \rangle = \langle (Tc)a, w \rangle = \langle Tc, aw \rangle$$

it follows that $a \circ w = 0$ implies $a \circ (T^* w) = 0$ and hence $E \otimes T^*$ is a linear operator from $A \circ W$ into $A \circ A^*$.

Consequently, if $T \in \underset{A}{\mathrm{Hom}}\,(A, W^*)$ and the condition of the theorem is satisfied then it follows from (6) that $F_T \in (A \circ W)^*$ Q.E.D.

Compared Theorem 1 and [11] Theorem 8.9, it is natural to ask:

Identifying W^* with its image $\{F_{w^*};\ w^* \in W^*\}$ by (1), what is the condition of

$$\underset{A}{\mathrm{Hom}}\,(A, W^*) = W^*$$

in the case when only a quasi-bounded approximate identity exists in A?

The answer is included in

THEOREM 2. *The following assertions are equivalent:*

I. $\underset{A}{\mathrm{Hom}}\,(A, W^*) \cong W^*$;

II. $\underset{A}{\mathrm{Hom}}\,(A, W_1^*) \cong W^*$;

III. $\|T_{w^*}\| = \|w^*\|$ *for every* $w^* \in W^*$ *and* T_{w^*} *defined by* (3).

PROOF. II⇒III is clear.

I⇒II follows from that $T_{w^*} \in \underset{A}{\mathrm{Hom}}\,(A, W^*)$ for every $w^* \in W^*$.

III⇒I: It follows from Theorem 1 that III. is the same as $(A \circ W)^* \cong W^*$.

If L is the injection of $A \circ W$ into W then L is a continuous 1—1 operator with dense range hence the dual operator L^* is a continuous 1—1 operator from W^* into $(A \circ W)^*$ with dense range in the weak* topology of $(A \circ W)^*$.

So, if we show that $L^* W^*$ is also closed in the weak* topology then the theorem will be proved.

From the KREIN—SMULIAN theorem ([2] p. 151—152) it follows that it is sufficient to show that $L^* W^* \cap S^*$ is weak-closed where S^* is the unit sphere in $(A \circ W)^*$.

Let $L^* w_\alpha^* \in S^*$ and let the net $\{\langle a \circ w, L^* w_\alpha^* \rangle\}$ be convergent for every $a \circ w \in A \circ W$. Then from

$$\langle a \circ w, L^* w_\alpha^* \rangle = \langle L(a \circ w_\alpha), w_\alpha^* \rangle = \langle aw, w_\alpha^* \rangle$$

it follows that $\{w_\alpha^*\}$ is pointwise convergent on a dense subset of W. Moreover $\{w_\alpha^*\}$ is bounded. Consequently, according to the Banach—Steinhaus Theorem, there exists a $w^* \in W^*$ such that weak*-lim $w_\alpha^* = w^*$ and hence

$$\lim_\alpha \langle a \circ w, L^* w_\alpha^* \rangle = \lim \langle L(a \circ w), w_\alpha^* \rangle = \langle L(a \circ w), w^* \rangle = \langle a \circ w, L^* w^* \rangle$$

for every $a \circ w \in A \circ W$ Q.E.D.

REMARK. If we suppose *merely the equivalence* of the norms $\|T_{w^*}\|$ and $\|w^*\|$ instead of III, then $\cong$ must be replaced by homeomorphism in I and II, and the same proof applies. Moreover it follows from the Banach homeomorphism theorem that in this case III. can be replaced by

III′. $\{T_{w^*}; w^* \in W^*\}$ is closed in $B(A, W^*)$ resp. $\underset{A}{\mathrm{Hom}}\,(A, W^*)$.

§ 4.

Let W be *any* Banach A-modul. On the basis of [11] Theorem 4.4, it may be thought, that $A \circ W$ is the essential part of W. In the following it will be shown that in the case where A contains a quasi-bounded approximate identity in A, W_1 and $A \circ W$ can be very different.

Let L^p and $L^{p'}$ $1 < p \leqq p' < \infty$ be two different Lebesgue spaces of functions on a compact group and $A = L^p$, $W = L^{p'}$. Then for any approximate identity e_α of L^1 consisting of continuous functions we have $e_\alpha * f \in L^{p'}$ and $e_\alpha * f \to f$ for every $f \in L^{p'}$ (see § 1 Example 1) and hence in this case $W_1 = W$.

On the other hand, $(L^p \circ L^{p'})^*$ is defined with the L^p-multipliers by Theorem 1 and for *any* $h \in L^1$, the convolution operator $T_h : T_h g = h^* g$ is an L^p-multiplier. Hence in this case $(A \circ W)^* \neq W^* = W_1^*$ which implies that $A \circ W \neq W_1$.

As a contrast, it can be proved, similarly to [11] Proposition 3.4, the followin

PROPOSITION 3. *If there is a bounded approximate identity in A then*

$$A \circ W \cong W_1$$

where $\cong$ *means that the natural mapping,* $a \circ w \to aw$ *is an isometric isomorphism.*

§ 5.

We conclude with remarks concerning the case of *bounded* approximate identity. They can be directly obtained from [11], but they seem worth mentioning.

We recall that A^* is a (two-sided) Banach A-module and hence so is A^{**}. Moreover, the module operation and the Arens product ([6], [11], [12]) restricted to the pairs $\{(a, a^{**})\colon a \in A,\ a^{**} \in A^{**}\}$ coincide.

So if $W = A^*$ in Theorem 1 and Proposition 3, then we obtain:

If there is a bounded approximate identity in A then

$$\mathrm{Hom}(A, A) \subseteqq (A_1^*)^* \tag{7}$$

and the algebra of A-multipliers is isometrically isomorphic to a certain quotient algebra of A^{**} with Arens product.

CIVIN and YOOD ([13], Lemma 3.8) have proved that there is a bounded weak approximate identity in A iff there is an identity in A^{**} with Arens product. Since the identity operator is an A-multiplier, it is clear that a necessary condition for (7) is the existence of a bounded *weak* approximate identity in A.

For a more detailed discussion of the connections between multipliers and Arens product we refer to [6], [7] and [9].

I am indebted to J. SIMPSON and J. POWELL for calling my attention to an inaccurancy in [6] Thm. 5 which led to the discussions in this §.

REFERENCES

[1] NORTHCOTT, D. G.: *An Introduction to Homological Algebra,* Cambridge Univ. Press 1962
[2] SCHAFFER, H. H.: *Topological Vector Spaces,* Macmillan Co. 1966.
[3] LARSEN, R.: *The multiplier Problem,* Lecture Notes in Mathematics. No 105. Springer Verlag 1969.
[4] FIGÁ-TALAMANCA, A. and GAUDRY, G.: Multipliers and sets of uniqueness of L_p, *Michigan Math. J.* (1970).
[5] MÁTÉ, L.: Multipliers and topological tensor product (preliminary report). *Notices of the AMS* **14** oct 1967. No T-366.
[6] MÁTÉ, L.: The Arens product and multiplier operators. *Studia Math.* **28** (1967), 227—234.
[7] MÁTÉ, L.: Embedding multiplier operators of a Banach algebra B into its second conjugate space B^{**}, *Bull. Ac. Polonaise* **13** (1965), 809—812.
[8] MÁTÉ, L.: On the connections between approximation by convolution and the multiplier problem, *Studia Sci. Math. Hung.* **6** (1971), 75—79.
[9] POWELL, J.: Multipliers on locally convex algebras, *Studia Math.* (to appear).
[10] RIEFFEL, M.: Induced Banach representations of Banach algebras and locally compact groups. *J. of Functional Analysis,* **1** (1967), 443—491.
[11] SIMPSON, J.: *Locally convex algebras, Arens multiplication,* Lexington, Ky. 1968. (Mimeographed).
[12] CIVIM, P. and YOOD, B.: The second conjugate space of a Banach algebra as an algebra, *Pac. J. Math.* **11** (1961), 847—870.
[13] FIGÁ—TALAMANCA, A. and GAUDRY, G. I.: Density and representation theorems for multipliers of type (p, q), *J. Austr. Math. Soc.* **7** (1967), 1—9.

Technical University, Budapest

(Received January 20, 1971, revised May 21, 1973.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 193—198.

ON THE RAREFACTION OF RENEWAL PROCESSES. V.

by

J. MOGYORÓDI

8. We have described a possible rarefaction procedure in section 2, paper I. We can generalize this model as follows: let $Z_i^{(n)}$ $(i, n=1, 2, \ldots)$ be a double-entry system of random variables taking the positive integers as values and suppose that for a fixed n the variables $Z_i^{(n)}$ are independent and identically distributed. Let further $\tau_0 \equiv 0 \leqq \tau_1 \leqq \tau_2 \leqq \ldots$ be the original renewal process to be rarefied. The n-th step of rarefraction is then defined as follows: the i-th renewal point of the n-th step of the rarefaction, denoted by $\tau_i^{(n)}$, is

$$\tau_0^{(n)} \equiv 0, \quad \tau_i^{(n)} = \tau_{Z_1^{(n)}+Z_2^{(n)}+\ldots+Z_i^{(n)}} \qquad (i = 1, 2, \ldots)$$

where on the right-hand side we have put that renewal point of the original renewal process, the index of which is $Z_1^{(n)}+\ldots+Z_i^{(n)}$. We shall also suppose that the variables $Z_i^{(n)}$ $(i=1, 2, \ldots)$ and the original renewal process are independent.

It is again easy to see that the differences

$$\tau_i^{(n)} - \tau_{i-1}^{(n)} \qquad (i = 1, 2, \ldots)$$

are independent and identically distributed random variables with the common distribution function

$$\sum_{k=1}^{\infty} \mathsf{P}(Z_i^{(n)} = k) F^{(*k)}(x),$$

where $F^{(*k)}(x)$ is the k-fold convolution of the distribution function $F(x)$ of the original renewal process.

We will show in this paper that under some conditions the asymptotic behaviour is distribution of $\tau_i^{(n)} - \tau_{i-1}^{(n)}$, as $n \to +\infty$, is independent of the distribution function $F(x)$ and depends only on the asymptotic behaviour of the variables $Z_i^{(n)}$. Namely, we prove the following

THEOREM 17. *If the mean-value μ of the original renewal process exists and if for a sequence of positive numbers B_n (where B_n tends to infinity as $n \to +\infty$) we have for every fixed i*

$$\lim_{n \to +\infty} \mathsf{P}\left(\frac{Z_i^{(n)}}{B_n} < x\right) = G(x),$$

where $G(x)$ is a distribution function, then

$$\lim_{n \to +\infty} \mathsf{P}\left(\frac{\tau_i^{(n)} - \tau_{i-1}^{(n)}}{\mu \beta_n} < x\right) = G(x).$$

Moreover, every distribution which is concentrated on $[0, +\infty)$ *can occur as a possible limit distribution for this model of rarefaction.*

PROOF. The random variable $\tau_i^{(n)} - \tau_{i-1}^{(n)}$ is a sum of the independent and identically distributed random variables $\xi_j = \tau_j - \tau_{j-1}$ $(j=1, 2, \ldots)$, where the number of summands is $Z_i^{(n)}$; $Z_i^{(n)}$ is independent of the summands. Now Theorem 1 of section 1 can be applied. We obtain thus the first assertion of our theorem. Concerning the second, let us put an arbitrary distribution function $G(x)$, which is concentrated on the positive axis. Put, further, a sequence $z_1, z_2, \ldots$ of non-negative, independent random variables which are distributed according to $G(x)$ and define

$$Z_i^{(n)} = [nz_i] \qquad (i, n = 1, 2, \ldots)$$

where the symbol $[\cdot]$ means the integer part. It is obvious that for fixed n the variables $Z_i^{(n)}$ are independent and identically distributed. Also it can be easily seen that

$$\lim_{n \to +\infty} \mathsf{P}\left(\frac{Z_i^{(n)}}{n} < x\right) = G(x).$$

Now if the mean-value μ of the original renewal process is finite we obtain by Theorem 1 of section 1, that

$$\lim_{n \to +\infty} \mathsf{P}\left(\frac{\tau_i^{(n)} - \tau_{i-1}^{(n)}}{\mu n} < x\right) = G(x).$$

This proves our theorem.

Before formulating our next assertion we shall need some notions and auxiliary results. Let us consider a double sequence

$$\xi_{n1}, \xi_{n2}, \ldots, \xi_{nv_n} \qquad (n = 1, 2, \ldots)$$

where v_n is a positive integer random variable, which is independent of the variables ξ_{nk} $(k=1, 2, \ldots)$ and the variables ξ_{nk} $(k=1, 2, \ldots)$ are also independent. We suppose that $\mathsf{M}(v_n) < +\infty$, $n=1, 2, \ldots$. We call that the variables

$$\zeta_{v_n}^{(n)} = \xi_{n1} + \xi_{n2} + \ldots + \xi_{vn_n} \qquad (n = 1, 2, \ldots)$$

form a stable sequence if there exists a double sequence $A_k^{(n)}$ $(k, n=1, 2, \ldots)$ of numbers such that

$$\zeta_{v_n}^{(n)} - A_{v_n}^{(n)}$$

converges in probability to zeros as $n \to +\infty$. About stability of the sequence $\zeta_{v_n}^{(n)}$ we have proved the following result which we formulate now as ([1]).

LEMMA 4. In order that the variables $\zeta_{v_n}^{(n)}$ be stable it is (necessary and) sufficient that the relation

$$\lim_{n \to +\infty} \mathsf{M}\left(\sum_{k=1}^{v_n} \int_{-\infty}^{+\infty} \frac{x^2}{1+x^2}\, dF_{nk}(x + m_{nk})\right) = 0$$

holds, where $F_{nk}(x)$ is the distribution function of the variable ξ_{nk} and m_{nk} denotes the median of ξ_{nk}. The variables $A_{\nu_n}^{(n)}$ can be chosen is this case as follows

$$A_{\nu_n}^{(n)} = \sum_{k=1}^{\nu_n} \left(m_{nk} + \int_{|x| \leqq 1} x \, dF_{nk}(x + m_{nk}) \right).$$

Let us consider now a sequence of independent and identically distributed random variables $\xi_1, \xi_2, \ldots$ and put

$$\zeta_{\nu_n}^{(n)} = \frac{1}{B_n} \sum_{j=1}^{\nu_n} \xi_j .$$

The result of the preceding lemma can now be formulated in the following form

COROLLARY. *The sequence*

$$\frac{1}{B_n} \sum_{j=1}^{\nu_n} \xi_j , \qquad (B_n \to +\infty),$$

is stable and only if

$$\lim_{n \to +\infty} \mathsf{M}(\nu_n) \int_{-\infty}^{+\infty} \frac{x^2}{B_n^2 + x^2} \, dF(x + m) = 0,$$

where $F(x)$ is the common distribution function of the random variables ξ_j and m denotes their median. The variables $A_{\nu_n}^{(n)}$ can now be chosen according to the formula

$$A_{\nu_n}^{(n)} = \frac{\nu_n}{B_n} \left(m + \int_{|x| \leqq B_n} x \, dF(x + m) \right), \qquad (n = 1, 2, \ldots).$$

Let us now consider the original reneval process $\tau_0 \equiv 0 \leqq \tau_1 \leqq \tau_2 \leqq \ldots$ and consider the rarefaction model described in this paper.

THEOREM 18. *If the variables $Z_i^{(n)}$ have finite mean-value, $\mathsf{M}(Z_i^{(n)}) < +\infty$, further if for a sequence B_n of constants such that $B_n \to +\infty$ we have*

$$\lim_{n \to +\infty} \mathsf{M}(Z_i^{(n)}) \int_{-\infty}^{+\infty} \frac{x^2}{B_n^2 + x^2} \, dF(x + m) = 0,$$

where

$$F(x) = \mathsf{P}(\tau_i - \tau_{i-1} < x) \qquad (i = 1, 2, \ldots)$$

and m is the median of the variables $\tau_i - \tau_{i-1}$, then the variables

$$\frac{\tau_i^{(n)} - \tau_{i-1}^{(n)}}{B_n}$$

have a limit distribution if and only if the random variables

$$\frac{Z_i^{(n)}}{B_n} \left(m + \int_{|x| \leqq B_n} x \, dF(x + m) \right)$$

have limit distribution. In this case the limit distribution is the same for the two sequences.

Moreover, the random variables

$$\frac{\tau_i^{(n)} - \tau_{i-1}^{(n)}}{B_n}$$

converge in probability to a random variable if and only if the variables

$$\frac{Z_i^{(n)}}{B_n}\left(m + \int\limits_{|x| \leqq B_n} x\, dF(x+m)\right)$$

converge in probability. The limit random variables are in this case the same for the two sequences.

PROOF. The assumption

$$\lim_{n \to +\infty} \mathsf{M}(Z_i^{(n)}) \int\limits_{-\infty}^{+\infty} \frac{x^2}{B_n^2 + x^2}\, dF(x+m) = 0$$

implies that the variables

$$\frac{\tau_i^{(n)} - \tau_{i-1}^{(n)}}{B_n} - \frac{Z_i^{(n)}}{B_n}\left(m + \int\limits_{|x| \leqq B_n} x\, dF(x+m)\right) \qquad (n = 1, 2, \ldots)$$

converge in probability to zero. Thus, if one of the variables

$$\frac{\tau_i^{(n)} - \tau_{i-1}^{(n)}}{B_n}$$

and

$$\frac{Z_i^{(n)}}{B_n}\left(m + \int\limits_{|x| \leqq B_n} x\, dF(x+m)\right)$$

converges in probability, or in distribution, it does the same the other.

A special case of the Corollary of Lemma 4 is when we put $B_n = \mathsf{M}(\nu_n)$. We have then the following

THEOREM 19. *Let $\xi_1, \xi_2, \ldots$ be a sequence of independent and identically distributed random variables and ν_n $(n = 1, 2, \ldots)$ a sequence of positive integer-valued random variables ξ_i. Let us suppose further that $\mathsf{M}(\nu_n)$ exists. Then the random variables*

$$\zeta_{\nu_n}^{(n)} = \frac{1}{\mathsf{M}(\nu_n)} \sum_{j=1}^{\nu_n} \xi_j \qquad (\mathsf{M}(\nu_n) \to +\infty)$$

are stable if and only if

$$\lim_{n \to +\infty} \mathsf{M}(\nu_n)\mathsf{P}(|\xi_j - m| \geqq \mathsf{M}(\nu_n)) = 0.$$

In this case the random variables $A_{\nu_n}^{(n)}$ can be chosen as follows:

$$A_{\nu_n}^{(n)} = \frac{\nu_n}{\mathsf{M}(\nu_n)}\left(m + \int\limits_{|x| \leqq \mathsf{M}(\nu_n)} x\, dF(x+m)\right).$$

PROOF. By our lemma the necessary and sufficient condition for

$$\frac{1}{\mathsf{M}(v_n)} \sum_{i=1}^{v_n} \xi_i$$

to be stable is the following

$$\lim_{n \to +\infty} \mathsf{M}(v_n) \int_{-\infty}^{\infty} \frac{x^2}{\mathsf{M}^2(v_n) + x^2} \, dF(x+m) = 0.$$

We have thus to show that this is equivalent to our condition. It is known ([2]) that the last limit relation is equivalent to the following:

$$\lim_{n \to +\infty} \mathsf{M}(v_n) \mathsf{P}\big(|\xi_j - m| \geqq M(v_n)\big) = 0$$

and

$$\lim_{n \to +\infty} \frac{1}{\mathsf{M}(v_n)} \int_{|x| \leqq \mathsf{M}(v_n)} x^2 \, dF(x+m) = 0.$$

Thus we have to prove that the first limit relation implies the second one. Let for this purpose the positive integer $N = N(n)$ be defined such that

$$N \leqq \mathsf{M}(v_n) < N+1$$

hold. Then

$$\frac{1}{\mathsf{M}(v_n)} \int_{|x| \leqq \mathsf{M}(v_n)} x^2 \, dF(x+m) \leqq \frac{1}{N} \int_{|x| \leqq N+1} x^2 \, dF(x+m).$$

On the other hand

$$\int_{|x| \leqq N+1} x^2 \, dF(x+m) = \sum_{k=1}^{N+1} \int_{(k-1) \leqq |x| < k} x^2 \, dF(x+m) \leqq$$

$$\sum_{k=1}^{N+1} k^2 \mathsf{P}(k-1 \leqq |\xi_j - m| < k) \leqq 2 \sum_{k=1}^{N+1} \left(\sum_{i=1}^{k} i \right) \mathsf{P}\big((k-1) \leqq |\xi_j - m| < k\big) =$$

$$= 2 \sum_{i=1}^{N+1} i \mathsf{P}(i-1 \leqq |\xi_j - m| < N+1) \leqq 2 \sum_{i=1}^{N+1} i \mathsf{P}(i-1 \leqq |\xi_j - m|).$$

Thus

$$\frac{1}{N} \int_{|x| \leqq N+1} x^2 \, dF(x+m) \leqq \frac{2}{N} \sum_{i=1}^{N+1} i \mathsf{P}\big(|\xi_j - m| \geqq i-1\big).$$

By the definition of N we have

$$\mathsf{M}(v_n) \int_{|x| \geqq \mathsf{M}(v_n)} dF(x+m) = \mathsf{M}(v_n^v) \mathsf{P}\big(|\xi_j - m| \geqq \mathsf{M}(v_n)\big) \geqq N \mathsf{P}(|\xi_j - m| \geqq N+1).$$

So, by our assumption

$$N \mathsf{P}(|\xi_j - m| \geqq N+1) \to 0.$$

It follows from this that

$$\frac{1}{\mathsf{M}(v_n)} \int\limits_{|x| \leqq \mathsf{M}(v_n)} x^2 \, dF(x+m) \leqq \frac{1}{N} \int\limits_{|x| \leqq N+1} x^2 \, dF(x+m) \to 0.$$

This proves our assertion.

This result enables us to formulate the following result for the rarefaction model described in this paper.

THEOREM 20. *If the random variables $Z_i^{(n)}$ have finite mean-value and $\mathsf{M}(Z_i^{(n)}) \to +\infty$ as $n \to +\infty$, further if*

$$\lim_{n \to +\infty} \mathsf{M}(Z_i^{(n)}) \mathsf{P}\left(|\tau_i - \tau_{i-1} - m| \geqq \mathsf{M}(Z_i^{(n)})\right) = 0,$$

then the random variables

$$\frac{\tau_i^{(n)} - \tau_{i-1}^{(n)}}{M(Z_i^{(n)})}$$

have a limit distribution if and only if the random variables

$$\frac{\mathsf{M}(Z_i^{(n)})}{Z_i^{(n)}} \left(m + \int\limits_{|x| \leqq \mathsf{M}(Z_i^{(n)})} x \, dF(x+m)\right)$$

have a limit distribution.

Moreover, the last assertion remains true if we replace the words "have a limit distribution" by the words "converge in probability".

The PROOF of this theorem is the same as that of Theorem 18.

REMARK. The condition

$$\lim_{n \to +\infty} \mathsf{M}(v_n) \mathsf{P}\left(|\xi_j - m| \geqq \mathsf{M}(v_n)\right) = 0$$

of Theorem 19 and the corresponding condition of Theorem 20 trivially holds if the mean-value $\mathsf{M}(\xi_j)$ exists.

REFERENCES

[1] Бикялис, А. П., Модьороди, Й: Об одной задаче Б. В. Гнеденко. *Acta Scientiarum Mathematicarum. Szeged.* **30** (1969) 241—245.

[2] Гнеденко, Б. В., Колмогоров, А. Н.: *Предельные распределения для сумм независимых случайных величин.* Гостехиздат. Москва, 1949.

Eötvös L. University, Chair of Probability Theory

(Received March 10, 1971.)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 199—209.

ON THE RAREFACTION OF RENEWAL PROCESSES. VI

by

J. MOGYORÓDI

9. The following rarefaction model often arises in physical applications. Let $\tau_0 \equiv 0 \leqq \tau_1 \leqq \tau_i \leqq \ldots$ be a renewal process with finite mean-value μ, and $0 \leqq \varrho(x) \leqq 1$ be a function defined on the halaxis $[0, +\infty)$. Under the condition that $\tau_1 = x$ the renewal point τ_1 remains in the process with probability $\varrho(x)$ and is deleted with probability $1-\varrho(x)$. Independently of this we make the same with the renewal point τ_i, i.e. if $\tau_i = y$ then τ_i remains in the process with probability $\varrho(y)$ and will be deleted with probability $1-\varrho(y)$. The same operation is made for every renewal point τ_i and for different indices i it is made independently each of other. This model of rarefaction will be called the "position-dependent rarefaction".

We shall suppose that $\varrho(x)$ is continuous. Let ν denote the minimal index i ($i=1, 2, \ldots$) such that after the position-dependent rarefaction the renewal point τ_i remains in the process. We shall be interested in studying the behaviour of ν and τ_ν. The quantity $\nu-1$ indicates the number of the rarefied renewal points and τ_ν is the time interval spent between $\tau_0 \equiv 0$ and the first renewal point not deleted in our rarefaction procedure.

If $\varrho(x) \equiv 0$ then each renewal point of the original renewal process is rarefied with probability 1, while if $\varrho(x) \equiv 1$ then there is no rarefaction with probability 1. The last case shows that the random variable ν can take on the value $+\infty$ with probability 1.

First we shall be interested in the distribution of the random variable ν. The probability that $\nu = n$ ($n=1, 2, \ldots$) can be calculated as follows. Let $F(x)$ denote the distribution function of $\tau_i - \tau_{i-1}$ ($i=1, 2, \ldots$) and $Q_n(x)$ the probability

$$\mathsf{P}(\tau_n < x,\ \nu \geqq n).$$

Then evidently

$$Q_1(x) = F(x),$$

and for $n = 2, 3, \ldots$ we have the following recursive formula

$$Q_n(x) = \int_0^x F(x-y)\big(1-\varrho(y)\big)\, dQ_{n-1}(y).$$

In fact, under the condition that the renewal points $\tau_1, \tau_2, \ldots, \tau_{n-2}$ are deleted (except $\tau_0 \equiv 0$) and $\tau_{n-1} = y$ the probability that τ_{n-1} will also be deleted from the process and at the same time $\tau_n < x$, equals $F(x-y)\big(1-\varrho(y)\big)$. The recursive formula can be obtained now by using the theorem of total probabilities.

The knowledge of the probabilities $Q_n(x)$ $(n=1, 2, \ldots)$ enables us now to calculate the others. For example we have

$$\mathsf{P}(\tau_\nu < x,\ \nu = n) = \int_0^x \varrho(y)\,dQ_n(y), \qquad (n = 1, 2, \ldots),$$

$$\mathsf{P}(\nu \geqq n) = \lim_{x \to +\infty} Q_n(x),$$

further

$$\mathsf{P}(\nu = n) = \int_0^{+\infty} \varrho(y)\,dQ_n(y).$$

We call that the variable X is proper if $\mathsf{P}(X < +\infty) = 1$.

THEOREM 21. *We have*

$$\mathsf{P}(\nu \geqq n) = \mathsf{M}\left(\prod_{i=1}^{n-1} (1 - \varrho(\tau_i))\right), \qquad (n = 1, 2, \ldots)$$

where the symbol M *denotes mathematical expectation and the empty product means* 1.

In order that the random variable ν *be proper it is necessary and sufficient that the random variables*

$$\prod_{i-1}^{n-1} (1 - \varrho(\tau_i))$$

converge to zero with probability 1. *If* ν *is proper then the integral*

$$\int_0^{+\infty} \varrho(x)\,dH(x)$$

diverges, where $H(x)$ *is the renewal function of the original renewal process.*[1]

PROOF. The event $\{\nu \geqq 1\}$ is of probability 1. For $n=2, 3, \ldots$ the event $\{\nu \geqq n\}$ consists in the delation from the original renewal process of the renewal points $\tau_1, \tau_2, \ldots, \tau_{n-1}$. The conditional probability

$$\mathsf{P}(\nu \geqq n | \xi_1, \xi_2, \ldots, \xi_{n-1})$$

is with probability 1 equal to

$$(1 - \varrho(\xi_1))(1 - \varrho(\xi_1 + \xi_2)) \ldots (1 - \varrho(\xi_1 + \ldots + \xi_{n-1})) = \prod_{i=1}^{n-1} (1 - \varrho(\tau_i)),$$

[1] Let $N(x)$ denote the number of the renewal points lying in the interval $(0, x)$. The renewal function is then defined as the mean-value of $N(x)$:

$$H(x) = \mathsf{M}(N(x)).$$

It is known that under our assumptions $H(x)$ is finite and

$$H(x) = \sum_{i=1}^{+\infty} F^{(i*)}(x),$$

where $F^{(i*)}(x)$ denotes the i-fold convolution of $F(x)$.

MAGYAR TUDOMÁNYOS AKADÉMIA KÖNYVTÁRA

where $\xi_i = \tau_i - \tau_{i-1}$ $(i=1, 2, \ldots)$. From this it follows that

$$\mathsf{P}(\nu \geqq n) = \mathsf{M}\big(\mathsf{P}(\nu \geqq n | \xi_1, \ldots, \xi_{n-1})\big) = \mathsf{M}\left(\prod_{i=1}^{n-1} (1 - \varrho(\tau_i))\right),$$

which proves our first assertion. The random variable ν is proper if and only if

$$\lim_{n \to +\infty} \mathsf{P}(\nu \geqq n) = 0.$$

Now the sequence $\prod_{i=1}^{n-1} (1-\varrho(\tau_i))$ monotonically decreases and is bounded with probability 1. Consequently, $\mathsf{P}(\nu \geqq n) = \mathsf{M}\left(\prod_{i=1}^{n-1} (1-\varrho(\tau_i))\right) \to 0$, if and only if we have with probability 1

$$\lim_{n \to +\infty} \prod_{i=1}^{n-1} (1 - \varrho(\tau_i)) = 0.$$

This can be shown by using Lebesgue's and Fatou's convergence theorem. This is our second assertion.

If the last limit relation holds then we have with probability 1

$$\lim_{n \to +\infty} \sum_{i=1}^{n} \varrho(\tau_i) = +\infty.$$

From this, using the Beppo Levi theorem, we conclude that the sequence

$$a_n = \sum_{i=1}^{n} \mathsf{M}\big(\varrho(\tau_i)\big) = \sum_{i=0}^{n} \int_0^{+\infty} \varrho(x)\, dF^{(i*)}(x)$$

diverges, where $F^{(i*)}(x)$ is the i-fold convolution of $F(x)$. Now the divergence of a_n implies the divergence of

$$\int_0^{+\infty} \varrho(x)\, dH(x).$$

This proves our theorem.

The divergence of the last integral is thus a necessary condition for ν to be proper. The following sufficient condition is not so far from this and ensures also the existence of all moments of the random variable ν. This is expressed in

THEOREM 22. *Let us suppose that*

$$\liminf_{x \to +\infty} \varrho(x) = \varrho > 0.$$

Then the random variable ν is proper and has moments of every order.

PROOF. It suffices to show that the power series

$$G(z) = \sum_{n=1}^{\infty} \mathsf{P}(\nu \geqq n) z^n$$

has a convergence radius greater than 1. Let $x_0>0$ be such a number that for $x \geqq x_0$ the inequality

$$\varrho(x) > \frac{\varrho}{2}$$

hold. Let further $N(x_0)$ denote the number of the renewal points in the interval $(0, x_0)$. It is known that

$$\mathsf{P}(N(x_0) = k) = F^{(k*)}(x_0) - F^{((k+1)*)}(x_0), \qquad (k = 0, 1, 2, \ldots),$$

where

$$F^{(0*)}(x_0) = \begin{cases} 1, & \text{if} \quad x > 0. \\ 0, & \text{if} \quad x \leqq 0. \end{cases}$$

Let us suppose for the moment that $F(x_0)<1$ and choose a positive number $c<1$ such that

$$\max\left(F(x_0),\ 1-\frac{\varrho}{2}\right) < c < 1$$

be satisfied. If χ_k denotes the indicator of the event $\{N(x_0)=k\}$ then for $n \geqq 2$ we have the estimation

$$\mathsf{M}\left(\prod_{i=1}^{n-1}(1-\varrho(\tau_i))\right) = \sum_{k=0}^{n-1}\mathsf{M}\left(\prod_{i=1}^{n-1}(1-\varrho(\tau_i))\chi_k\right) + \sum_{k=n}^{\infty}\mathsf{M}\left(\prod_{i=1}^{n-1}(1-\varrho(\tau_i))\chi_k\right) \leqq$$

$$\leqq \sum_{k=0}^{n-1} c^{n-1-k}\mathsf{M}(\chi_k) + \sum_{k=n}^{\infty}\mathsf{M}(\chi_k).$$

Now

$$\mathsf{M}(\chi_k) = F^{(k*)}(x_0) - F^{((k+1)*)}(x_0) \leqq (F(x_0))^k = q^k,$$

where $q = F(x_0) < 1$. Thus

$$\mathsf{M}\left(\prod_{i=1}^{n-1}(1-\varrho(\tau_i))\right) \leqq c^{n-1}\sum_{k=0}^{n-1}\left(\frac{q}{c}\right)^k + \sum_{k=n}^{\infty} q^k \leqq \frac{c^{n-1}}{1-\frac{q}{c}} + \frac{q^n}{1-q}.$$

From this it follows that

$$|G(z)| = \left|\sum_{n=1}^{\infty}\mathsf{P}(\nu \geqq n)z^n\right| \leqq \sum_{n=1}^{\infty}{}'|z|^n\left(\frac{c^{n-1}}{1-\frac{q}{c}} + \frac{q^n}{1-q}\right) =$$

$$= \frac{|z|}{1-\frac{q}{c}}\frac{1}{1-|z|c} + \frac{q|z|}{(1-q)(1-q|z|)},$$

if $|z| \max(c, q) = |z|c < 1$. Thus $G(z)$ converges in the circle $|z|<1$.

If $F(x_0)=1$, then we can choose a positive integer $m>1$ such that $F^{(m*)}(x_0)<1$. Such an integer exists. By the aid of $q'=F^{(m*)}(x_0)$ the probability

$$\mathsf{P}(N(x_0)=k)=\mathsf{M}(\chi_k)$$

can be estimated by $q'^{\left[\frac{k}{m}\right]}$. From this point on the proof is the same as above.

Let us consider for example the homogeneous Poisson process as original renewal process. In this case

$$F(x)=\begin{cases}1-e^{-\lambda x}, & \text{if } x>0\\ 0\quad\ , & \text{if } x\leqq 0,\end{cases}$$

where $\lambda>0$ is a fixed number. Then

$$Q_1(x)=1-e^{-\lambda x}\qquad (x>0)$$

and using the recursion formula we see by induction that

$$Q_n(x)=\int_0^x \frac{\left(\lambda\int_0^y(1-\varrho(u))\,du\right)^{n-1}}{(n-1)!}\lambda e^{-\lambda y}\,dy\qquad (n=1,2,\ldots).$$

Now

$$\mathsf{P}(\nu\geqq n)=\int_0^{+\infty}\frac{\left(\lambda\int_0^y(1-\varrho(u))\,du\right)^{n-1}}{(n-1)!}\lambda e^{-\lambda y}\,dy$$

and

$$\mathsf{P}(\nu=n)=\int_0^{+\infty}\varrho(y)\frac{\left(\lambda\int_0^y(1-\varrho(u))\,du\right)^{n-1}}{(n-1)!}\lambda e^{-\lambda y}\,dy.$$

An easy calculation shows that

$$\mathsf{P}(\nu<+\infty)=\sum_{n=1}^{\infty}\mathsf{P}(\nu=n)=1-e^{-\lambda\int_0^{+\infty}\varrho(u)\,du}.$$

We have in this case

$$H(x)=\lambda x.$$

So, the divergence of the integral

$$\int_0^{+\infty}\varrho(x)\,dH(x)=\lambda\int_0^{+\infty}\varrho(x)\,dx$$

is in this case not only necessary but also sufficient for ν to be a proper random variable.

If for a $c>0$ we have

$$\int_0^y\varrho(u)\,du\geqq cy,$$

then the random variable ν has moments of every order. In fact, after a formal calculation we see that

$$\sum_{n=1}^{\infty} \mathsf{P}(\nu = n) z^n = \int_0^{+\infty} \lambda z \varrho(y) \exp\left(\lambda z \int_0^y (1 - \varrho(u))\, du\right) e^{-\lambda y}\, dy.$$

From this and from our assumption it follows that

$$\left| \sum_{n=1}^{\infty} \mathsf{P}(\nu = n) z^n \right| \leqq |z| \int_0^{+\infty} e^{-\lambda y(1-|z|(1-c))}\, dy,$$

which is convergent if $|z| < 1/(1-c)$.

Now we shall be interested in the distribution of τ_ν and of its moments. Formally, the distribution function of τ_ν can be expressed as follows:

$$\mathsf{P}(\tau_\nu < x) = \sum_{n=1}^{\infty} \mathsf{P}(\tau_n < x,\ \nu = n) = \sum_{n=1}^{\infty} \int_0^x \varrho(y)\, dQ_n(y).$$

It follows from this that τ is a proper random variable if and only if ν is proper.

In the case of the Poisson process, as original renewal process, we have

$$\mathsf{P}(\tau_\nu < x) = \int_0^x \varrho(y) \lambda e^{-\lambda y} e^{\lambda \int_0^y (1-\varrho(u))\, du} = 1 - e^{-\lambda \int_0^x \varrho(u)\, du} \qquad (x > 0).$$

The moments of τ_ν, if they exist, can be calculated by the aid of the so called Wald's identity. We deduce first this identity. Let us, for this purpose, denote by $\varphi(s)$ the Laplace transform of the random variables ξ_j $(j = 1, 2, \ldots)$:

$$\varphi(s) = \mathsf{M}(e^{-s\xi_j}) \qquad (s = \sigma + i\tau,\ \sigma \geqq 0),$$

and let η_n be the indicator of the event $\{\nu > n\}$ and γ_n the indicator of the event $\{\nu = n\}$. The event $\{\nu = n\}$ is independent of the variables $\xi_{n+1}, \xi_{n+2}, \ldots$. It follows from this that the event $\{\nu > n\}$ is also independent of the variables $\xi_{n+1}, \xi_{n+2}, \ldots$. We have thus the formula

$$\mathsf{M}(\eta_n e^{-s\tau_{n+1}}) = \varphi(s) \mathsf{M}(\eta_n e^{-s\tau_n}).$$

Applying this we obtain

$$\sum_{n=1}^{m} \varphi(s)^{-n} \mathsf{M}(\eta_{n-1} e^{-s\tau_n}) = 1 + \sum_{n=2}^{m} \varphi(s)^{-(n-1)} \mathsf{M}(\eta_{n-1} e^{-s\tau_{n-1}}) =$$

$$= 1 + \sum_{l=1}^{m-1} \varphi(s)^{-l} \mathsf{M}(\eta_l e^{-s\tau_l}).$$

On the other hand

$$\sum_{n=1}^{m} \varphi(s)^{-n} \mathsf{M}(\eta_{n-1} e^{-s\tau_n}) = \sum_{l=1}^{m} \varphi(s)^{-l} \mathsf{M}(e^{-s\tau_l} \gamma_l + e^{-s\tau_l} \eta_l).$$

Comparing the right-hand sides we obtain

$$\sum_{l=1}^{m} \varphi(s)^{-l} \mathsf{M}(e^{-s\tau_l}\gamma_l) = 1 - \varphi(s)^{-m} \mathsf{M}(\eta_m e^{-s\tau_m}). \tag{31}$$

We have thus proved ([4]).

THEOREM 23. *The random variables ν and τ_ν, determined by the position-dependent rarefaction, satisfy the Wald's identity*

$$\mathsf{M}\big(\varphi(s)^{-\nu} e^{-s\tau_\nu}\big) \equiv 1 \qquad (s = \sigma + i\tau,\ \sigma \geqq 0)$$

if and only if

$$\lim_{m \to +\infty} \varphi(s)^{-m} \mathsf{M}(\eta_m e^{-s\tau_m}) = 0.$$

A sufficient condition is given in

THEOREM 24. *Let us suppose that $\varrho(x)$ is such that*

$$\liminf_{x \to +\infty} \varrho(x) = \varrho > 0.$$

Then Wald's identity

$$\mathsf{M}\big(\varphi(s)^{-\nu} e^{-s\tau_\nu}\big) \equiv 1$$

holds for those real s for which $0 \leqq s < \varphi^{-1}(1-\varrho)$.

PROOF. On the basis of (31) it is enough to show that for $0 \leqq s < \varphi^{-1}(1-\varrho)$ the limit relation

$$\lim_{n \to +\infty} \frac{\mathsf{M}(\eta_m)}{\varphi(s)^m} = \lim_{m \to +\infty} \frac{\mathsf{P}(\nu \geqq m+1)}{\varphi(s)^m} = 0$$

holds. For this purpose it sufficies to prove that

$$G(z) = \sum_{n=1}^{\infty} \mathsf{P}(\nu \geqq n) z^n$$

converges at the point $z = 1/\varphi(s)$. This is satisfied for those $s \geqq 0$ for which $\varphi(s) = 1$. Let us suppose that $0 \leqq s < \varphi^{-1}(1-\varrho)$ and $0 < \varphi(s) < 1$. Then

$$1 > \varphi(s) > 1 - \varrho.$$

Let $\varepsilon > 0$ be such that

$$1 - \varrho + \varepsilon < \varphi(s) < 1$$

be satisfied and define the number x_0 such that for $x \geqq x_0$ the inequality $\varrho(x) \geqq \varrho - \frac{\varepsilon}{2}$ be satisfied. Further, let us choose the index k such that $F^{(k*)}(x_0) < 1 - \varrho + \varepsilon$ hold. Now the argumentation of the proof of Theorem 22 can be repeated by choosing $c = 1 - \varrho + \varepsilon$. Since in this case $\frac{1}{\varphi(s)} < \frac{1}{c}$, we see that at the point $z = \frac{1}{\varphi(s)}$ the power series converges.

The derivatives of Wald's identity at the point $s = 0$ give the moments of the random variable τ_ν. The following sufficient condition makes this procedure legitimate.

THEOREM 25. *Let us suppose that*

$$\int_0^{+\infty} x^k \, dF(x) < +\infty.$$

If

$$\lim_{m \to +\infty} m^k \mathsf{M}(\eta_m \tau_m) = 0,$$

then for $j=0, 1, 2, \ldots, k$ *we have*

$$\mathsf{M}\left(\frac{d^j}{ds^j} \varphi(s)^{-\nu} e^{-s\tau_\nu} \right)\Big|_{s=0} = \begin{cases} 1, & \text{if } j = 0 \\ 0, & \text{if } j \geqq 1. \end{cases}$$

PROOF. The formal derivatives of (31) at $s=0$ give

$$\frac{d^j}{ds^j} \sum_{l=1}^{m} \varphi(s)^{-l} \mathsf{M}(e^{-s\tau_l} \gamma_l)\Big|_{s=0} = \frac{d^j}{ds^j} \left(1 - \varphi(s)^{-m} \mathsf{M}(\eta_m e^{-s\tau_m})\right)\Big|_{s=0}.$$

To prove our assertion we have to show that

$$\frac{d^j}{ds^j} \varphi(s)^{-m} \mathsf{M}(\eta_m e^{-s\tau_m})\Big|_{s=0} \to 0,$$

if $m \to +\infty$; $j=0, 1, 2, \ldots, k$. But this is ensured by our conditions.

In many applications it often happens that the probability $\varrho(x)$ is uniformly near to 0. This means that ν is with great probability great. To investigate this case we shall suppose that $\varrho_k(x)$ is such a sequence $(k = 1, 2, \ldots)$ of rarefying probabil ities which converges uniformly to zero. Let ν_k denote the corresponding random variable, i.e. the first index i in the position-dependent rarefaction characterized by t he probability $\varrho_k(x)$, for which the corresponding τ_i remains in the process, $(i=1, 2, \ldots)$. Let further

$$r_k = \inf_x \varrho_k(x), \quad R_k = \sup_x \varrho_k(x).$$

To make precise our investigations we give

THEOREM 26. *If* $r_k > 0$, $k=1, 2, \ldots$ *and* $\lim_{k \to +\infty} R_k = 0$, *further*

$$\lim_{k \to +\infty} \frac{R_k}{r_k} = 1$$

then we have

$$\lim_{k \to +\infty} \mathsf{P}(\nu_k < x\mathsf{M}(\nu_k)) = 1 - e^{-x} \qquad (x > 0).$$

PROOF. Since

$$\mathsf{P}(\nu_k \geqq n) = \mathsf{M}\left(\prod_{i=1}^{n-1} (1 - \varrho_k(\tau_i)) \right) \qquad (k, n = 1, 2, \ldots),$$

we have

$$(1 - R_k)^{n-1} \leqq \mathsf{P}(\nu_k \geqq n) \leqq (1 - r_k)^{n-1}.$$

Further, sinece $r_k>0$, the mean-value $\mathsf{M}(v_k)$ exists and

$$\frac{1}{R_k} \leqq \mathsf{M}(v_k) \leqq \frac{1}{r_k}.$$

Thus

$$(1-R_k)^{\frac{x}{r_k}} \leqq \mathsf{P}(v_k \geqq xM(v_k)) \leqq (1-r_k)^{\frac{x}{R_k}}.$$

Taking into account our conditions we obtain from the assertion.

We formulate now another system of sufficient conditions. Let us consider the sequence $\varrho_k(x)$ of the rarefying probabilities which satisfies the following

CONDITION a. For $k=1, 2, \ldots$ there exists the limit

$$\lim_{x\to+\infty} \varrho_k(x) = \varrho_k,$$

where $\varrho_k>0$, $k=1, 2, \ldots$, and this is uniform in k, i.e. for an arbitrary $\varepsilon>0$ there exists a number $x_0=x_0(\varepsilon)$, not depending on k, such that for $x\geqq x_0$ we have

$$|\varrho_k(x)-\varrho_k| < \varepsilon.$$

THEOREM 27. *If* $\lim_{k\to+\infty} R_k=0$ *and condition* a. *is satisfied, then*

$$\lim_{k\to+\infty} \mathsf{P}\big(v_k < xM(v_k)\big) = 1-e^{-x} \qquad (x>0).$$

PROOF. Let us choose x_0 such that for $x\geqq x_0$ the inequality

$$\frac{1-\varrho_k-\varepsilon}{1-\varepsilon} \leqq 1-\varrho_k(x) \leqq \frac{1-\varrho_k+\varepsilon}{1+\varepsilon}$$

be satisfied and let χ_l denote the indicator of the event $\{N(x_0)=l\}$. Then

$$\mathsf{M}\left(\prod_{i=1}^{n-1}(1-\varrho_k(\tau_i))\right) = \sum_{l=0}^{n-1}\mathsf{M}\left(\prod_{i=1}^{n-1}(1-\varrho_k(\tau_i))\chi_l\right) + \sum_{l=n}^{\infty}\mathsf{M}\left(\prod_{i=1}^{n-1}(1-\varrho_k(\tau_i))\chi_l\right).$$

Taking into account the above inequalities we obtain

$$\mathsf{M}\left(\prod_{i=1}^{n-1}(1-\varrho_k(\tau_i))\right) \leqq \sum_{l=0}^{n-1}\left(\frac{1-\varrho_k+\varepsilon}{1+\varepsilon}\right)^{n-1-l}\mathsf{P}\big(N(x_0)=l\big)+\mathsf{P}\big(N(x_0)\geqq n\big),$$

and

$$\mathsf{M}\left(\prod_{i=1}^{n-1}(1-\varrho_k(\tau_i))\right) \geqq \sum_{l=0}^{n-1}\left(\frac{1-\varrho_k-\varepsilon}{1-\varepsilon}\right)^{n-1-l}(1-R_k)^l\mathsf{P}\big(N(x_0)=l\big).$$

Since

$$\mathsf{M}(v_k) = \sum_{n=1}^{\infty}\mathsf{P}(v_k\geqq n) = \sum_{n=1}^{\infty}\mathsf{M}\left(\prod_{i=1}^{n-1}(1-\varrho_k(\tau_i))\right),$$

we obtain by summing the preceding inequalities with respect to n and changing the order of summation that

$$\frac{(1-\varepsilon)^2}{\varrho_k} \leqq \mathsf{M}(v_k) \leqq \frac{1+\varepsilon}{\varrho_k}+H(x_0).$$

It follows from this that

$$(1-\varepsilon)^2 \leqq \mathsf{M}(\nu_k)\varrho_k \leqq 1+\varepsilon+\varrho_k H(x_0).$$

Since x_0 is fixed, we obtain finally by choosing k suitably, that

$$1-2\varepsilon \leqq \mathsf{M}(\nu_k)\varrho_k \leqq 1+2\varepsilon.$$

So, we have to investigate the asymptotic behaviour of the probability

$$\mathsf{P}\left(\nu_k \geqq \frac{x}{\varrho_k}\right) \qquad (x>0)$$

as $k \to +\infty$. Let for this purpose the positive integer p determined such that

$$\mathsf{P}(N(x_0) \geqq p) < \varepsilon$$

be satisfied. Then similarly as above we can write

$$\sum_{l=0}^{p}\left(\frac{1-\varrho_k-\varepsilon}{1-\varepsilon}\right)^{\left[\frac{x}{\varrho_k}\right]-l}\mathsf{P}(N(x_0)=l) \leqq \mathsf{P}\left(\nu_k \geqq \frac{x}{\varrho_k}\right) \leqq$$

$$\leqq \sum_{l=0}^{p}\left(\frac{1-\varrho_k+\varepsilon}{1+\varepsilon}\right)^{\left[\frac{x}{\varrho_k}\right]-l}\mathsf{P}(N(x_0)=l)+\varepsilon.$$

Now, if $k \to +\infty$, the left and right hand sides of this inequality converge to

$$e^{-\frac{x}{1-\varepsilon}}\left(1-\mathsf{P}(N(x_0)>p)\right)$$

and

$$e^{-\frac{x}{1+\varepsilon}}\left(1-\mathsf{P}(N(x_0)>p)\right)$$

respectively. Taking into account this we see that

$$e^{-\frac{x}{1-\varepsilon}}-\varepsilon \leqq \lim_{k\to+\infty}\inf_{\sup}\mathsf{P}\left(\nu_k \geqq \frac{x}{\varrho_k}\right) \leqq e^{-\frac{x]}{1+\varepsilon}}+\varepsilon.$$

Since $\varepsilon>0$ was chosen arbitrarily we obtain from this inequality our assertion.

The knowledge about the asymptotic behaviour in distribution of the random variables ν_k enables us to formulate a limit distribution theorem concerning τ_{ν_k}. Before formulating this we need a result of [2] and [3], page 148, which we give here as

LEMMA 5. *Let $\xi_1, \xi_2, \ldots$ be a sequence of independent and identically distributed random variables with the finite mean-value μ. Let further ν_n $(n=1, 2, \ldots)$ be a sequence of positive integer-valued random variables, which are not not necessarily independent from those of the sequence $\xi_1, \xi_2, \ldots$. If for a sequence B_n $(B_n \to +\infty)$ we have*

$$\lim_{n\to+\infty}\mathsf{P}\left(\frac{\nu_n}{B_n} < x\right) = G(x),$$

where $G(x)$ is a distribution function $(G(+0) = 0)$, then the random variables

$$\frac{\xi_1 + \ldots + \xi_{\nu_n}}{\nu_n}$$

converge in probability to μ.

Now we have

THEOREM 28. *Under the condition of Theorem 26 or Theorem 27 we have*

$$\lim_{k \to +\infty} \mathsf{P}\left(\frac{\tau_{\nu_k}}{\mu \mathsf{M}(\nu_k)} < x\right) = 1 - e^{-x} \qquad (x > 0).$$

PROOF. Now

$$\tau_{\nu_k} = \xi_1 + \ldots + \xi_{\nu_k},$$

where $\xi_i = \tau_i - \tau_{i-1}$ $(i = 1, 2, \ldots)$. By our supposition $\mathsf{M}(\xi_i) = \mu > 0$ exists. So, by Lemma 5

$$\frac{\tau_{\nu_k}}{\nu_k}$$

converges in probability to μ. We have

$$\frac{\tau_{\nu_k}}{\mu \mathsf{M}(\nu_k)} = \frac{\tau_{\nu_k}}{\mu \nu_k} \cdot \frac{\nu_k}{\mathsf{M}(\nu_k)}.$$

It follows from an assertion of Cramér that under the conditions of Theorem 6 or Theorem 7

$$\lim_{k \to +\infty} \mathsf{P}\left(\frac{\tau_{\nu_k}}{\nu \mathsf{M}(\nu_k)} < x\right) = \lim_{k \to +\infty} \mathsf{P}\frac{\nu_k}{\mathsf{M}(\nu_k)} < x\Big) = 1 - e^{-x},$$

which proves our assertion.

REFERENCES

[1] CRAMÉR, H.: *Mathematical methods of statistics,* Princeton University Press, 1946.
[2] MOGYORÓDI, J.: On the law of large numbers for the sum of a random number of independent random variables *Annales Univ. Sci. Budapest,* **8** (1965) 33—38.
[3] RÉVÉSZ, P.: *The laws of large numbers,* Akadémiai Kiadó, Budapest, 1967.
[4] KEMPERMANN, J. H. B.: *The passage for a stationary Markov chain,* The University of Chicago Press, 1961.

Eötvös L. University, Chair of Probability Theory. Budapest

(Received March 10, 1971)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 211—224.

JOINT DISTRIBUTIONS BASED ON RUNS AND ON THE NUMBER OF INTERSECTIONS

by

SUJAN SRIVASTAVA

1. Introduction

A simple and elementary approach to the problem of the number of intersections not only on the axis but also at a height t and whith the endpoint of the random walk unspecified was given by CSÁKI [1], generalizing an earlier paper of his with VINCZE [2].

In [2] the authors determine the exact and the limiting distribution of the random variable $\lambda_{n,n}$, where $\lambda_{n,n}$ is the number of intersections of the random path on the axis (or $\lambda_{n,n}+1$ is the number of "waves"). The random variable $\lambda_{n,n}$ is equal to the number of intersections between the two empirical distribution functions. Further, they have considered the joint distribution of $\lambda_{n,n}$ and the Galton statistic, as well as that of the statistic $\lambda_{n,n}^{(t)}$ (intersection in the height t) and a modification of the Galton statistic. The joint distribution of $\lambda_{n,n}$ and the number of strictly positive terms in the sequence $(s_0, s_1, \ldots, s_{2n})$ has also been dealt with in [3]. Similar work has also been done by BATCHELET [4], by FELLER [5] and by MIHALEVICH [6].

KANWAR SEN in his paper [7] considered the distribution of the number of intersections, $\lambda_{n+k,n-k}$ and $\lambda_{n+k,n-k}^{(t)}$, when the path ends in some fixed point $(2n, 2k)$ other than the origin. He further generalized the results of [1] and [2].

In the present paper, the joint distribution and the joint limiting distribution of the statistic based on the number of runs and that based on the number of intersections have been derived. Further, certain results concerning the latter statistic obtained papers [2] and [7] have been generalized by pairing the statistic with that based on the number of runs. Concerning a possible statistical application of the results we refer to the paper written by REIMANN and VINCZE [8] in which — for samples of slightly different sizes — advantageous statistical properties of sample functions connected with the process $nF_n(x)-mG_m(x)$ are treated. The random quantities occurring in the present paper can also be considered as two-sample statistics in the same respect.

2. Notations

Following the usual notation, sequences $\theta=(\theta_1, \theta_2, \ldots, \theta_{m+n})$ of m $(+1)$'s and n (-1)'s will be considered, each possible sequence having the same probability $\binom{m+n}{n}\cdot\binom{m+n}{m}$. The partial sum of θ_i's will be denoted by s_i, i.e.,

$$s_i = \theta_1+\theta_2+\ldots+\theta_i, \qquad (i = 1, 2, \ldots, m+n)$$

$$s_0 = 0 \quad \text{and} \quad s_{m+n} = m-n.$$

The array $\{s_0, s_1, s_{2i}, \ldots s_{m+n}\}$ will be called the path of the particle. Thus to each sequence $(\theta_1, \theta_2, \ldots, \theta_{2i}, \ldots \theta_{m+n})$ there corresponds a random path, starting at the origin and reaching after $2n$ steps the point $(m+n, m-n)$, $(0<n\leqq m)$. A figure illustrating the path of the particle can be got by representating the points (i, s_i) in a plane, and then connecting each one of them with the next one.

In the sequel we shall consider the joint distributions of the pairs $(\lambda_{n,n}, R_{n,n})$, $(\lambda_{n+k,n-k}, R_{n+k,n-k})$, $(\lambda^{(t)}_{n,n}, R_{n,n})$ and $(\lambda^{(t)}_{n+k,n-k}, R_{n+k,n-k})$ where we denote by:

$\lambda_{m,n}$: the number of changes of sign in the sequence $s_1, s_2, \ldots, s_{m+n}$, $(s_{m+n} = = m-n)$ for random samples of sizes m and n.

$\lambda^{(t)}_{m,n}$: the number of changes os sign in the sequence $s_1-t, s_2-t, \ldots, s_{m+n}-t$ for random samples of sizes m and n.

$R_{m,n}$: the number of runs in the sequence $\theta_1, \theta_2, \ldots, \theta_{m+n}$

$R^{+}_{m,n}$: $\begin{cases} R_{m,n} & \text{if the first run is positive} \\ 0 & \text{if the first run is negative} \end{cases}$

$R^{-}_{m,n}$: $\begin{cases} R_{m,n} & \text{if the first run is negative} \\ 0 & \text{if the first run is positive} \end{cases}$

We shall employ the random walk model as suggested in [2] and [7] and consider paths $E_{2n,2k}$ starting at the origin and reaching after $2n$ steps the point $(2n, 2k)$, $0\leqq k\leqq n$.

It can easily be verified that there exists a 1:1 correspondence between $\lambda_{n+k,n-k}$ and the number of intersections of an $E_{2n,2k}$ path on the axis and between $\lambda^{(t)}_{n+k,n-k}$ and the number of intersections of an $E_{2n,2k}$ path on the height $y=t$.

If we have two samples, both of size n, coming from a common continuous distribution, we arrange the unified sample in order of magnitude. If the i-th value in this series comes from the first sample we define $\theta_i=1$, otherwise we put $\theta_i=-1$. In this case $\lambda_{n,n}$, $\lambda^{(t)}_{n,n}$ and $R_{n,n}$, as defined above, will be called number of intersections, number of intersections on the height t between the two empirical distribution functions and runs, respectively (see [2]).

We will use in the sequel the notion of the composed path which is different from the ordinary path as defined above. The composed path is up of runs $(\theta_{i_j+1}= =\theta_{i_j+2}=\ldots=\theta_{i_{j+1}}, j=0,1, \ldots, R-1, i_0=0)$ where two consecutive runs are not necessarily of different kind (i.e. $\theta_{i_j}\neq\theta_{i_j+1}$ need not hold). For specifying a composed path its runs are to be specified.

For case in writting we introduce the following symbols:

$T^{(t)}$-point: a point (i, s_i) of the path $E_{2n,2k}$ for which either $(s_{i-1} = = t-1, s_i = t, s_{i+1} = t+1)$ or $(s_{i-1} = t+1, s_i = t, s_{i+1} = = t-1)$ holds. This will be called the intersection point in the height $t(t > 0)$, $T^{(0)} = T$.

$F^{l}_{2n,2k}$: an $E_{2n,2k}$ path with exactly l T-points.

$F^{R,l}_{2n,2k}$: an $F^{l}_{2n,2k}$ path with exactly R runs.

$F^{R+,l}_{2n,2k}$: an $F^{R,l}_{2n,2k}$ path starting with a positive step.

$F^{R,l}_{2n,2k}$: an $F^{R,l}_{2n,2k}$ path starting with a negative step.

$F^{2r+,\cdot,j}_{2n,2k}$ ($F^{(2r+1)+,\cdot,j}_{2n,2k}$): a composed path from $(0, 0)$ to $(2n, 2k)$ starting with a positive step, having $r(r+1)$ positive runs and $(j+r)$ negative runs where the last j runs are negative and the remainder of the runs alternate, the $(j+1)$th run from the end being negative (positive).

$F^{2r+,l,j}_{2n,2k}$ ($F^{(2r+1)+,l,j}_{2n,2k}$): an $F^{2r+,\cdot,j}_{2n,2k}$ ($F^{(2r+1)+,\cdot,j}_{2n,2k}$) path having l T-points.

$F^{l}_{2n,2k,t}$: an $E_{2n,2k}$ path having l $T^{(t)}$-points.

$F^{R,l}_{2n,2k,t}$: an $F^{l}_{2n,2k,t}$ path having exactly R runs.

$$F^{R,l}_{2n,2k,0} = F^{R,l}_{2n,2k};\quad F^{R,0}_{2n,2k,0} = F^{R,0}_{2n,2k}.$$

$F^{R+,l}_{2n,2k,t}$: an $F^{R,l}_{2n,2k,t}$ path starting with a positive step.

$$F^{R+,0}_{2n,2k,0} = F^{R+,0}_{2n,2k}.$$

$F^{R-,l}_{2n,2k,t}$: an $F^{R,l}_{2n,2k,t}$ path starting with a negative step.

$$F^{R-,0}_{2n,2k,0} = F^{R-,0}_{2n,2k}.$$

$F^{2r+,l,j}_{2n,2k,t}$ ($F^{(2r+1)-,l,j}_{2n,2k}$): an $F^{2r+,\cdot,j}_{2n,2k}$ ($F^{(2r+1)+,\cdot,j}_{2n,2k}$) path with exactly l $T^{(t)}$ points, $t>0$.

G_{2n}: a path $\{s_0, s_1, \ldots, s_{2n}\}$ with the point of termination unspecified.

$G^{R,l}_{2n}$: a G_{2n} path having R runs and l T-points.

$G^{R,l}_{2n,t}$: a G_{2n} path having R runs and l $T^{(t)}$-points.

$N(A)$: number of all possible paths whose type is A.

3. Joint distributions of $(\lambda_{n+k,n-k}, R_{n+k,n-k})$ and $(\lambda_{n,n}, R_{n,n})$

LEMMA. *The number of composed paths* $F^{2r+,0,j}_{2n,2k}$ *is given by*

$$N(F^{2r+,0,j}_{2n,2k}) = \binom{n+k-1}{r-1}\binom{n-k-1}{r+j+1} - \binom{n+k-1}{r-2}\binom{n-k-1}{r+j} \qquad k>0 \tag{3.1}$$

$$N(F^{2r+,0,j}_{2n,0}) = \binom{n}{r-1}\binom{n}{r+j}\frac{j+1}{n} \tag{3.2}$$

PROOF. By considering the reversed path, it can be shown that

$$N(F^{2r+,0,j}_{2n,2k}) = N(F^{2r-,\cdot,j}_{2n,2k}) - N(F^{2r-,2k,j}_{2n,2k})$$

where $F^{2r-,\cdot,j}_{2n,2k}$ denotes a composed path from $(0, 0)$ to $(2n, 2k)$ having r positive runs and $(j+r)$ negative runs where the first $(j+1)$ runs are negative and the remainder of the runs alternate; the number of such path is $\binom{n+k-1}{r-1}\binom{n-k-1}{r+j-1}$, and $E^{2r-,2k,j}_{2n,2k}$ denotes an $E^{2r-,\cdot,j}_{2n,2k}$ path crossing $y=2k$. To determine their number we apply

a transformation at the last point of intersection of the path with $y=2k$, which we may refer to as "procedure 1" and is as follows. Change the order of the θ's of the segment from the origin to this point and then reflect the remaining portion of the path about $y=2k$. The resulting is an $E_{2n,2k}^{(2r-1)-,\cdot,j+1}$ path and the number of such paths is $\binom{n+k-1}{r-2}\binom{n-k-1}{r+j}$.

This gives the required result (3.1) for the case $k>0$. The same relation holds for the case $k=0$ too and hence we have

$$N(F_{2n,0}^{2r+,0,j}) = \binom{n-1}{r-1}\binom{n-1}{r+j-1} - \binom{n-1}{r-2}\binom{n-1}{r-2}$$

which on simplification reduces to (3.2).

For an ordinary path $F_{2n,2k}^{2r+,0}$, the corresponding expressions, got by putting $j=0$ in (3.1) are:

$$N(F_{2n,2k}^{2r+,0}) = \binom{n+k-1}{r-1}\binom{n-k-1}{r-1} - \binom{n+k-1}{r-2}\binom{n-k-1}{r} \qquad k>0 \tag{3.3}$$

$$N(F_{2n,0}^{2r+,0}) = \binom{n}{r-1}\binom{n}{r}\cdot\frac{1}{n} \tag{3.4}$$

THEOREM 1.

$$N(F_{2n,2k}^{2r+,2l,j} = \binom{n+k-1}{r-l-1}\binom{n-k-1}{r+l+j-1} - \binom{n+k-1}{r-l-2}\binom{n-k-1}{r+l+j} \tag{3.5}$$

$$1 \leqq l \leqq \frac{n-k-1}{2} \quad \textit{if} \quad n-k$$

is odd

$$1 \leqq l \leqq \frac{n-k-2}{2} \quad \textit{if} \quad n-k$$

is even

$$l+1 \leqq r \leqq n-k-j-l, \quad k>0$$

$$N(F_{2n,0}^{2r+,2l,j}) = \binom{n}{r-l-1}\binom{n}{r+l+j}\cdot\frac{2l+j+1}{n} \tag{3.6}$$

$$1 \leqq l \leqq \left[\frac{n-1}{2}\right], \quad l+1 \leqq r \leqq n-l-j$$

where $[x]$ *denotes the greatest integer contained in* x.

PROOF. Let $QP_1Q_1P_2Q_2 \ldots P_{2l}Q_{2l}R$ denote the $F_{2n,2k}^{2r+,2l,j}$ path where $Q_1, Q_2, \ldots, \ldots, Q_{2l}$ are the points of intersection of the path with the axis. $P_1, P_2, \ldots, P_{2l}$ are the last "turning points" of the path before intersecting the axis at $Q_1Q_2, \ldots, Q_{2l}$ where a "turning point" is a point where the path changes its direction from positive to negative or vica versa (see Fig. 1).

Now, make the following transformation to the path. To the end P_1, attach in order the segments $Q_1P_2, Q_2P_3, \ldots, Q_{2l}R$ with the signs of the θ's changed for all such segments lying below the axis. Then attach to this path in order the remaining segments $P_{2l}Q_{2l}, P_{2l-1}Q_{2l-1}, \ldots, P_1Q_1$ again such that the direction of all those segments lying below the axis is changed. Part of the transformed path between the origin and its last turning point is considered to be an ordinary path. In the

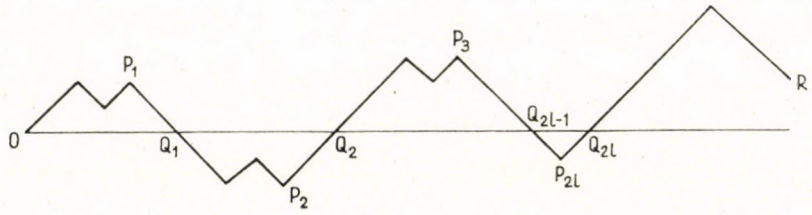


Fig. 1

remaining part the run endpoints are kept as they are, moreover each section attached forms a separate run. The resulting is an $F_{2n,2k}^{2r-2l+,\,0,\,j+2l}$ path (see Fig. 2).

Consider the last $2l$ negative runs of this composed path. Then draw parallel lines through their endpoints intersecting the path. By diminishing the ordinates of these points of intersection by $2k$, the points $P_1, P_2, \ldots, P_{2l}$ of the original path

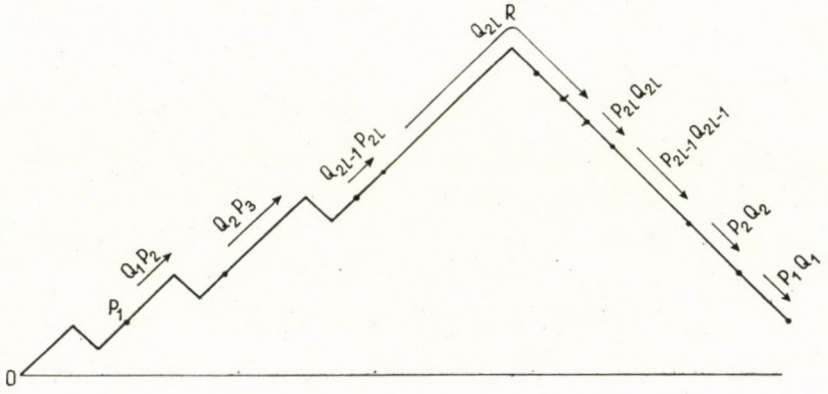


Fig. 2

can be traced back. Then the reversing of the procedure gives back the complete original path. Hence the transformation is 1:1.

The theorem now follows from lemma.

COROLLARY. For an ordinary path the corresponding expressions are got by putting $j=0$ in (3.5) and in (3.6).

The procedure applied in the proof of the above theorem will henceforth be referred to as "procedure 2".

THEOREM 2.

$$N(F_{2n,2k}^{(2r+1)+,2l}) = \binom{n+k-1}{r-l}\binom{n-k-1}{r+l-1} - \binom{n+k-1}{r-l-1}\binom{n-k-1}{r+l} \tag{3.7}$$

$$1 \leqq l \leqq \frac{n-k-1}{2} \quad \text{if} \quad (n-k)$$

is odd

$$1 \leqq l \leqq \frac{n-k}{2} \quad \text{if} \quad (n-k)$$

is even.

$$l \leqq r \leqq n-k-l, \quad k > 0$$

$$N(F_{2n,0}^{(2r+1)+,2l+1}) = \frac{2l+2}{n} \cdot \binom{n}{r-l-1}\binom{n}{r+l+1} \tag{3.8}$$

$$0 \leqq l \leqq \left[\frac{n-2}{2}\right], \quad l+1 \leqq r \leqq n-l-1$$

PROOF. As before applying procedure 2, we can establish a 1:1 correspondence between the $F_{2n,2k}^{(2r+1)+,2l}$ paths and the $F_{2n,2k}^{(2r-2l+2)+,0,2l-1}$ paths, $k>0$. The required result then follows from lemma.

Proceeding in exactly the same way for the case $k=0$, a 1:1 correspondence between the $F_{2n,0}^{(2r+1)+,2l+1}$ and $F_{2n,0}^{(2r-2l)+,0,2l+1}$ paths can be established. The required result then follows from lemma.

THEOREM 3.

$$N(F_{2n,2k}^{2r-,2l+1}) = \binom{n+k-1}{r-l-1}\binom{n-k-1}{r+l-1} - \binom{n+k-1}{r-l-2}\binom{n-k-1}{r+l} \tag{3.9}$$

$$0 \leqq l \leqq \frac{n-k-1}{2} \quad \text{if} \quad n-k$$

is odd

$$0 \leqq l \leqq \frac{n-k-2}{2} \quad \text{if} \quad n-k$$

is even

$$l+1 \leqq r \leqq n-k-l, \quad k > 0$$

$$N(F_{2n,0}^{2r-,2l}) = \binom{n}{r-l-1}\binom{n}{r+l} \cdot \frac{2l+1}{n} \tag{3.10}$$

$$0 \leqq l \leqq \left[\frac{n-1}{2}\right], \quad l+1 \leqq r \leqq n-l$$

PROOF. Repeating again procedure 2, an $F_{2n,2k}^{(2r-,2l+1}$ $(k>0)$ path can be reduced to an $F_{2n,2k}^{2r-2l+,0,2l}$ $(k>0)$ path and vice versa. The required result then follows from lemma.

Similarly, for the case $k=0$, an application of procedure 2 establishes a 1:1 correspondence between the $F_{2n,0}^{2r-,2l}$ paths and the $F_{2n,0}^{2r-2l+,0,2l}$ paths thus giving the required result using lemma.

THEOREM 4.

$$N(F_{2n,2k}^{(2r+j)-,2l+1,j)}) = \binom{n+k-1}{r-l-1}\binom{n-k-1}{r+l+j} - \binom{n-k-1}{r-l-2}\binom{n-k-1}{r+l+j+1} \tag{3.11}$$

$$0 \leqq l \leqq \frac{n-k-3}{2} \quad \textit{if} \quad n-k$$

is odd

$$0 \leqq l \leqq \frac{n-k-2}{2} \quad \textit{if} \quad n-k$$

is even

$$l+1 \leqq r \leqq n-k-l-j-1$$

$$N(F_{2n,0}^{(2r+1)-,2l+1,j)}) = \binom{n}{r-l-1}\binom{n}{r+l+j+1}\cdot\frac{2l+j+2}{n} \tag{3.12}$$

$$0 \leqq l \leqq \left[\frac{n-2}{2}\right], \quad l+1 \leqq r \leqq n-l-j-1$$

PROOF. Again applying procedure 2, an $F_{2n,2k}^{(2r+1)-,2l+1,j}$ $(k>0)$ path can be reduced to an $F_{2n,2k}^{2r-2l,0,2l+j+1}$ $(k>0)$ path and vica versa. The required result then follows from lemma.

The above holds for the case $k=0$ too, which on simplification reduces to (3.12).

COROLLARY. For an ordinary path the corresponding expressions are got by inserting $j=0$ in (3.11) and in (3.12).

The preceding give immediately the joint distributions of $(\lambda_{n+k,n-k}, R_{n+k,n-k})$ and $(\lambda_{n,n}, R_{n,n})$.

For the limiting distribution we have for $k\sim a\sqrt{2n}$

$$\lim_{n\to\infty} \mathsf{P}\left(\lambda_{n+k,n-k} < v\sqrt{2n},\ R_{n+k,n-k} < n+u\frac{\sqrt{n}}{\sqrt{2}}\right) = \tag{3.13}$$

$$= \frac{1}{\sqrt{2\pi}}(1-e^{-2(a+v)^2}) \int_{-\infty}^{u} e^{-u^2/2}\,du \qquad -\infty \leqq u \leqq \infty, \quad v \geqq 0$$

which shows that $\lambda_{n+k,n-k}$ and $R_{n+k,n-k}$ are independent in the limit.

The joint limiting distribution of $(\lambda_{n,n}, R_{n,n})$ follows by putting $a=0$ in (3.13).

For the number of paths with given property and with unspecified end point

we obtain the following results:

$$(3.14)\quad N(G_{2n}^{2r,2l}) = 2\sum_{k=1}^{n}\left\{\binom{n+k-1}{r-l-1}\binom{n-k-1}{r+l-1}-\binom{n+k-1}{r-l-2}\binom{n-k-1}{r+l}\right\}+$$
$$+2\frac{2l+1}{n}\binom{n}{r-l-1}\binom{n}{r+l}\qquad 1\leqq l\leqq\left[\frac{n-1}{2}\right]$$

$$(3.15)\quad N(G_{2n}^{2r+1,2l}) = 2\sum_{k=1}^{n}\left\{\binom{n+k-1}{r-l}\binom{n-k-1}{r+l-1}-\binom{n+k-1}{r-l-1}\binom{n-k-1}{r+l}\right\}$$
$$1\leqq l\leqq\left[\frac{n-1}{2}\right]$$

$$(3.16)\quad N(G_{2n}^{2r,2l+1}) = 2\sum_{k=1}^{n}\left\{\binom{n+k-1}{r-l-1}\binom{n-k-1}{r+l-1}-\binom{n+k-1}{r-l-2}\binom{n-k-1}{r+l}\right\}$$
$$0\leqq l\leqq\left[\frac{n-2}{2}\right]$$

$$(3.17)\quad N(G_{2n}^{2r+1,2l+1}) = 2\sum_{k=1}^{n}\left\{\binom{n+k-1}{r-l-1}\binom{n-k-1}{r+l}-\binom{n+k-1}{r-l-2}\binom{n-k-1}{r+l+1}\right\}+$$
$$+2\frac{2l+2}{n}\cdot\binom{n}{r-l-1}\binom{n}{r+l+1}\qquad 0\leqq l\leqq\left[\frac{n-2}{2}\right]$$

4. Joint distributions of $(\lambda_{n+k,n-k}^{(t)}, R_{n+k,n-k})$ and $(\lambda_{n,n}^{(t)}, R_{n,n})$

THEOREM 5.

$$(4.1)\quad N(F_{2n,2k,t}^{2r+,2l+1,j}) = \binom{n+k-1}{r-l-1}\binom{n-k-1}{r+l+j-1}-\binom{n+k-1}{r-l-2}\binom{n-k-1}{r+l+j}$$

$$0\leqq l\leqq\frac{n-k-1}{2}\quad \textit{if}\quad n-k$$

is odd

$$0\leqq l\leqq\frac{n-k-2}{2}\quad \textit{if}\quad n-k$$

is even

$$l+1\leqq r\leqq n-k-l,\quad 0<l<2k,\quad k>0$$

PROOF. Let $OO_1P_1Q_1P_2Q_2\ldots, P_{2l}Q_{2l}R$ be the $F_{2n,2k,t}^{2r+,2l+1,j}$ path with $O_1, Q_1, Q_2,\ldots$ $\ldots, Q_{2l}$ as its points of intersection with $y=t$ and with $P_1, P_2, \ldots P_{2l}$ as the last "turning points" of the path before intersecting $y=t$ in $Q_1, Q_2, \ldots, Q_{2l}$ respectively (see Fig. 3.).

Next reverse the segment OO_1, i.e. change the order of the θ's in it and then attach to it the segment O_1R after applying procedure 2 to the latter. Thus we get

an $F_{2n,2k}^{2r-2l+,0,2l+j}$ path (see Fig. 4.). As before the fact that the transformation is 1:1 can be verified easily.

The theorem now follows from lemma. The procedure used will henceforth be referred to as "procedure 3".

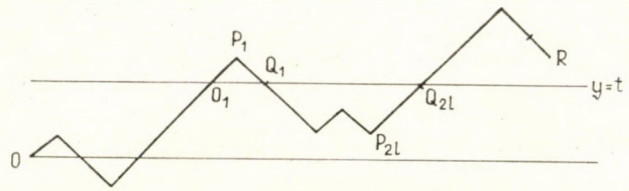


Fig. 3

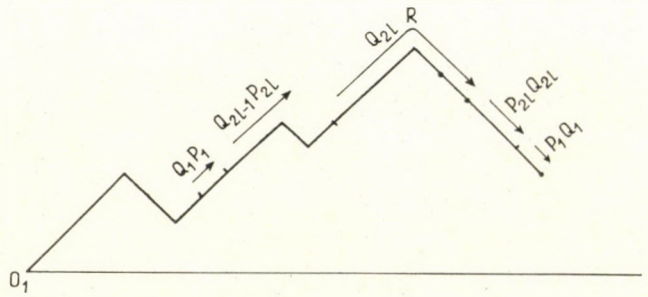


Fig. 4

COROLLARY. For an ordinary path $F_{2n,2k,t}^{2r+,2l+1}=F_{2n,2k,t}^{2r-,2l+1}$; the corresponding result is a special case of the above theorem, got by putting $j=0$ in it.

Hence

$$N(F_{2n,2k,t}^{2r+,2l+1}) = N(F_{2n,2k,t}^{2r-,2l+1}) = \tag{4.2}$$

$$= \binom{n+k-1}{r-l-1}\binom{n-k-1}{r+l-1} - \binom{n+k-1}{r-l-2}\binom{n-k-1}{r+l}$$

$$0 \leqq l \leqq \frac{n-k-1}{2} \quad \text{if} \quad (n-k)$$

is odd

$$0 \leqq l \leqq \frac{n-k-2}{2} \quad \text{if} \quad (n-k)$$

is even

$$0 < l < 2k, \quad l+1 \leqq r \leqq n-k-l, \quad k > 0$$

THEOREM 6.

$$(4.3)\qquad N(F_{2n,2k,t}^{(2r+1)+,2l+1}) = \binom{n+k-1}{r-l}\binom{n-k-1}{r+l-1} - \binom{n+k-1}{r-l-1}\binom{n-k-1}{r+l}$$

$$0 \leqq l \leqq \frac{n-k-1}{2} \quad \textit{if} \quad (n-k)$$

is odd

$$0 \leqq l \leqq \frac{n-k}{2} \qquad \textit{if} \quad (n-k)$$

is even.

$$l \leqq r \leqq n-k-l, \quad 0 < l < 2k, \quad k > 0$$

PROOF. Again applying procedure 3 it can be shown that there exists a 1:1 correspondence between the $F_{2n,2k,t}^{(2r+1)+,2l+1}$ paths and the $F_{2n,2k}^{2r-2l+2,0,2l-1}$ paths. The required result then follows from lemma.

THEOREM 7.

$$(4.4)\qquad N(F_{2n,2k,t}^{(2r+1)-,2l+1,j}) = \binom{n+k-1}{r-l-1}\binom{n-k-1}{r+l+j} - \binom{n+k-1}{r-l-2}\binom{n-k-1}{r+l+j+1}$$

$$0 \leqq l \leqq \frac{n-k-3}{2} \quad \textit{if} \quad n-k$$

is odd

$$0 \leqq l \leqq \frac{n-k-2}{2} \quad \textit{if} \quad n-k$$

is even

$$l+1 \leqq r \leqq n-k-l, \quad 0 < l < 2k$$

PROOF. Let $OPO_1P_1Q_1, \ldots, P_{2l}Q_{2l}R$ be an $F_{2n,2k,t}^{(2r+1)-,2l+1,j}$ path where the points $O_1, Q_1, Q_2, \ldots, Q_{2l}$ and the points $P_1, P_2, \ldots P_{2l}$ are as defined before and where P is the point where the first negative run ends (see Fig. 5).

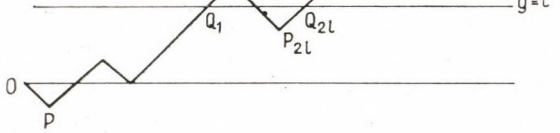


Fig. 5

We apply procedure 3 to the segment PR shifting its starting point P to the origin and then attach the segment OP to the end point of the transformed path. The last segment is considered as a separate run. Thus an $F_{2n,2k}^{2r-2l+,0,2l+j+1}$ path is obtained. That the transformation is 1:1 can easily be verified. The required result now follows from lemma.

COROLLARY. The corresponding result for an ordinary path is got by putting $j=0$ in (4.4).

The exact distribution of $\lambda^{(t)}_{n+k,n-k}$ and $R_{n+k,n-k}$ for $t<2k$ then follows from theorems 7 to 9. Also, it can be seen that the results in the above theorems are independent of t.

For the limiting distribution we obtain for $k\sim a\sqrt{2n}$

$$\lim_{n\to\infty} \mathsf{P}\left(v \leqq \frac{\lambda^{(t)}_{n,n}}{\sqrt{2n}} < v+dv,\ R_{n+k,n-k} < n+\frac{u\sqrt{n}}{\sqrt{2}}\right) = \tag{4.5}$$

$$= 4(v+a)e^{2a^2-2(a+v)^2}\,dv \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\,du \qquad a, v \geqq 0 \quad -\infty \leqq u \leqq \infty$$

THEOREM 8. *For* $t = 2k$

$$N(F^{2r+,2l+1}_{2n,2k,2k}) = \binom{n+k-1}{r-l-1}\binom{n-k-1}{r+l-1} - \binom{n+k-1}{r-l-2}\binom{n-k-1}{r+l} \tag{4.6}$$

$$0 \leqq l \leqq \frac{n-k-1}{2} \quad \textit{if} \quad n-k$$

is odd

$$0 \leqq l \leqq \frac{n-k-2}{2} \quad \textit{if} \quad n-k$$

is even

$$l+1 \leqq r \leqq n-k-l, \quad k > 0.$$

PROOF. The proof is trivial and follows from theorem 3 by showing a 1:1 correspondence between the paths $F^{2r+,2l+1}_{2n,2k,2k}$ and $F^{2r-,2l+1}_{2n,2k}$ by reversing the order of the θ's.

In a similar way it can be shown that

$$N(F^{2r-,2l}_{2n,2k,2k}) = \binom{n+k-1}{r-l-1}\binom{n-k-1}{r+l-1}\Big) - \binom{n+k-1}{r-l-2}\binom{n-k-1}{r+l} \tag{4.7}$$

$$1 \leqq l \leqq \frac{n-k-1}{2} \quad \text{if} \quad n-k$$

is odd

$$1 \leqq l \leqq \frac{n-k-2}{2} \quad \text{if} \quad n-k$$

is even

$$l+1 \leqq r \leqq n-k-l, \quad k > 0.$$

$$N(F^{(2r+1)+,2l}_{2n,2k,2k}) = \binom{n+k-1}{r-l}\binom{n-k-1}{r+l-1} - \binom{n+k-1}{r-l-1}\binom{n-k-1}{r+l} \tag{4.8}$$

$$1 \leqq l \leqq \frac{n-k-1}{2} \quad \text{if} \quad n-k$$

is odd

$$1 \leqq l \leqq \frac{n-k}{2} \quad \text{if} \quad n-k$$

is even

$$l \leqq r \leqq n-k-l, \quad k > 0.$$

$$N(F_{2n,2k,2k}^{(2r+1)-,2l+1}) = \binom{n+k-1}{r-l-1}\binom{n-k-1}{r+l} - \binom{n+k-1}{r-l-2}\binom{n-k-1}{r+l+1} \tag{4.9}$$

$$0 \leqq l \leqq \frac{n-k-3}{2} \quad \text{if} \quad n-k$$

is odd

$$0 \leqq l \leqq \frac{n-k-2}{2} \quad \text{if} \quad n-k$$

is even

$$l+1 \leqq r \leqq n-k-l-1.$$

The exact distribution of $(\lambda_{n+k,n-k}^{(2k)}, R_{n+k,n-k})$ follows from (4.6) to (4.9). For the limiting distribution we have for $k \sim a\sqrt{2n}$

$$\lim_{n\to\infty} \mathsf{P}\left(v \leqq \frac{\lambda_{n+k,n-k}^{(t)}}{\sqrt{2n}} < v+dv,\ R_{n+k,n-k} < n + \frac{u\sqrt{n}}{\sqrt{2}}\right) = \tag{4.10}$$

$$= 4(v+a)e^{2a^2-2(v-a)^2}\,dv \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\,du. \qquad a, v \geqq 0, \quad -\infty \leqq u \leqq \infty$$

THEOREM 9. *For* $t > 2k$

$$N(F_{2n,2k,t}^{2r+,2l}) = \left(\frac{n+t-k-1}{r-l}\right)\left(\frac{n-t+k-1}{r-l-2}\right) - \left(\frac{n+t-k-1}{r-l-1}\right)\left(\frac{n-t+k-1}{r+l-1}\right) \tag{4.11}$$

$$1 \leqq l \leqq \frac{n-t+k+1}{2} \quad \textit{if} \quad n-t+k$$

is odd

$$1 \leqq l \leqq \frac{n-t+k}{2} \quad \textit{if} \quad n-t+k$$

is even

$$l+1 \leqq r \leqq n-t+k-l+1, \quad t = 2k+1,\ 2k+2, \ldots, k \geqq 0$$

PROOF. Applying procedure 1 to the $F_{2n,2k,t}^{2r+,2l}$ path, a 1:1 correspondence between the $F_{2n,2k,t}^{2r+,2l}$ and the $F_{2n,2t-2k}^{2r-,2l-1}$ paths can be set up. The required result then follows from theorem 3.

THEOREM 10.

$$(4.12)\quad N(F_{2n,2k,t}^{2r-,2l}) = \binom{n+t-k-1}{r-l-1}\binom{n-t+k-1}{r+l-1} - \binom{n+t-k-1}{r-l-2}\binom{n-t+k-1}{r+l}$$

$$1 \leqq l \leqq \frac{n-t+k-1}{2} \quad \textit{if} \quad n-t+k$$

is odd

$$1 \leqq l \leqq \frac{n-t+k-1}{2} \quad \textit{if} \quad n-t+k$$

is even

$$l+1 \leqq r \leqq n-t+k-l, \quad t = 2k+1,\ 2k+2,\ \ldots,\ k \geqq 0.$$

PROOF. Let OPR be an $F_{2n,2k,t}^{2r-,2l}$ path with P as the point where the first negative run ends. Apply to the segment PR procedure 1, shifting its starting point P to

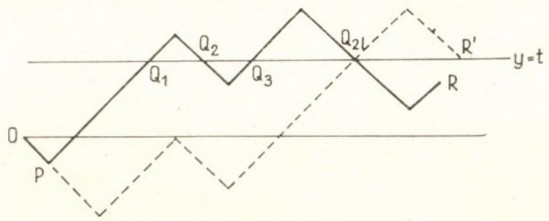


Fig. 6

the origin, and then to the end point of the transformed segment PR attach the segment OP. The last segment will be counted as a separate run. Thus an $F_{2n,2t-2k}^{(2r-1)-,2l-1,1}$ path is obtained (see Fig. 6.).

That the correspondence is 1:1 van be easily verified. The proof then follows from theorem 4.

THEOREM 11.

$$(4.13)\quad N(F_{2n,2k,t}^{(2r+1)+,2l}) = N(F_{2n,2k,t}^{(2r+1)-,2l}) =$$

$$= \binom{n+t-k-1}{r-l}\binom{n-t+k-1}{r+l-1} - \binom{n+t-k-1}{r-l-1}\binom{n-t+k-1}{r+l}$$

$$1 \leqq l \leqq \frac{n-t+k-1}{2} \quad \textit{if} \quad n-t+k$$

is odd

$$1 \leqq l \leqq \frac{n-t+k}{2} \quad \textit{if} \quad n-t+k$$

is even

$$l \leqq r \leqq n-t+k-l, \quad t = 2k+1,\ 2k+2,\ \ldots,\ k \geqq 0$$

PROOF. Proceeding in exactly the same way as in the preceding two theorems, a 1:1 correspondence can be established between the $F_{2n,2k,t}^{(2r+1)+,2l}$ and the $F_{2n,2t-2k}^{(2r+1)-,2l-1}$ paths and between the $F_{2n,2k,t}^{(2r+1)-,2l}$ and the $F_{2n,2t-2k}^{(2r+1)-,2l-1}$ paths. The required result then follows from the corollary to theorem 4.

THEOREM 12.

$$\binom{2n}{n} \mathsf{P}(\lambda_{n,n}^{(t)} = 2l,\ R_{n,n} = 2r) = \tag{4.14}$$

$$= \binom{n+t-1}{r-l}\binom{n-t-1}{r+l-2} - \binom{n+t-1}{r-l-2}\binom{n-t-1}{r+l}$$

$$1 \leqq l \leqq \left[\frac{n-t+1}{2}\right] \quad l+1 \leqq r \leqq n-t-l+1, \quad t > 0$$

and

$$\binom{2n}{n} \mathsf{P}(\lambda_{n,n}^{(t)} = 2l,\ R_{n,n} = 2r+1) = \tag{4.15}$$

$$= 2\left\{\binom{n+t-1}{r-l}\binom{n-t-1}{r+l-1} - \binom{n+t-1}{r-l-1}\binom{n-t-1}{r+l}\right\}$$

$$1 \leqq l \leqq \left[\frac{n-t}{2}\right] \quad l \leqq r \leqq n-t-l, \quad t > 0$$

PROOF. The theorem follows directly from (4.11) to (4.13) by considering the case $k=0$.

Similarly, the exact distribution of $(\lambda_{n+k,n-k}^{(t)}, R_{n+k,n+k})$ for $t>2k$ can be written down.

For the limiting distribution we obtain for $k \sim a\sqrt{2n}$ and $t \sim b\sqrt{2n}$:

$$\lim_{n\to\infty} \mathsf{P}\left(v \leqq \frac{\lambda_{n+k,n-k}^{(t)}}{\sqrt{2n}} < v+dv,\ R_{n+k,n-k} < n + \frac{u\sqrt{n}}{\sqrt{2}}\right) = \tag{4.16}$$

$$= 4(v+b-a)e^{2a^2-2(v+b-a)^2}\,dv \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\,du, \quad a, b \geqq 0, \quad v > 0, \quad -\infty < u < \infty.$$

REFERENCES

[1] CSÁKI, E.: On the number of intersections in the onedimensional random walk. *Publ. Math. Inst. Hungar. Acad. Sci.* **6** (1961), 281—286.

[2] CSÁKI, E. and VINCZE, I.: On some problems connected with the Galton-test. *Publ. Math. Inst. Hungar. Acad. Sci.* **6** (1961), 97—109.

[3] CSÁKI, E. and VINCZE, I.: Two joint distribution laws in the theory of order statistics. *Mathematica* (Cluj) **5** (1963), 27—37.

[4] BATCHELET, E.: Zur Theorie der wiederkehrenden Ereignisse, I—II. *Archiv der Math.* **8**, (1957) I: 184—194, II: 294—297.

[5] FELLER, W.: Fluctuation theory of recurrent events. *Trans. Amer. Math. Soc.* **67** (1949), 98—119.

[6] MICHALEVICH, V. S.: On the mutual disposition of two empirical distribution functions. *Sel. Transl. Math. Statist. Prob.* I, (1952) 63—68.

[7] KANWAR SEN: On some combinatorial relations concerning the symmetric random walk. *Publ. Math. Inst. Hungar. Acad. Sci.* **9** (1964), 335—356.

[8] REIMANN, J.—VINCZE, I.: On the comparison of two samples with slightly different sizes. *MTA Mat. Kut. Int. Közl.* **5** (1960), 293—309.

Hyderabad, Judie

(Received July 16, 1971)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 225—243.

DIRECT COMPARISON OF DIFFERENT PRIORITY QUEUING DISCIPLINES

by

A. SIMONOVITS

Summary

In the present paper we compare different priority queuing disciplines on the basis of their expected avarage waiting times or losses. Among the general *endogenous* priority disciplines the exogenous *discretionary* discipline plays a fundamental role. The simplest special cases of the discretionary discipline are the *nonpreemptive* and the *preemptive* disciplines.

A lot of papers investigated the optimization of the priority disciplines under different restrictions comparing different disciplines. Except the nonpreemptive discipline in all cases some assumptions are needed on the service time distributions, but for the sake of the simplicity we mention only the strongest one: *the service completion rate is a nondecreasing function of the elapsed service time for each unit.*

Under this and other conditions it looks like evident *the optimality ot the minimum principle:* At each decision point a unit of the minimal ratio of the expected remaining service time and the unit waiting cost is chosen. Authors assuming Poisson arrivals deduced it by complicated explicit formulae. Some authors excluded new arrivals or assumed non stochastic service times and proved the optimality of the minimum principle *directly*. Papers [1], [2], [4], [5], [5a], [7], [9], [11], [14], [17b], [18] belong to the first class, while [10b], [15], [16]. [17] belong to the second class.

I have noticed if the number of the priority classes equals to *two*, then the previous results hold without the mentioned restrictions. Furthermore, I proved the corrected version of BRONSTEIN—VEKLEROV's [6] conjecture: If the arrivals form a Poisson process and if the completion rates are nondecreasing, then *the optimal endogenous discipline is determined by the minimum principle, it is the optimal discretionary discipline* investigated in an other paper of the present author (17*b*), too.

1. Introduction

JAISWAL's monography "Priority Queues" [11] writes the following: "Situations commonly occur when a given service facility is made use of by $r(>1)$ classes of units which may be distinguished according to some »measure of importance«". Let be integer i, $(1 \leqq i \leqq r)$ the *priority index* of the i-th most important class, *a* unit of this class is denoted by L_i. "The discipline according to which the server selects the next unit and serves it is termed a priority discipline". Any *priority discipline* specifies the following two decisions:

1. Which unit to select for service, if the server is free.
2. Whether to continue or discontinue the service of the unit being serviced.

If Decision 1. depends only upon the knowledge of the priority class to which the unit belongs, the discipline is *exogenous*. The complementer type is the *endogenous* discipline.

In the case of exogenous priority L_i is always prior to L_j, $(i<j)$ at Decision 1. However, if L_j is in service and L_i arrives, there are the following disciplines (D):

D_I: *Non-preemptive:* The service of L_j is *continued.*

D_{II}: *Preemptive:* The service of L_j is immediately *interrupted* and the server starts to serve L_i.

D_{III}: *Discretionary:* The server uses his *discretion* to follow D_I or D_{II} depending upon the elapsed service-time of L_j.

To avoid any misunderstanding, let us remark that the discretionary discipline is exogenous, because Decision 1. depends only on the priority classes; but it is *dynamical*, because Decision 2. depends on the state of the system, too.

We shall use the following notations: τ_i denotes the random variable of the service-time of a unit in the i-th priority class, its distribution function is $B_i(t)$. The first moment, the second moment and the standard deviation of $B_i(t)$ is denoted by β_i, γ_i and σ_i; $\gamma_i=\beta_i^2+\sigma_i^2$, respectively.

In this paper the *waiting-time* of a unit equals to the end of its service minus the time of its arrival. We assume that the waiting-loss of a unit L_i is the waiting-time multiplied by the constant c_i, $c_i>0$.

We shall see that there is an interesting connection between the "necessary" assumptions on the character of the arrival process and the number of the priority classes. Therefore we need three alternative assumptions.

A(i): Arbitrary arrival process, two priority classes. (As our results for two-class priority disciplines are independent of the actual realization of the stochastic arrival process, we shall fix an arbitrary realization of it.)

A(ii): The number of priority classes is arbitrary, the arrival dates are arbitrary, but the probability of the event that a unit of a given date belongs to C_i, is a_i/a, independently of other arrival events. Clearly, a is the total arrival intensity, the existence of which is assumed.

A (iii): The specification of A (ii) — the arrival dates form a Poisson stream.

In this paper we always assume that the arrival process and the service process are stochastically *independent*. Furthermore, we assume that the service times of arbitrary finite collection of the units are *independent*, too, and the preempted unit *resumes* service from the point where it was interrupted.

It may happen that two different priority disciplines are actually identical for certain realizations of the arrival process, because these realizations exclude any actual difference between the disciplines. We shall neglect this case as non interesting.

The application of the mentioned priority disciplines does not change either the busy period or the utilization factor of the system, it influences the avarage waiting loss and its expected value. In the case of limited number of arriving units these quantities exist, but for infinite source they may not exist. However, in this case we shall assume that there exists the *asymptotical arrival rate* a_i for each i, $1\leqq i\leqq r$. Moreover, we shall assume that the queue is *not saturated*, i.e. its utilization factor is smaller than one: $\sum_{i=1}^{r} a_i\beta_i<1$. (If the system is saturated, then the avarage expected waiting loss equals to infinity.)

In the following we shall not mention these assumptions!

If L_i is prior to L_j, then notation $L_i \succ L_j$ is used. If discipline D is better than discipline D', i.e. the avarage expected waiting time of D is smaller than that of D', then notation $D \succ D'$ is used. If equality holds, then notation $D \sim D'$ is used.

If infinitely many arrivals take place, then an optimal priority discipline may contain non optimal finite parts. But if each finite part is optimal, then the whole discipline is optimal. Between two optimal disciplines the first one is said to be optimal, if the second one contains all the preemptions of the first discipline.

2. Nonpreemptive priority discipline

We compare the nonpreemptive priority discipline with the natural (first come first served) discipline and hence we shall deduce the optimal priority ranking in special cases.

A (i): First we fix an arbitrary arrival process (realization) and assume that there are only two priority classes. As any preemption is excluded, the server may select, if there are both ordinary and priority units in the queue, when a service is completed. (As the arrival process is deterministic, it may occur that the server waits a priority unit, while ordinary units wait for him. But we exclude this possibility a priori!) Let us fix an arbitrary ordinary unit L_2, which is served with positive probability after a priority unit L_1, which arrived later. Shortly: *A priority change* $L_1 \succ L_2$ takes place with positive probability. Under the service of L_1 L_2 must wait time τ_1, its loss is $c_2\tau_1$. We assume that every *overtaking* causes loss d (with no time loss). If we prevent this overtaking, then L_1 must wait for L_2 time τ_2, its loss is $c_1\tau_2$. We say that an *overtaking* $L_1 \succ L_2$ *is advantageous in itself*, if the total waiting loss decreases by this overtaking, i.e.: $c_1\tau_2 + d < c_2\tau_1$. We shall say, that *an overtaking is expectedly advantageous* (=*good*) *in itself*, if the total expected waiting-loss decreses, i.e.: $c_1\beta_2 + d < c_1\beta_2$.

We shall prove *if an overtaking is advantageous or good in itself, then it is advantageous or good in global.* This statement is the fundament of the whole paper. If we exclude other overtakings after this overtaking for a moment, then this overtaking does not influence the waiting-times (losses) of all units behind L_1 and L_2. If a unit arrives time t after the selection, then it waits time $(\tau_1 + \tau_2 - t)^+$ for L_1 and L_2 together, independently of the service order of L_1 and L_2.

Now we drop the previous assumption according to which after this overtaking no other overtaking takes place. Let us remark, that the possibility of a given change $L_1' \succ L_2'$ depends on the realization of the arrival process (which is fixed here) and on the realization of service-times of those units, which are served before both L_1' and L_2'. Clearly, if a change is possible when our change $L_1 \succ L_2$ does not take place, then it remains possible, if $L_1 \succ L_2$ takes place.

Thus if one change is good in itself, then all other changes are good in themselves; one change increases or at least does not decrease the possibility of the subsequent changes, hence we proved

THEOREM 1 (i). *The nonpreemptive priority discipline is better than the natural discipline (for an arbitrary arrival process and two-class disciplines) if and only if* $c_1\beta_2 + d < c_2\beta_1$.

Remarks: Let us introduce the following notation: If L_1 is prior to L_2, then a discipline is denoted by $D_{()}$, if L_2 is prior to L_1, then it is denoted by $\bar{D}_{()}$. Thus $D_1 \succ D_0$ implies $\beta_1/c_1 < \beta_2/c_2$, and if $d=0$, then the converse is also true. If $D_1 \succ D_0$, then $D_0 \succ \bar{D}_1$; but it may happen, that $D_1 \prec D_0$ and $\bar{D}_1 \prec D_0$ hold together. For the sake of the simplicity, we assume $\beta_1/c_1 < \beta_2/c_2$, when $D_1 \succ D_0 \succ \bar{D}_1$, if the change is costless: $d=0$.

We have mentioned in the Summary and in the Introduction, that our method becomes weaker, if there are more than two priority classes. Now we present an example to illustrate the appearing problems:

Let constant service-times and equal waiting costs be given, $\beta_1 < \beta_2 < \beta_3$. Units L_2 and L_3 may be selected at a given time, L_1 arrives after time t, where $\beta_2 < t < \beta_3$. If we select L_2 against L_3, then we prevent a later selection $L_1 \succ L_3$, while the selection L_3 against L_2 yields the selection $L_1 \succ L_2$. In the first case the total waiting loss is $\beta_2 + (\beta_2 + \beta_3 - t)$, while in the second case the total loss is $(\beta_3 + \beta_1) + (\beta_3 - t)$. Clearly, if $\beta_1 + \beta_3 < 2\beta_2$, then the latter is smaller than the former. A selection, which is good in itself, excludes a better selection, thus it is bad in global.

Somebody may think that usage of *endogenous nonpreemptive disciplines* can improve the possibilities. What happens, if every selection depends on the state of the system, i.e. on the number of different classes, their largeness at the moment of the selection, etc.? — We have seen that each selection $L_2 \succ L_1$ is bad in itself ($d=0$!) and does not increase the possibility of subsequent good selection of type $L_1 \succ L_2$. Thus the optimal endogenous discipline is the optimal exogenous discipline!

It is easy to see if all the units are present before the service begins, then Theorem 1 (i) is true for arbitrary number of priority classes, which was proved by ROTHKOPF [15] in 1966.

A (ii): We have shown that Theorem 1 (i) is not true, if $r>2$, even if $c_i=1$, $1 \leqq i \leqq r$. Therefore we use A (ii), i.e. *the type of each arrival unit is random.* Furthermore, we assume that all unit waiting costs equal to 1.

It is well-known and evident that the waiting-time of a unit of C_j is independent of the priority ranking of the higher priority classes C_i, $1 \leqq i < j$. That is the reason for the basic importance of two-class priority disciplines. Let us define the united class of priority classes of higher priority than j: $C^{[j]} = \bigcup_{i=1}^{j-1} C_i$. According to our assumption A (ii), the following equations hold:

$$B^{[j]}(t) = \sum_{i=1}^{j-1} a_i B_i(i) / \sum_{i=1}^{j-1} a_i, \quad \beta^{[j]} = \sum_{i=1}^{j-1} a_i \beta_i / \sum_{i=1}^{j-1} a_i.$$

The priority ranking $C^{[j]} \succ C_j$ is good if and only if $\beta^{[j]} < \beta_j$, by Theorem 1 (i), because the presence of lower priority classes than j does not influence the optimal relation of $C^{[j]}$ and C_j. So we have proved

THEOREM 1 (ii). *If $L_{j-1} \succ L_j$ and $\beta^{[j]} < \beta_j$ $2 \leqq j \leqq r$, then the non-preemptive discipline is better than the natural discipline.*

Remark. Unfortunately, Theorem 1 (ii) does not tell which priority ranking is the optimal one. The problem is the following: While the priority ranking of higher priorities is independent of the lower priorities, the ranking of the lower priorities depends on the existence of higher priorities (see our counter example).

There is a *conjecture*, that *under* A (ii) *the ranking of lower priorities does not depend on the existence of higher priority classes*, but we are not to prove it. It would imply that the priority ranking $L_{j-1} \succ L_j$ is good in itself if and only if $\beta_{j-1}/c_j < \beta_j/c_j$.

Let us remark there is a unique priority ranking which is good and independent of the arrival intensities, the ranking of our conjecture.

A (iii): Finally we remark the following well-known result:

THEOREM 1 (iii). *If the arrival process is a Poisson stream, then the priority ranking $L_j \succ L_{j+1}$ is good in itself (i.e. the classes of higher priority than j and the classes of lower priority than $j+1$ are fixed) if and only if $\beta_j/c_j < \beta_{j+1}/c_{j+1}$. Therefore the unique optimal ranking is the order of weights β_j/c_j's. The optimal nonpreemptive discipline is better than the natural discipline.*

Remarks: Theorem 1 (iii) was found and proved by M. ACZÉL [1] in 1960 for the special case of equal cost of waiting. COX and SMITH proved the first and the second statements of Theorem 1 (iii) in paper [9] in 1961. The third statement of Theorem 1 (iii) was proved by BRONSTEIN and RYKOV [5] in 1965. The method used was rather complicated. See also [11] pp. 152—156.

In 1967 RYKOV and LEMBERG [14] defined the notion of endogenous nonpreemptive priority disciplines. They proved our recent result only for Poisson arrival, but for arbitrary number of priority classes, that the optimal endogenous discipline is the optimal exogenous discipline of Theorem 1 (iii).

Let us complete this Part with the following remark of FIFE [10]: The assumption of Poisson arrival excludes the idleness of the server in the optimal discipline, when some units are waiting.

3. Preemptive priority discipline

Now we analyse the question: When is the preemptive priority better than the nonpreemptive priority? According to the general method of the Queuing Theory

1) if we deal with a *general* arrival process, then we assume that the distributions of service-times are *exponential.*

2) if we deal with *general* service-time distributions, then we assume that the arrival process consists of r independent *Poisson* streams with intensities a_i, $1 \leqq i \leqq r$.

We assume that each preemption causes loss e (with no time-loss), which is not smaller than loss d due to a nonpreemptive priority-change: $e \geqq d$.

1) A (i): If $B_2(t) = 1 - e^{-b_2 t}$, then the distribution of the *remaining service time* τ_i^* is also $B_2(t)$. (Therefore we can study the *preemptive repeat-different* discipline, where the preempted unit on its reentry requires a random service-time independent of past preemptions and vasted service times. Now this disciplines is reduced to the *preemptive resume discipline*, where the preempted unit resumes the service from the point where it was interrupted.)

We analyse the influence of an arbitrary preemption $L_1 \succ L_2$: Without the change L_1 must wait τ_2 for L_2, its loss is $c_1\beta_2$, ($\beta_2 = 1/b_2$). With the change L_2 must wait τ_1 for L_1, its loss is $c_2\beta_1 + e$. As we have seen in the previous part, a preemption is good if and only if $c_1\beta_2 \succ c_2\beta_1 + e$. As $e \geqq d$, $c_1\beta_2 \succ c_2\beta_1 + d$ also holds, thus an *induced* priority-change $L_1 \succ L_2$ in the queue is also good, if it is induced at all. Hence if $c_1\beta_2 \succ c_2\beta_1 + e$, then $D_{II} \succ D_{II} \succ D_0 \succ \bar{D}_I$.

If $c_2\beta_1+d<c_1\beta_2 \leqq c_2\beta_1+e$ holds, then the preemption is not good in itself, but the induced changes are good. Which factor is the stronger one? This depends on the expected number of the induced priority-changes.

We have proved

THEOREM 2 (i). *If* $B_2(t) = 1-e^{-b_2 t}$ *and* $d \leqq e$, *then* $c_2\beta_1+e<c_1\beta_2$ *implies* $D_{II} \succ D_I \succ D_0 \succ \bar{D}_1$.

In the remaining part of this paper we assume that all changes are *costless:* $d=e=0$; and $B_i(t) = 1-e^{-b_i t}$. Now $D_{II} \succ D_I$ is equivalent to $D_1 \succ D_0$. As in the previous Section, we may prove that *the optimal endogenous preemptive discipline is identical with the exogenous one.*

A (ii): Clearly, it is parallel to the previous case: $(c_i=1)$

THEOREM 2 (ii): *If* $B_j(t) = 1-e^{-b_j t}$, $\beta^{[j]}<\beta_j$ *and* $L_{j-1} \succ \bar{L}_j$, $2 \leqq j \leqq r$, *then the preemptive priority discipline is better than the nonpreemptive one.*

A (iii): Finally we mention the following known result:

THEOREM 2 (iii). *If* $B_i(t) = 1-e^{-b_i t}$, $1 \leqq i \leqq r$ *and the arrivals form Poisson stream, then the preemptive-change* $L_i \succ L_{i+1}$ *is good iff the priority-change in the queue* $L_i \succ L_{i+1}$ *is good, i.e. iff* $\beta_i/c_i<\beta_{i+1}/c_{i+1}$. *The optimal priority ordering is the same as in Theorem* 1 (iii) *and the optimal preemptive priority is better then the nonpreemptive priority.* ("Iff" stands for "if and only if".)

The problem was analysed by BROSH and NAOR [7] in 1963, see also [11] pp. 119—122. The authors dealt with the question of the optimal priority ranking and only Bronstein and RYKOV (1965) [5] proved that the preemptive priority is better than the nonpreemptive one.

In 1967 VEKLEROV [18] investigated the problem of *endogenous preemptive priority.* He proved that *the optimal endogenous preemptive discipline is identical with the optimal exogenous discipline analysed in Theorem* 2 (iii). (To be precise let us remark that VEKLEROV proved this statement for generalized Poisson arrivals; the arrival-times form a Poisson-process, but at a given arrival-point k units L_i arrive with probability $p_k^{(i)}$, $\sum_{k=1}^{\infty} p_k^{(i)} = 1$, $1 \leqq i \leqq r$.)

2) If we deal with general service-time distributions, then we "must" assume that the arrival process is a Poisson stream. Now we cannot fix any realization of the arrival process, thus our method becomes problematical. For the sake of the completness we cite some known results.

We need the *conditional expected value of the remaining service-time* of an ordinary unit the service of which has not yet completed; denote it β^*. To avoid certain misunderstanding we emphasise that this quantity is defined independently of the number of the preemptions which took place earlier. In 1963 AVI—ITZHAK and NAOR [3] proved the following result: $\beta^*=\gamma/2\beta$. It is very important that β^* depends on the first and the second moments of $B(t)$, but it is independent of other characteristics of $B(t)$ as well as of the arrival-intensities.

From now on we assume that the changes are costless. As SCHRAGE and MILLER [16] remarked in 1966, one can suggest

THEOREM 3 (i) *If the arrival-process of the priority units is a Poisson stream, then the preemptive priority is better than the nonpreemptive priority if and only if* $\beta_2^*/c_2 \succ \beta_1/c_1$.

Remark: SCHRAGE and MILLER [16] assumed that the arrivals of ordinary units form a Poisson stream, too. Of course, we assumed that the arrivals of the priority units and the ordinary units are *independent*.

If we fix an ordinary unit and the stochastically varying priority units which disciplines, then Theorem 3 (i) becomes evident.

In the case of the costless change we can assume $D_I \succ D_0 \succ \bar{D}_I$. However, it is possible that $D_{II} \succ D_I$ and $\bar{D}_{II} \succ \bar{D}_I$ hold together or $D_I \succ D_{II}$ and $\bar{D}_I \succ \bar{D}_{II}$ hold together. In the case of equal waiting-loss per time these relations are equivalent to $\beta_2^* > \beta_1$ and $\beta_1^* > \beta_2$ or $\beta_2^* < \beta_1$ and $\beta_1^* < \beta_2$. Clearly, $\beta^* < \beta$ is equivalent to $\beta > \sigma$. If $\sigma < \beta$, then $\beta_1/c_1 < \beta_2/c_2$ implies $\beta_1^*/c_1 < \beta_2/c_2$, hence $D_I \succ D_0 \succ \bar{D}_I \succ \bar{D}_{II}$. As $\beta^* \geqq \beta/2$, $\beta_1 < \beta_2/2$ implies $\beta_2^* > \beta_1$. Hence, if $\beta_1/c_1 < \beta_2/2c_2$, then $D_{II} \succ D_I \succ D_0$. In the case of constant service time of L_2 the converse is also true.

We must admit that our direct method cannot tell the condition of $D_{II} \succ D_0$, though it is an easy consequence of the explicit formulae.

A (iii): If there are r different priority classes, then BRONSTEIN [4] defined *the mixed preemptive — nonpreemptive discipline* as follows:

Let $i(j)$ be an integer between 1 and j for each j, $1 \leqq j \leqq r$. An arriving unit L_i is prior to L_j which is before the start of its service if and only if $i < j$. An arriving unit L_i is prior to L_j^* which is after the start of its service if and only if $i < i(j)$.

Evidently, if $i(j) \equiv 1$, then we get the nonpreemptive discipline; if $i(j) \equiv j$, then we get the preemptive discipline.

BRONSTEIN [4] proved the following generalization of Theorem 3 (i) in 1969:

THEOREM 3 (iii): *If there are r independent Poisson arrival processes and if $\sigma_i < \beta_i$, $1 \leqq i \leqq r$, then the optimal mixed preemptive — nonpreemptive discipline chooses the unit of the minimal β_i/c_i, or β_i^*/c_i, corresponding to the state of the unit whether it is before or after the start of its service.*

If $\beta_i/c_i < \beta_{i+1}/c_{i+1}$, then $i(j) = \min\{i; \beta_j^*/c_j \leqq \beta_i/c_i\}$.

Remark: If $\beta_j^*/c_j > \beta_i/c_i$ holds for each i and j, $1 \leqq i < j < r$, then L_i^* is prior to L_j as well as to L_j^* if and only if $i < j$; the preemptive discipline is the optimal mixed discipline.

4. Discretionary priority discipline

A (i): *First we fix one realization of the arrival process.*

It is easy to imagine that ... "in preemptive discipline, we may preempt an ordinary unit whose service is almost completed, or, in the nonpreemptive discipline, we may allow a priority unit to wait even if the ordinary unit has just entered service. Obviously, such possibilities can be avoided by allowing the server to use his discretion to continue or discontinue the service of an ordinary unit." [11] p. 161.

To be precise let us denote u_L the amount of service which unit has recieved until a given moment. A nonnegative number $u^0 \in [0, \infty]$ is called *the discretionary point*, if the service of an ordinary unit is continued at the arrival of a priority unit if and only if $u_{L_2} \geqq u^0$. According to our notations: $u_{L_2} < u^0$ implies $L_1(0) \succ L_2(u_2)$ and $u_{L_2} \geqq u^0$ implies $L_2(u_2) \succ L_1(0)$.

Denote by $\beta_L(u)$ the *conditional expected remaining service-time* of L — under

the condition that its elapsed service time is u and its service has not yet finished. Clearly, $\beta(u) = \int_{u=0}^{\infty} (t-u)dB(t)/\bar{B}(u+0)$, where $\bar{B}(u) = 1-B(u)$ and $\bar{B}(u+0)>0$. (This quantity differs from β^* defined in Part 3!)

In this part we look for the optimal discretionary point. According to our method a preemption $L_1(0) \succ L_2(u_2)$ is good in itself if and only if $c_2\beta_1 < c_1\beta_2(u_2)$. Certain complications arises if there is a pair (u', u''), where $u' < u''$ and $\beta_2(u') < \beta_2(u'')$, because a preemption may be bad at u' and it may be good later, at u''. To avoid this complication we assume the existence of a point $u^0 \in [0, \infty]$, for which $u < u^0$ implies $c_2\beta_1 \leqq c_1\beta_2(u)$ and $u \geqq u^0$ implies $c_2\beta_1 \geqq c_1\beta_2(u)$. As $\beta(u-0) \leqq \beta(u)$ holds, $\beta(u^0-0)=\beta(u^0)$ also holds, i.e. the service-time distribution of the ordinary unit must be continuous at the discretionary point. If $c_2\beta_1 < c_1\beta_2(u)$ for each $u \geqq 0$, then $u^0=\infty$, i.e. the optimal discretionary discipline is the preemptive discipline. As $\beta(0) \leqq \beta(0+0)$ and $c_2\beta_1 < c_1\beta_2$, the optimal discretionary discipline is never identical with the nonpreemptive discipline: $u^0>0$.

If $u^0 \in (0, \infty)$, then the smaller discretionary point is the better if both points are smaller than the optimum and the greater one is the better if both points are greater than the optimum.

Thus we have proved

THEOREM 4 (i): *If there exists a point $u^0 \in (0, \infty]$, such, that $u<u^0$ implies $c_2\beta_1 \leqq c_1\beta_2(u)$ and $u>u^0$ implies $c_2\beta_1 \geqq c_1\beta_2(u)$; then u^0 is an optimal discretionary point.*

Examples:

— The most trivial case is when the service-times are constant: β_i. (Clearly, it is enough if the ordinary service-time is constant.) Then $\beta_2(u) = (\beta_2-u)^+$, hence $u^0 = \beta_2 - \beta_1 c_2/c_1$, where $c_2\beta_1 < c_1\beta_2$.

— If $B_2(t) = 1-e^{-b_2 t}$, then $\beta_2(u)=\beta_2$, hence $u^0=\infty$. (See also 2) of Part 3.)

The first example was studied by AVI-ITZHAK, BROSH and NAOR [2] in 1964, other examples were studied by ETSCHMAIER in 1966, referred by JAISWAL [11] pp. 169—173. To prove Theorem 4(i) it is not necessary to assume that the arrivals form a Poisson stream and ther method is also rather complicated.

JAISWAL failed to discuss the global optimality of the discretionary point, he analysed only the local optimality. His optimum condition is the following: $p_2(u^0) < 1/\beta_2(u^0) = c_1/c_2\beta_1$, where $p(u)=b(u)/\bar{B}(u)$, $b(u)=B'(u)$.

Of course, if there is only one local optimum, including 0 and ∞ then it is the global optimum, too. If the inequality of JAISWAL's condition is valid for each u, then $\beta_2(u)$ is a strictly decreasing function, and vice versa; then all right.

JAISWAL himself emphases if $\beta_2(u)$ is decreasing, then the priority unit preempts the ordinary unit if and only if $\beta_2(u_2)/c_2 > \beta_1/c_1$; but he does not even explain it instead of deducing the optimum from this simple inequality.

If the number of the priority classes is greater than two, then there exist two variants of the discretionary discipline. We shall use the first variant, if A (ii) holds, and the second variant, if A (iii) holds.

We shall see later it is "necessary" to assume that *the expected remaining service-time does not increase, when the elapsed service-time increases.*

A (ii): JAISWAL [11] p. 174. defined the general discretionary discipline as follows: L_i preempts $L_j(u_j)$ if and only if $i<j$ and $u_j<u_j^0$. In this case we shall prove

THEOREM 4 (ii): *If $\beta_i(u_i)$ is a nonincreasing function of u_i, $1\leqq i\leqq r$; $c_i\equiv 1$ and one of the good priority rankings of Theorem* 1 (ii) *is applied, then the optimal discretionary point of L_j exists and defined by*

$$u_j^0 = \inf\{u;\ \beta_j(u) \leqq \beta^{[j]}\} \qquad 2 \leqq j \leqq r.$$

PROOF. This proof is similar to the proofs of Theorems 1 (ii) and 2 (ii). If the server uses his discreation to continue or discontinue the service of $L_j(u_j)$ when $L^{[j]}$ has just arrived, then the existence of lower priority classes $(C_{j+1}, \ldots, C_r)$ is indifferent, because each unit of $\bigcup_{k=j+1}^{r} C_k$ must wait both L_j and $L^{[j]}$. Thus we can use Theorem 4 (i), according to which the optimal discretionary point between L_j and $L^{[j]}$ is defined as

$$u_j^0 = \inf\{u;\ \beta_j(u) = \beta^{[j]}\}.$$

Let us remark that the optimal discretionary points depend on $\beta_i(u_i)/c_i$'s as well as on the arrival intensities a_i. There is only a unique priority ranking, which is independent of the arrival intensities: the order of β_i/c_i.

It is hoped to be the optimal priority ranking in the case of the discretionary discipline.

A (iii): BRONSTEIN and VEKLEROV defined the general discretionary disciplines by the interruption rule: The arriving L_i preempts $L_j(u_j)$ if and only if $i<j$ and $u_j<u_{ij}^0$. Moreover, we assume $u_{ij}^0\leqq u_{i-1,j}^0$ and $u_{ji}^0=0$ holds for each i, j; $2\leqq i\leqq j\leqq r$. How can we determine the selection rule, where the server becomes free?

If $L_i\succ L_j(u_j)$, then $i<j$, consequently between two subsequent completions at most $r-1$ preemptions may take place. Therefore the last preemption exists and this preemptive unit is not interrupted during its total service. We shall show, that after the removal of this unit we can replace units arriving during this period to the start of the next service — with no influence on the priority order.

To prove this, suppose L_h was the last unit, and it preempted $L_j(u_j)$. During the servise of L_h the arriving L_i and L_k did not preempt L_h. Let us forget about L_h and assume L_i and L_k arrived at u_j. Now the selection rule is reduced to the preemption rule. We shall show that the arrival order of L_i and L_k does not influence the final priority order.

If $i=k$, then it is trivial. Thus we may assume that $i<k$. If L_i arrives earlier then L_k, then $L_i\prec L_j(u_j)$ if $u_j\geqq u_{ij}^0$ and $L_i\succ L_j(u_j)$ if $u_j<u_{ij}^0$. In the first case $u_j\geqq u_{ij}^0\geqq$ $\geqq u_{kj}^0$ i.e. $L_k\prec L_j(u_j)$, while in the second case $L_i\succ L_k$. The presence of L_k is indifferent.

If L_k arrives earlier than L_i, then $L_k\prec L_j(u_j)$ or $L_k\succ L_j(u_j)$. In the first case $L_i\prec L_j(u_j)$ really independent of L_k, while in the second case $u_j<u_{kj}^0\leqq u_{ij}^0$ and $L_i\succ L_j(u_j)$. (If u_{ij}^0 is independent of i, $i<j$, then we get JAISWAL's definition as a special case.)

In an other paper the present author stated and proved the following

THEOREM 4 (iii): *If the arrivals form a Poisson stream, and if $\beta_j(u_j)$ is nonincreasing for $j=1, \ldots, r$, then there exists the optimal discretionary discipline, where*

$$u_{ij}^0=\inf\{u_j;\ \beta_j(u_j)/c_j\leqq\beta_i/c_i\}.$$

Remarks: This theorem is the generalization of the theorems of BRONSTEIN and VEKLEROV [5a] and [6], which concerned non random service times. We publish the proof in a separate paper [17b], because its method differs from that of this paper.

Though Theorem 4 (iii) is a simple special case of Theorem 5 (iii) of the next part, the method of (17b) is interesting in itself, because the expected value of the waiting time of each priority class is determined, generalizing the corresponding results of JAISWAL [11] and BRONSTEIN—VEKLEROV [5a].

Theorem 3(iii), which is due to BRONSTEIN [4], is the "special case" of Theorem 4 (iii), where $u_{ij}^0=\infty$ if $i>i(j)$ and $u_{ij}^0=0$ if $i\geqq i(j)$, $1\leqq i(j)\leqq j\leqq r$ and $\{i(j)\}_{j=1}^r$ is optimized.

5. Endogenous priority discipline

Though we have mentioned the problem of the endogenous disciplines at the end of Part 2 and Part 3 in the simplest cases, up to now we concentrated on exogenous disciplines. Now we concentrate on the endogenous disciplines and we prove new results in this field, generalizing some ideas of SCHRAGE and MILLER [16], [17], BRONSTEIN and VEKLEROV [6], and ROTHKOPF [15].

We need the following notations: Denote $n(t)$ the number of units which are present at moment t. The arrival-index of L is denoted by $m_L(t)$, i.e. $m_L(t)-1$ units have arrived earlier than L among those units which are present at t. The *original priority-index* (the type) of $L_{m(t)}$ is denoted by $i_{m(t)}$, the amount of the service which it has received until t is $u_{m(t)}$. The *state-vector* of the queue at t is $v_t=(u_{m(t)}, i_{m(t)})_{m(t)=1}^{n(t)}$.

From the practical point of view we may define a *general endogenous priority discipline* as follows: Each state-vector determines the *temporary priority-index* m_v, a time period of length Δ_v, and L_{m_v} is served during this period unless its service is finished or a new unit arrives. If its service is completed or a new unit arrives, then the new state-vector determines the following step, etc.

In 1966 SCHRAGE and MILLER [17] defined a similar, but rather special discipline for the case of *non random* service-times. (We remark, it was assumed that all waiting costs were the same.) They introduced the following discipline: *At every moment the highest priority is given to a unit of the shortest remaining service-time.* We shall call this discipline the *discipline of the minimum-principle.* (Let us remark, that the priority-index changes only at arrival points and if the newly arriwing unit has the minimal remaining service-time: $\Delta_v=\infty$.)

The authors emphasised that the minimum-principle generated the optimal endogenous priority discipline for every arrival process! As we have mentioned earlier, this perception is exceptional in the literature of the priority queues! This paper follows the method of Schrage and Miller.

In 1969 BRONSTEIN and VEKLEROV [6] wrote the following: There is no doubt about the optimality of the minimum-principle for the general case of stochastic service-times, but there is no explicit formula for the avarage expected loss, how can we prove optimality.

For the sake of the correctness we must emphasize that BRONSTEIN and VEKLEROV also wrote that the application of these endogenous disciplines might be very expensive, therefore we had to minimize the *netto* loss instead of the brutto loss.

Thus we had to know the brutto loss and the cost of each discipline and we had to minimize their difference.

But there is a serious problem: BRONSTEIN and VEKLEROV have failed to define the general endogenous disciplines and the optimum-principle for random service-times. It seems to us there is only one solution: to substitute the remaining service-loss $(\tau-u)/c$ by its expected values $\beta(u)/c$. Now we have a general principle which we shall also call minimum-principle.

1) First, let us show by a counter example that *the minimum principle does not generate an optimal discipline in general.* There are two units: L_1 and L_2. (These indices are the original priority indices now.) Suppose that the server is free and $u_1=u_2=0$. The service-time of L_1 is a constant $\beta_1>0$, while the random service-time of L_2 is 0 and $2\beta_2$ with probability $^1/_2$-$^1/_2$, resp. If $\beta_1<\beta_2$, then according to the minimum-principle we must begin with the service of L_1. As $\beta_1(u)=(\beta_1-u)^+$ decreases, $\beta_1(u_1)<\beta_2$, thus we must finish the service of L_1 before the service of L_2. But the common sense is a better adviser than the minimum principle. If we start with L_2, with probability $^1/_2$ its service is completed immediately, nobody must wait. If its service not completed within an infinitesimal period, then we interrupt its service and the minimum principle follows. Thus the loss is reduced to $\beta_1/2$ from β_1. So the minimum principle is not optimal.

To avoid any misunderstanding we need to emphasize that our counter example is valid, even if the arrivals form a Poisson process.

2) Moreover, this conjecture is not correct for every realization of the arbitrary arrival process, though the service-times are constant, if the waiting-costs are different. Let us assume $\beta_1/c_1<\beta_2/c_2<\beta_3/c_3$ and $\beta_3<\beta_2$. Similarly to the counter example of section 1, units L_2 and L_3 are present, when the server becomes free, and L_1 arrives after time t, where $\beta_3<t<\beta_2$, etc. I.e. The principle of SCHRAGE and MILLER is not correct, if the waiting costs are different.

3) The minimum-principle is not optimal for an arbitrary arrival process (realization), even if each service-time distribution has nonincreasing expected remaining service-time and equal waiting cost. Let be the arrival situation the same as in the previous counter-example, the service-times of L_1 and L_2 are constant and the service-time of L_3 is uniformly distributed over $[0, 2\beta_3]$, etc.

But we are interested in theorems and not in counter-examples. And we must admit that in the previous problems the optimal priority disciplines were determined by the minimum-principle at every selection point.

We must mention ROTHKOPF's interesting paper [15], published in 1966. *Though he assumed no arrival after the beginning of the whole service-process, but under this assumption and under the assumption of nonincreasing $\beta(u)$'s, he found and proved the optimality of the minimum principle.* As there is no arrival, the minimum-principle excludes any preemption, thus it is a simple nonpreemptive priority discipline discussed in Theorem 1.

At this point we must touch some general problems concerning distribution functions of nonincreaing remaining expected service-time. We need the following important quantity: $p(u)=b(u)/\overline{B}(u)$, $\overline{B}(u)>0$; which is called the *completion rate of the service*, when the elapsed service-time is u, and the service is not yet completed. It is clear, $p(u)du$ is the probability of the completion until $u+du$, under the conditions mentioned above. (In the theory of reliability $p(u)$ is called the *failure rate.*) Let

us analyse the *renewal process*, generated by random service-times of identical and independent distribution functions $B(u)$. It is well-known, that the density function of the asymptotical remaining service-time is $b^*(u)=\bar{B}(u)/\beta$, and the corresponding completion rate is $p^*(u) = \bar{B}(u)/\int_u^\infty \bar{B}(t)dt = 1/\beta(u)$. It is easy to prove, *if* $p(u)$ *is nondecreasing, then* $p^*(u)$ is also nondecreasing, i.e. $\beta(u)$ *is nonincreasing.* See GNEDENKO, BELIAEV and SOLOVIOV [10c] on the theory of reliability.

By the way, let us remark, that the result of AVI-ITZHAK and NAOR [3], according to which $\beta^*=\gamma/2\beta$ (see p. 230 of the present paper), is based on the following relations:

$$\beta^* = \int_0^\infty ub^*(u)\,du = \int_0^\infty u\bar{B}(u)\,du/\beta = \int_0^\infty \beta(u)\bar{B}(u)\,du/\beta = (\beta^2+\sigma^2)/2\beta = \gamma/2\beta.$$

Clearly, if $\beta(u)$ is nonincreasing, then $\beta^* = \int_0^\infty \beta(u)\bar{B}(u)du/\beta \leqq \int_0^\infty \bar{B}(u)du = \beta$, or which is the same: $\sigma \leqq \beta$. Consequently, if $\beta(u)$ is nonincreasing, then the condition of Theorem 3 (iii) is valid.

Let us return to $p(u)$ and $\beta(u)$. As $p(u) = -\frac{d}{du}\log \bar{B}(u)$, nondecreasing $p(u)$ is equivalent to log-concave $\bar{B}(u)$. According to PRÉKOPA's fundamental theorem [12], [13]: *if the density function* $b(u)$ *is log-concave, then the distribution function* $B(u)$ *and its complementer,* $\bar{B}(u)$ *are also log-concave.* Furthermore, in this case $p(u)$ *is nondecreasing, hence* $\beta(u)$ *is nonincreasing.* We remark that PRÉKOPA's fundamental theorem is very useful to verify these conditions, but the proof of the theorem is rather complicated. Therefore the direct and simple proof due to PRÉKOPA and ZOLOTARIOV — "If $b(u)$ is log-concave, then $\beta(u)$ is nonincreasing" — is interesting in itself. It is easy to see that $b(u)$ is log-concave in the following important cases: Uniform (rectangular) distribution, Gamma distribution (including the exponential distribution), Beta distribution, the normal distribution, etc.

LEMMA 1. *If the service time distribution* $B(u)$ *has non-decreasing completion rate* $b(u)/\bar{B}(u)$ *then both the conditional density function* $b(t+u)/\bar{B}(u)$ *and the conditional distribution function* $[B(t+u)-B(u)]/\bar{B}(u)$ *are also non-decreasing function of the elapsed service time* u.

From now on we always assume that the service completion rates are non-decreasing. Then we can prove

LEMMA 2. *If the completion rate of each unit is non-decreasing, then the optimal endogenous discipline (if it exists at all) does not begin to serve a unit until another unit of the same priority class is after its start but before its completion.*

PROOF. Let us assume the opposite of it; there exists a moment t_0, and there exist two units L_1 and L_2, both belonging to G_i and $u_1(t_0)>0=u_2(t_0)$ and the service of L_2 starts at this moment. For a moment let us fix the realization of the service time of every other unit. As they belong to common class, we deal with waiting times instead of waiting losses. The (netto) waiting time of L_1 and L_2 is the

sum of three types of periods: 1) one waits for the other; 2) both units are waiting for other units; 3) one has already left the queue, but the second one still waits.

Since both units wait during the second periods, it is doubled in the total waiting time. Since the total length of the second and the third periods is independent of the relative order of L_1 and L_2, the total waiting time under these periods is minimal if the second subtotal is minimal. By Lemma 1 this is the case if L_2 does not preempt L_1.

Similarly, by Rothkopf's theorem, the first subtotal is minimal in the same case. Consequently, the total waiting time is minimal if L_2 does not preempt L_1, contrary to our assumtion.

Now we restrict the analysis to the case of two-class disciplines. We shall prove

THEOREM 5 (i). *If there are two priority classes and if each service time distribution has nondecreasing completion rate, then there exists the optimal endogenous priority discipline, which is determined by the minimum principle: the optimal discretionary discipline.*

Remark: Our theorem holds for constant service time "distributions". — If $r=2$, then the knowledge of the arrival process does not influence the optimal discipline, thus Theorem 5 (i) is the generalization of Theorem 4 (i). Theorem 5 (i) proves the same result as the mentioned theorems of SCHRAGE—MILLER and ROTHKOPF, but the conditions are different: Theorem 5 (i) and the theorem of SCHRAGE and MILLER hold even if arrivals take place after the beginning of the service, contrary to ROTHKOPF's theorem. Theorem 5 (i) and the theorem of ROTHKOPF hold even if the service-times are random and the waiting costs are different, contrary to the assumption of SCHRAGE and MILLER. But Theorem 5 (i) holds for only two-class disciplines, contrary to the other theorems mentioned above.

PROOF. If the arrivals take place infinitely many times, then an optimal priority discipline may contain non optimal finite parts. But if each finite part of a discipline is optimal, then the whole discipline is optimal, moreover, it is *the* optimal discipline. Thus we may assume, there are units of finite number.

Let us assume the opposite of the statement of Theorem 5 (i): There exists the optimal discipline and the minimum-principle does not hold at least one time, and the minimum-principle does not hold at least one time. According to the previous lemmas there are only two possibilities:

1) An arriving priority unit preempts an ordinary unit too late;

$$L_1(0) \succ L_2(u_2), \qquad u_2 > u^0$$

2) A too "young" ordinary unit is prior to a priority unit:

$$L_2(u_2) \succ L_1(u_1), \qquad u_2 < u^0$$

As the proofs of 2) is the same as the proof of 1), we omit the former. Fix an L_2 which is preempted by priority units with positive probability. Let L_1 be the last priority unit of them. (As the number of the units is finite, there is such a unit.) Now the newly arriving units do not influence this preemption, hence Theorem 4 (i) holds, i.e. $\beta_2(u_2')/c_2 > \beta_1/c_1$, hence $u_2' < u^0$. But $u_2' \equiv u_2$ and $u_2 \equiv u^0$, i.e. $u_2' \equiv u^0$, which is a contradiction.

Remark: After the completion of the first version of this paper Gittins's paper [10 a] was published which is related to the Part 5 (i).

A (iii): In the remaining section of this Part we assume that *the arrival process of each class is an independent Poisson process.* We are interested in the stationary case, therefore we assume that the priority function is also stationary. We can simplify the earlier notations: we omit the sign of the exogenous time. By Lemma 2. the optimal priority discipline does not begin to serve a unit if the service of an other unit of the same class has already begun but has not completed. Therefore, we may assume that at every state at most one unit of a given class has got any service, which has arrived first. Let u_i be the elapsed service time of this unit, and let q_i be the number of C_i units in the queue, including the unit under service. Evidently, the state is described by $v = (u_i, q_i)_{i=1}^{r} = (u, q)$. The priority index (of the optimal discipline) is an integer $m(v)$ between 1 and r, which determines that in state v the first unit of $C_{m(v)}$ is served.

At this point we shall disprove SEVCIK's following statement: „When all distribution functions have decreasing remaining expected service times, then preemption at times other than the arrival instants is unnecesary to minimize expected total loss." ([17/a], p. 32.)

It is sufficient to modify our counter example 2 on p. 235 as follows: L_1 arrives immediately after the start with probability almost 1 *or* time t after the start with the remaining, almost zero probability. It is easy to prove that the optimal discipline starts with L_2, but in the unlikely, though possible event, when L_1 does not arrive immediately, (but in t), then L_3 preempts L_2 immediately.

Similarly to the previous situations the absence of the total randomness of the arrival process prevents the minimum principle to be optimal. In contrast, we have the following.

CONJECTURE. *If the units of the different classes arrive in independent Poisson streams and if the completion rate of each class is non-decreasing, then the optimal discipline (if it exists at all) permits interruptions only at arrival times and only by arriving units.*

The first part is proved in Lemma 4, it is needed only to the second part. From now on we consider only disciplines satisfying the statement of the conjecture. We should refer to the optimal discipline within this class as the *relative* optimal one, but we drop the word „relative". But before this drastic limitation we state

CONJECTURE 5 (iii). *Under the condition of the previous conjecture the relative optimal discipline is the absolute one.*

In the following part we shall prove through the bulk of lemmas that the optimal priority discipline is a discretionary discipline discussed in 4 (iii). Moreover, Theorem 4 (i) implies the main result of this paper:

THEOREM 5 (iii). *If the units of different classes arrive in independent Poisson streams and if the completion rate of the service of each class is nondecreasing, then there exists the optimal endogenous discipline, which is determined by the minimum principle: the optimal discretionary discipline of Theorem* 4 (iii).

Remark: We have mentioned that the optimality of the minimum principle was suggested by BRONSTEIN and VEKLEROV [6], but they did not define the endogenous discipline and did not notice the crucial condition of the nondecreasing completion rate.

BRONSTEIN and VEKLEROV [6] worried that the application of the minimum-principle would have been too complicated and expensive. Let us remark, if our conditions hold, then the minimum-principle determines a simple and cheap priority discipline. If our condisions do not hold, then the minimum-principle is not optimal in general, so there is no reason to apply it, consequently its complicatedness has not any practical significance.

VEKLEROV's theorem [18] cited at the end of 2 (iii) is a special case of our Theorem 5 (iii): when each service time has an exponential distribution.

We shall prove six lemmas:

LEMMA 3. *Under the conditions Theorem* 5 (iii) *the optimal priority index is independent of the number of the units in the different queuing classes, it depends on these classes.*

PROOF. According to our notations, Lemma 3 states the following: $m(u, q) = m(u, \sum_{q_h>0} e_h)$, where e_h is the h-th r dimensional unit vector. The statement is trivial, if $q_k \leqq 1$ for each k, $1 \leqq k \leqq r$. We shall show that $m(u, q+e_h) = m(u, q)$ if $q_h > 0$. According to Lemma 1, $m(u, q+e_h)$ equals to h or $m(u, q)$. By Lemma 2 the arriving L_h does not preempt $L_{m(u,q)}$, because there was at least another L'_h at his arrival. Thus $m(u, q+e_h) = m(u, q)$.

LEMMA 4. *Under the mentioned conditions the optimal priority index does not change if only the elapsed service time of the priority unit increases.*

PROOF. By notations: if $m(u, q) = i$ and if $u'_j = u_j$, $j \neq i$, $u'_i > u_i$, then $m(u', q) = i$. Of course, we assume the service has not yet completed: $B_i(u'_i) < 1$. With positive probability it may happen that the actual service is before its completion at u'_i and there is no arrival between u_i and u'_i. In this case $m(u, q)$ does not change by Lemma 1: $m(u', q) = m(u, q)$.

LEMMA 5. *Under the mentioned conditions we can choose such an index-system, that if every unit in the queue is before start, then the optimal priority index is the minimal integer the class of which is not empty in the queue.*

PROOF. If $m(0, \sum_{h=1}^{r} e_h) \neq 1$, let us change the original notation to satisfy $m(0, \sum_{h=1}^{r} e_h) = 1$. (We do not use any notation for the distinction between the old and the new indices.) If we achieved $m(0, \sum_{k=i}^{r} e_k) = i$, $i < r$, then we can also achieve $m(0, \sum_{k=i+1}^{r} e_k) = i+1$ by replacing the left index by $i+1$ and vice versa. As $m(0, \sum_{k=i+1}^{r} e_k) \geqq i+1$, the replacing does not change $\sum_{k=i+1}^{r} e_k$.

Next we shall show that $m(0, \sum_{h \in H} e_h) = \min \{h, h \in H\} = \min H$. Let us assume

the opposite. Moreover, let us assume our set H has the greatest number of elements with this property. Clearly, there exist at least one integer k between 1 and r, for which $k>\min H=g$ and $k\notin H$. With positive probability L_k arrives first, in time Δ. By the maximum property of H and by $k>g$, $m\left(0, e_k+\sum_{h\in H} e_h\right) = \min\left[\{k\}\cup H\right] = g$.

Let be $j=m\left(0, \sum_{h\in H} e_h\right)$. If $L_j(\Delta)\succ L_k$ then $m\left(e_j\Delta, e_k+\sum_{h\in H} e_h\right)=j$. As Δ is arbitrary small, $m\left(0, e_k+\sum_{h\in H} e_h\right) = j$, which implies $j=g$. Contradiction.

If $L_k\succ L_j(\Delta)$, then, similarly to Lemma 4, $m\left(0, e_k+\sum_{h\in H} e_h\right) = k$, i.e. $k=g$, contradiction.

By Lemma 3 $m(0, q)=m\left(0, \sum_{q_h>0} e_h\right)$, and $m\left(0, \sum_{q_h>0} e_h\right) = \min\{h, q_h>0\}$, thus $m(0, q)=\min\{h; q_h>0\}$.

LEMMA 6. *Under the mentioned conditions if an arriving unit does not preempt the priority unit, then a later arriving unit preempts or does not preempt the priority unit independently of the presence of the earlier unit. The same holds for the relation between the two arriving units.*

PROOF. Suppose that just before the arrival of L_k $v=(u, q)$ and $m(u, q)=j$. After the arrival $m(u, q+e_k) = j$. We can assume there is no arrival between the arrivals of L_k and L_i. We state $m(u', q+e_k+e_i) = m(u', q+e_i)$. By the restriction both of them equals to i or j. If they are different, there are only two possibilities: $i\neq j$ or $j\neq i$. Suppose the first possibility. If $m(u', q+e_i) = j$ and L_k arrives Δ later than L_i, then $m(u'+e_j\Delta, q+e_i+e_k)$ equals to j or k. As Δ is arbitrary small, $m(u', q+e_i+e_k)$ equals to j or k, too; but differs from i, because by Lemma 2 we can assume $i\neq j$, $i\neq k$.

By Lemma 4 $m(u', q+e_k+e_i) = i$ implies $m(u'+u_ie_i, q+e_k+e_i) = i$. But the same state is generated if L_k arrives after L_i, when L_i preempted $L_j(u_j')$, i.e. $L_i(u_i)$ is prior to L_k.

Finally, we must prove the same if $L_j(u_j')$ is missing from the system.

Evidently, $m(u', q+e_i) = i$ implies $m(u'-u_j'e_j; q+e_i) = i$. According to the first part of the lemma, $m(u'-u_j'e_j, q-e_j+e_i) = i$. According to the first part of the lemma, $L_i\succ L_k$ is independent of L_j, i.e. $m(u'-u_j'e_j+u_ie_i, q+e_i+e_k) = i$. If our statement did not hold, then L_k would interrupt $L_i(u_i)$, if $L_j(u_j)$ were missing.

With positive probability L_j arrives Δ later than L_k, and as above $m(u'-u_j'e_j+u_ie_i+\Delta e_k, q+e_i+e_k)$ remains k or jumps on j. As Δ is arbitrary, $m(u'-u_j'e_j+u_ie_i, q+e_i+e_k)$ also equals to k or j, but anyway it differs from i, which contradicts to a previous statement.

LEMMA 7. *If we remove all but two units from the system, where one of the remaining two units is the original priority unit, then it remains the priority unit.*

First we shall prove that those units which took part in the generation of a state, but they have disappeared with completed service, do not influence the further history of the process.

By Lemma 2 at most $r-1$ interruptions may take place between two subsequent completions. As above, during this period the optimal priority index may change

service starts. According to Lemma 6 we can assume that those units which arrived during this nonpreempted service period arrived at the end of it. Furthermore, if this period is removed, then the further order of the service does not change. only at arrivals. Consequently, not later than the r-th arrival the first nonpreempted Repeating this procedure, we prove the statement. In the followings, we neglect the transient presence of the completed units.

It is easy to see, that state (u, q) has the following (reduced) history: L_{i_1} arrived into the empty system and its service was preempted by L_{i_2} at u_{i_1}. During this period some units arrived, but they did not interrupt the service of L_{i_1}. By Lemma 3 we can assume $q_h \leqq 1$, $1 \leqq h \leqq r$, thus it is enough to describe the index-set of these units: K_1. If the number of the positive u_h's is p, then we can define i_n and K_n for each n, $1 \leqq n \leqq p$.

By Lemma 6 $L_{i_2} \succ L_{i_1}(u_{i_1})$ is independent of the presence of L_k's, $k \in K_1$, furthermore we can assume they arrived during the service of L_{i_2}. This modification does not influence the subsequent priority changes, either. Repeating this modification $p-2$ times, we can assume that all units which are before their start arrived after $L_{i_p}(u_{i_p})$. By Lemma 6 $L_{i_p}(u_{i_p}) \succ L_k$ independently of $L_{i_{p-1}}(u_{i_{p-1}})$. If we eliminate $L_{i_{p-1}}$, then $L_{i_p}(u_{i_p}) \succ L_k$ independently of $L_{i_{p-2}}(u_{i_{p-2}})$, etc. Again by Lemma 6 $L_{i_p}(u_{i_p}) \succ L_k$ independently of other L_k''s. I.e. $m(u_{i_p} e_{i_p}, e_{i_p} + e_k) = i_p$.

Now we can drop all units which are before their start, without any consequence on the optimal priority index. We must prove that $m(u_{i_p} e_{i_p} + u_{i_n} e_{i_n}, e_{i_p} + e_{i_n}) = i_p$.

If $p=2$, then it is trivial. Let us assume that our statement holds for $p-1$, then we must prove the same for p. We can choose an integer l which is different from n and p. If we reduce u_{i_l} to zero, then the priority index does not change, but L_{i_l} can be eliminated and the problem is reduced to the case of $p-1$.

LEMMA 8. *The optimal endogenous priority discipline is the optimal discretionary discipline.*

According to Lemma 5 if all the units are before their start, then a unit of minimal index starts the service. According to Lemma 7 it is enough to deal with the arrival states. If at state (u, q) $m(u, q)=j$, and the arriving L_i does not preempt it, then $m(u+e_j \Delta, q+e_i) = j$, $\Delta \geqq 0$ by Lemma 5. Let be $u_{ij}^0 = \sup \{u_j;\ L_i \succ L_j(u_j)\}$, then $L_i \succ L_j(u_j)$ if and only if $u_j < u_{ij}^0$. By Lemma 7 this relation is independent of the other coordinates of the state-vector. According to 4 (i), u_{ij}^0 is the (i, j)-th discretionary point; $u_{ij}^0 = 0$ if $i > j$ by Lemma 5.

To complete the proof we shall show that $u_{ij}^0 \geqq u_{kj}^0$, for each i, j, k; $i < k < j$. The proof is the reverse of the procedure which reduces the selection rule to the interruption rule in 4 (iii). Let us assume the opposite of it: there exist indices i, j and k; where $i < k < j$ and $m(u_j e_j, e_j + e_k) = k$ while $m(u_j e_j, e_j + e_i) = j$; $u_{ij}^0 \leqq u_j < u_{kj}^0$. By Lemma 6 $m(u_j e_j, e_j + e_i + e_k) = k$, hence by Lemma 7 $m(0, e_i + e_k) = k$, which contradicts to Lemma 5.

Hence the optimal endogenous discipline is a discretionary discipline, consequently the optimal one.

Remark: In the proofs of Lemmas 3—8 we have used the following property of the Poisson arrival process: At each time the future of the process is independent of its past.

6. Queues with many servers

Until now we assumed there is only *one* server, following the other authors, except ROTHKOPF [15]. If each server has an independent queue, then there is no problem, but usually independent queues are not economical. If there is a unique queue for all the servers, but preemptions are excluded, then the problem of the optimal idleness appears again; it is worth to wait to the fast server, who is occupied, though the slower server is free. If all servers have the same service ability, then this problem disappears, but the following complication takes place (see ROTHKOPF [15], pp 709—710.):

There are two identical servers and four units with exponential service time distributions. Suppose, $\beta_1 \gg \beta_2 \gg \beta_3 = \beta_4$, $c_1 \gg c_3 = c_4 > c_2$ and $\beta_1/c_1 < \beta_2/c_2 \lesssim \beta_3/c_3 = \beta_4/c_4$. If L_2, L_3 and L_4 are present at the beginning, Theorem 1 (i) *suggests* to begin the service of L_2 and L_3. One can prove that L_4 waits for L_2 or L_3 *in avarage* $1/(b_2+b_3)$, its expected loss is $c_3/(b_2+b_3)$. But let us start with the service of L_3 and L_4. Then the expected loss is $c_2/2b_3$, which is smaller than the former loss. *Consequently, Theorem* 1 (i) *and Theorem* 2 (i) *do not hold, if there are two servers.*

If all units are present at the start, then L_1 "has a cost so large that it must be one of the tasks placed in service at once. However, b_1 is small enough that the processing of the other tasks can start on the remaining processor as it were the only processor. Since L_2 has the largest c_i/β_i of the remaining three tasks, it would start first. In the unlikely, but possible, event that L_1 completes service before L_2," we have the previous situation, thus *the service of* L_2 *must be interrupted, because the other server completed the service.*

This counter example shows that our Conjecture does not hold, if there are more than one servers.

The author of the present paper gave a lecture on this topic at a seminar of the Computer Center of the Hungarian Academy of Sciences. The participants, especially Professor PRÉKOPA helped to clarify some problems of this paper.

Addendum. After the completion of the present paper I have got SEVCIK's paper [17 a] which generalized ROTHKOPF's problem on the optimal endogenous priority discipline of the no-arrival case for arbitrary service-time distributions. He conjectured that his optimal priority index was optimal under A (iii) (Poisson arrivals), which was proved by GITTINS and NASH [10b], generalizing our Theorem 5 (iii) and conjecture 5 (iii).

REFERENCES

[1] ACZÉL, M.: The effect of introducing priorities, *Operations Research* **8** (1960), 730—733.
[2] AVI-ITZHAK, B., BROSH, I. and NAOR, P.: On discretionary priority queuing, *Zeitschrift für Angewandte Mathematik und Mechanik* **44** (1964), 235—242.
[3] AVI-ITZHAK, B. and NAOR, P.: Some queuing problem with service station subject to breakdown, *Operations Research* **11** (1963), 303—320.
[4] Бронштейн, И. О.: О многоэтапном обслуживании одним прибором, *Техническая Кибернетика* **5**, (1969), 59—64.
[5] Бронштейн, И. О. и Рыков, В. В.: Об оптимальных приоритетах в системах массового обслуживания, *Техническая Кибернетика* **6** (1965), 28—37.
[5a] Бронштейн, И. О. и Веклеров, Е. Б.: Об одном классе дисциплины обслуживания, *Техническая Кибернетика,* **5** (1967)

[6] Бронштейн, И. О. и Веклеров, Е. Б.: Об одном классе дисциплин обслуживания, — *Массовое обслуживание в системах передачи информации* — Наука, Москва 1969, стр. 54—59

[7] Brosh, I. and Naor, P.: On optimal disciplines in priority queuing, *Bulletin of Institute of International Statistics* **40** (1963), 593—609

[8] Cobham, A.: Priority assignement in waiting line problems, *Operations Research 2* (1954), 70—76.

[9] Cox, D. R. and Smith, W. L.: *Queues,* Meuthen, London, 1961.

[10] Fife, D. W.: "Scheduling with random arrivals and linear loss functions", *Management Science* **11** (1965) 429—437.

[10 a] Gittins, J. G.: On the optimal allocation of resources, *Journal of the Applied Probability,* 1972.

[10 b] Gittins, J. C. and Nash, P.: *A priority index for queueing systems,* Cambridge University, Engineering Deparment. (England) manuscript, January 1974

[10 c] Гнеденко, Б. В., Беляев, Ю. К., Соловьев, А. Д.: *Математические методы в теории надежности,* Изд. Наука, Москва, 1965.

[11] Jaiswal, N. K.: *Priority Queues,* Academic Press, New York and London, 1968. pp. 240.

[12] Prékopa, A.: „*Sztochasztikus rendszerek optimalizálási problémái*", Akadémiai doktori disszertáció, kézirat. („Optimization Problems of Stochastic Systems", manuscript, in Hungarian), 1971.

[13] Prékopa, A.: Contributions to the theory of stochastic programming, *Mathematical Programming* (to appear).

[14] Рыков, В. В. и Лемберг, Э. Е.: Об оптимальных динамических приоритетах систем массового обслуживания, *Техническая Кибернетика,* **2** (1967), 25—34.

[15] Rothkopf, M.: Scheduling with random service times, *Management Science,* **12** (1966). 707—713.

[16] Schrage, L. E.: A proof of the optimality of the Shorthest Remaining Processing Time Discipline, *Operations Research* **16** (1968), 687—689.

[17] Schrage, L. E. and Miller, L. W.: The queue M(G)1 with the Shorthest Remaining Processing Time Discipline, *Operations Research* **14** (1966), 670—683.

[17 a] Sevcik, K. C.: *The use of service time distributions in scheduling,* Technical Report CSRG-14 May 1972 Departments of Electrical Engineering and Computer Science of the University of Toronto.

[17b] Simonovits, A.: On discretionary queuing discipline, *Proceedings of the European Meeting of Statisticians* (to appear). Budapest, 1972.

[18] Веклеров, Е. Б.: Об оптимальных абсолютных динамических приоритетах в системах массового обслуживания, *Техническая Кибернетика,* **2** (1967), 87—90.

[19] White, H. and Christie, L. S.: Queuing with preemptive priorities, *Operations Research* **6** (1958), 79—95.

Institute of Economics, Hungarian Academy of Sciences

(Received May 15, 1972)

(Revised May 1, 1974)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 245—268.

PERTURBATIONS OF n DIMENSIONAL LINEAR SYSTEMS OF DIFFERENTIAL EQUATIONS

by

I. BIHARI

1. Introduction. The present paper is the continuation of the work [1] solving the problem for $n=3$. As in [1], here, too, the conditions, methods, and results correspond to a point of view different than in [2] or [3] and [4]. The main purpose aimed at was to preserve the asymptotic behaviour of the solutions of the corresponding linear system or, in other words, to establish a close one to one map between the solutions of the two systems. This will be detailed below.

The system we are concerned with has the form

$$\dot{x} = Ax + f(t, x), \qquad \left(\cdot = \frac{d}{dt}\right) \tag{1}$$

where $A = \text{const}$ is an $n \times n$ matrix, x and f are column vectors:

$$x = (x_1, \ldots, x_n), \quad x_i = x_i(t), \quad f = (f_1, \ldots, f_n), \quad f_i = f_i(t, x), \qquad (i = 1, 2, \ldots, n)$$

$f_i(t, x)$ is continuous for $t \geqq 0$ and every x.

The corresponding linear system is

$$\dot{y} = Ay. \tag{2}$$

Let $A = J_1 \oplus J_2 \oplus \ldots \oplus J_k$ be the Jordan normal form of A, J_i having the order $a_i \geqq 1$. Then $\sum_{i=1}^{k} a_i = n$. Here

$$J_i = \begin{bmatrix} \lambda_i & & & 0 \\ 1 & \lambda_i & & \\ & \ddots & \ddots & \\ 0 & & 1 & \lambda_i \end{bmatrix}, \quad \operatorname{Re} \lambda_i = \alpha_i, \quad \max \alpha_i = \alpha, \quad \alpha_i \neq 0. \tag{3}$$

If

$$Y(t) = e^{At} = e^{J_1 t} \oplus e^{J_2 t} \oplus \ldots \oplus e^{J_k t}, \quad e^{J_i t} = e^{\lambda_i t} \begin{bmatrix} 1 & & & & \\ t & 1 & & & \\ \frac{t^2}{2!} & t & 1 & & \\ \vdots & & \ddots & \ddots & \\ \frac{t^{a_i-1}}{(a_i-1)!} & & \frac{t^2}{2!} & t & 1 \end{bmatrix} \tag{4}$$

then the general solution of (2) is $y=Yc$, $c=y(0)$. Here c is an n-vector which will be partitioned (divided in subvectors) in the following way:

$$c = c_1 \oplus c_2 \oplus \ldots \oplus c_k \quad c_i = (c_{1i}, c_{2i}, \ldots, c_{a_i,i}),$$

where the order of c_i is a_i. Then

$$y = Yc = e^{J_1 t} c_1 \oplus e^{J_2 t} c_2 \oplus \ldots \oplus e^{J_k t} c_k. \tag{5}$$

Here

$$e^{J_i t} c_i = e^{\lambda_i t} \begin{bmatrix} c_{1i} \\ c_{1i} t + c_{2i} \\ c_{1i} \dfrac{t^2}{2!} + c_{2i} t + c_{3i} \\ \vdots \\ c_{1i} \dfrac{t^{a_i-1}}{(a_i-1)!} + c_{2i} \dfrac{t^{a_i-2}}{(a_i-2)!} + \ldots + c_{a_i-1,i} t + c_{a_i,i} \end{bmatrix}. \tag{6}$$

The corresponding partitionings of f, x, y are

$$f = f_1 \oplus f_2 \oplus \ldots \oplus f_k; \quad x = x_1 \oplus x_2 \oplus \ldots \oplus x_k; \quad y = y_1 \oplus y_2 \oplus \ldots \oplus y_k.$$

The orders of f_i, x_i and y_i are a_i.

The solution of (1) satisfies the integral equation

$$x(t) = Y(t)c + \int_0^t Y(t-s) f(s, x(s))\, ds, \quad x_0 = x(0) = c. \tag{7}$$

Here

$$Y(t-s)f = e^{J_1(t-s)} f_1 \oplus e^{J_2(t-s)} f_2 \oplus \ldots \oplus e^{J_k(t-s)} f_k.$$

So the vector equation (7) splits into the system of (sub) vector equations as

$$x_i(t) = e^{J_i t} c_i + \int_0^t e^{J_i(t-s)} f_i\, ds \qquad i = 1, 2, \ldots, k, \quad f_i = f_i(t, x) \tag{8}$$

where $e^{J_i t} c_i = y_i$ is the solution of $\dot{y}_i = J_i y_i$ and

$$e^{J_i(t-s)} f_i = e^{\lambda_i (t-s)} \begin{bmatrix} f_{1i} \\ f_{1i}(t-s) + f_{2i} \\ f_{1i} \dfrac{(t-s)^2}{2} + f_{2i}(t-s) + f_{3i} \\ \vdots \\ f_{1i} \dfrac{(t-s)^{a_i-1}}{(a_i-1)!} + f_{2i} \dfrac{(t-s)^{a_i-2}}{(a_i-2)!} + \ldots + f_{a_i-1,i}(t-s) + f_{a_i,i} \end{bmatrix}. \tag{9}$$

Now we are able to formulate the program of the present work. We state that under suitable assumptions there exists a one to one correspondence between the solutions of (1) and (2) in such a way that not only the relation

$$|x(t) - y(t)| = o(t^{a-1} e^{\alpha t}), \qquad t \to \infty, \quad a = \max_{\alpha_i = \alpha} a_i \tag{10}$$

is satisfied (as in [4]), but also the relations

$$x_{li}-y_{li}=o(t^{l-1}e^{\alpha_i t}), \qquad t\to\infty, \quad (i=1,\ldots,k;\ l=1,\ldots,a_i) \tag{11}$$

Relations (11) distinguish the behaviour of the components of the solution and assure a more intrinsic topological affinity of $x(t)$ and $y(t)$ than (10) (see [1]).

In order to reach this end suitable assumptions are needed concerning the perturbation $f(t, x)$. In 2—4 we suppose $\alpha_i<0$ $(i=1, \ldots k)$ and

$$|f_{li}(t,x)|\leqq\Phi_i(t,r), \qquad r=|x|, \quad i=1,2,\ldots,k; \quad l=1,2,\ldots,a_i, \quad r<r_1 \tag{12}$$

where $|x|$ is an appropriate vector norm of x, $r_1>0$ some constant and

$$\Phi_i(t,r)=(1+t)^{(a-1)(1-m)}r^{\frac{\alpha_i}{\alpha}}\theta(t,r), \quad m=\max\frac{\alpha_i}{\alpha}\geqq 1. \tag{13}$$

Here $\theta(t, r)$ is continuous for $t\geqq 0$, $r\geqq 0$, $r<r_1$ and nondecreasing in r for a fixed t, finally $\theta(t, 0)=0$ and it will be assumed that

$$\int^{\infty} t^{p-1}\theta(t, Ct^{a-1}e^{\alpha t})\,dt<\infty \tag{14}$$

where $C>0$ is an arbitrary constant. In Lemmas 1 and 2 $p=a$ while in Lemma 3 $p=\max a_i$. — Notice that for $\theta(t, r)=\omega(r)$ (14) is equivalent to

$$\int_{+0}\left(\log\frac{1}{r}\right)^{p-1}\frac{\omega(r)}{r}\,dr<\infty. \tag{14'}$$

2. LEMMA 1 (an a priori estimate of $r(t)$). *Along every path corresponding to a sufficiently small* $r_0<r_1(r_0=|c|=|x_0|=|x(0)|)$

$$r(t)<Cr_0(1+t)^{a-1}e^{\alpha t}, \qquad t\geqq 0, \quad C=2M=\text{const} \tag{15}$$

holds where M will be defined presently.

PROOF. First we remark that the estimate

$$|Y|<M(1+t)^{a-1}e^{\alpha t}, \qquad t\geqq 0 \tag{16}$$

holds with some $M=\text{const}>0$ depending on the matrix A. By (7) and (16)

$$r=|x|<Mr_0(1+t)^{a-1}e^{\alpha t}+M\int_0^t[1+(t-s)]^{a-1}e^{\alpha(t-s)}\left|f\big(s,x(s)\big)\right|ds$$

and by (13)

$$\left|f\big(s,x(s)\big)\right|<\sum_{i,l}|f_{li}|\leqq(1+s)^{(a-1)(1-m)}\theta\big(s,r(s)\big)\sum_{i=1}^{k}a_i r^{\frac{\alpha_i}{\alpha}}.$$

Since $\min\dfrac{\alpha_i}{\alpha}=1$, we have $\sum\limits_{i=1}^{k}a_i r^{\frac{\alpha_i}{\alpha}}\leqq nr$ provided $r<1$. (This can be supposed if r_0 is sufficiently small; see (15)). By introducing the notation

$$p(t)\equiv Mr_0(1+t)^{a-1}+nM\int_0^t[1+(t-s)]^{a-1}e^{-\alpha s}(1+s)^{(a-1)(1-m)}r(s)\theta\big(s,r(s)\big)\,ds$$

it is sufficient to show that

$$p(t) < Cr_0(1+t)^{a-1}, \qquad t \geqq 0, \quad C = 2M,$$

which holds for sufficiently small non-negative t's. If it were not true for every $t>0$, then there would exist a first (smallest) $T>0$ with

$$p(T) = Cr_0(1+T)^{a-1}, \quad \text{and} \quad p(t) < Cr_0(1+t)^{a-1}, \qquad 0 < t < T$$

or, what is the same,

(17)

$$S = nM \int_0^T [1+(T-s)]^{a-1} e^{-\alpha s}(1+s)^{(a-1)(1-m)} r(s)\theta\big(s, r(s)\big)\, ds = Mr_0(1+T)^{a-1}.$$

However, this cannot occur. Indeed, since $1-\frac{s}{1+T} < 1$ and $m \geqq 1$, we have

$$\frac{S}{(1+T)^{a-1}} <$$

$$< nM \int_0^T \left(1-\frac{s}{1+T}\right)^{a-1} (1+s)^{(a-1)(1-m)} e^{-\alpha s}\theta\big(s, Cr_0(1+s)^{a-1}e^{\alpha s}\big) Cr_0(1+s)^{a-1}e^{\alpha s}\, ds <$$

$$< nMCr_0 \int_0^T (1+s)^{a-1}\theta\big(s, Cr_0(1+s)^{a-1}e^{\alpha s}\big)\, ds.$$

Contrary to (17) $\frac{S}{(1+T)^{a-1}} < Mr_0$, provided

$$nC \int_0^T (1+s)^{a-1}\theta\big(s, Cr_0(1+s)^{a-1}e^{\alpha s}\big)\, ds < 1, \qquad C = 2M.$$

This is valid if we assume that

$$2nM \int_0^\infty (1+s)^{a-1}\theta\big(s, 2Mr_0(1+s)^{a-1}e^{\alpha s}\big)\, ds < 1$$

and (14) shows, that the lass inequality is fulfilled if only $r_0\ (<r_1)$ is small enough, since $\theta(t, 0)=0$.

3. THEOREM 1.

1° *If* (12)—(14) *hold and* $0>\alpha_1>\ldots>\alpha_k$, *then to every solution* $x(t)$ *of* (1) *there belongs a solution* $y(t)$ *of* (2) *satisfying*

$$x_{ji} - y_{ji} = o(t^{j-1}e^{\alpha_i t}), \qquad t \to \infty, \quad j = 1, \ldots, a_i; \quad i = 1, \ldots, k$$

or equivalently

$$x = e^{J_1 t}\big(c_1 + o(1)\big) \oplus \ldots \oplus e^{J_k t}\big(c_k + o(1)\big),$$

where c_j *is an arbitrary vector of order* a_j.

2° *For every fixed* l $(l=1, 2, \ldots, a_1-1)$ *there exists a subfamily of solutions of* (1) *along which*

$$r<Cr_0(1+t)^{a_1-l-1}e^{\alpha_1 t}, \qquad t>0$$

(in the sequel $C>0$ *means a constant which can vary from case to case) and to this there belongs a* $y(t)$ *(solution of* (2)*) with*

$$x_{ji}-y_{ji}=o(t^{j-l-1}e^{\alpha_i t}), \qquad j=1, \ldots, a_i; \quad i=1, \ldots, k$$

or

(I) $$x=e^{J_1 t}(c_1+o(t^{-l}))\oplus \ldots \oplus e^{J_k t}(c_k+o(t^{-l})),$$

where $c_1=(0, \ldots, 0, c_{l+1,1} \ldots, c_{a_1,1})$, *i.e.* $c_{i1}=0$ *for* $i=1, \ldots, l$.

3° *There exists a subfamily with*

$$r<Cr_0(1+t)^{a_s-l-1}e^{\alpha_s t}, \qquad s=1, \ldots, k$$

having the asymptotic form

(II) $$x=x_1\oplus \ldots \oplus x_{s-1}\oplus e^{J_s t}(c_s+o(t^{-l}))\oplus \ldots \oplus e^{J_k t}(c_k+o(t^{-l}));$$

$$x_{jm}=o(t^{j-l-1}e^{\alpha_s t}); \quad m=1, \ldots, s-1; \quad j=1, \ldots, a_m; \quad l=0, \ldots, a_m-1;$$

$$c_{is}=0, \quad i=1, \ldots, l.$$

If there are also equal α*'s denote the different* α_i*'s by* β_j $(j=1, 2, \ldots, h)$. *Then* $0>\beta_1>\ldots>\beta_h$, $h<k$.

If

4° $\alpha_1=\ldots=\alpha_m=\beta_1$ *and* $a_1\geqq a_2\geqq \ldots \geqq a_m$, *then there are solutions of* (1) *with*

$$r<Cr_0(1+t)^{a_1-l-1}e^{\beta_1 t}, \qquad l=0, 1, \ldots, a_1-a_2$$

having the asymptotic form (I) *with* c_1 *given there,*

5° *and a solution with*

$$r<Cr_0(1+t)^{a_s-l-1}e^{\beta_1 t}, \qquad s=1, \ldots, m; \quad l=0, 1, \ldots, a_s-a_{s+1}$$

having the asymptotic form (I) *where* $c_{p,q}=0$, $p=1, \ldots, s+l-1$; $q=1, \ldots, a_p-a_s+l$, *i.e. all the terms of the polynomial parts involved in* $\exp(J_i t)\cdot c_i$ *with higher degree than* a_s-l-1 *vanish. — The case of equal blocks is here included.*

6° *Corresponding to* $c_i=0$ $(i=1, \ldots, m)$ *there is a solution with*

$$r<Cr_0(1+t)^{A_2-1}e^{\beta_2 t}, \quad A_2=\max_{\alpha_i=\beta_2} a_i$$

having the asymptotic form

(III) $$x=x_1\oplus \ldots \oplus x_m\oplus e^{J_{m+1}t}(c_{m+1}+o(1))\oplus \ldots \oplus e^{J_k t}(e_k+o(1))$$

where $x_{ji}=o(t^{j-1}e^{\beta_2 t})$, $i=1, \ldots, m$; $j=1, \ldots, a_i$.

7° *and there are also solutions with*

$$r < Cr_0(1+t)^{A_2-l-1}e^{\beta_2 t}, \qquad l = 1, \ldots, A_2-1$$

etc.

So the set of the solutions of (1) can be arranged in classes (subsets, subfamilies) the members of which have the same asymptotic behaviour as the corresponding subfamily of the solutions of (2). The PROOF proceeds through several lemmas, but first we make some preparatory statements.

1° *Corollary of* LEMMA 1. *Asymptotic form of the solution.* By (12)—(14) we have

$$S = \left|\int_0^\infty f_{1i}e^{-\lambda_i s}\,ds\right| < \infty, \quad \text{moreover} \quad S < r_0,$$

provided r_0 is small enough. Indeed in this case

$$\int_0^t |f_{1i}|e^{-\alpha_i s}\,ds < \int_0^t (1+s)^{(a-1)(1-m)}r^{\frac{\alpha_i}{\alpha}}\theta(s,r)e^{-\alpha_i s}\,ds <$$

$$< \int_0^t [Cr_0(1+s)^{a-1}e^{\alpha s}]^{\frac{\alpha_i}{\alpha}}(1+s)^{(a-1)(1-m)}\theta\left(s, Cr_0(1+s)^{a-1}e^{\alpha s}\right)e^{-\alpha_i s}\,ds <$$

$$< Cr_0\int_0^\infty (1+s)^{(a-1)\left(1-m+\frac{\alpha_i}{\alpha}\right)}\theta\left(s, Cr_0(1+s)^{a-1}e^{\alpha s}\right)ds <$$

$$< Cr_0\int_0^\infty (1+s)^{a-1}\theta\left(s, Cr_0(1+s)^{a-1}e^{\alpha s}\right)ds < r_0.$$

Also it can be easily shown that

$$\int_0^t s^j f_{1i}e^{-\lambda_i s}\,ds = o(t^j), \qquad t \to \infty, \quad j = 1, 2, \ldots, a_i-1.$$

Let us consider the components of x_i

$$(18_1) \qquad x_{1i} = e^{\lambda_i t}\left(c_{1i} + \int_0^t f_{1i}e^{-\lambda_i s}\,ds\right) = e^{\lambda_i t}\left(\bar{c}_{1i} - \int_t^\infty f_{1i}e^{-\lambda_i s}\,ds\right)$$

where

$$\bar{c}_{1i} = c_{1i} + \int_0^\infty f_{1i}e^{-\lambda_i s}\,ds, \quad |\bar{c}_{1i}| < 2r_0, \quad \text{since} \quad |c_{1i}| < r_0.$$

This can be written as

$$x_{1i} = e^{\lambda_i t}(\bar{c}_{1i} + o(1)), \quad \text{or} \quad x_{1i} - \bar{y}_{1i} = o(e^{\alpha_i t}), \quad \bar{y}_{1i} = e^{\lambda_i t}\bar{c}_{1i}.$$

Similarly

$$(18_2) \qquad x_{2i} = e^{\lambda_i t}\left\{c_{1i}t + c_{2i} + \int_0^t [(t-s)f_{1i} + f_{2i}]e^{-\lambda_i s}\,ds\right\} =$$

$$= e^{\lambda_i t}\left[c_{1i}t + c_{2i} + t\left(\int_0^\infty - \int_t^\infty f_{1i}e^{-\lambda_i s}\,ds\right) - \int_0^t sf_{1i}e^{-\lambda_i s}\,ds + \left(\int_0^\infty - \int_t^\infty f_{2i}e^{-\lambda_i s}\,ds\right)\right] =$$

$$= e^{\lambda_i t}\left[\bar{c}_{1i}t + \bar{c}_{2i} - t\int_t^\infty f_{1i}e^{-\lambda_i s}\,ds - \int_0^t sf_{1i}e^{-\lambda_i s}\,ds - \int_t^\infty f_{2i}e^{-\lambda_i s}\,ds\right]$$

$$\bar{c}_{2i} = c_{2i} + \int_0^\infty f_{2i}e^{-\lambda_i s}\,ds.$$

Here the first and second integrals are $o(t)$, the third one is $o(1)$. Therefore

$$x_{2i} = e^{\lambda_i t}\left[(\bar{c}_{1i} + o(1))t + (\bar{c}_{2i} + o(1))\right]$$

or

$$x_{2i} - \bar{y}_{2i} = o(te^{\alpha_i t}), \quad \bar{y}_{2i} = e^{\lambda_i t}(\bar{c}_{1i}t + \bar{c}_{2i}).$$

In the same way

$$(18) \quad x_{li} = e^{\lambda_i t}\left[(\bar{c}_{1i} + o(1))\frac{t^{l-1}}{(l-1)!} + \ldots + (\bar{c}_{li} + o(1))\right], \quad \bar{c}_{li} = c_{li} + \int_0^\infty f_{li}e^{-\lambda_i s}\,ds$$

or

$$x_{li} - \bar{y}_{li} = o(t^{l-1}e^{\alpha_i t}), \quad \bar{y}_{li} = e^{\lambda_i t}\left(\bar{c}_{1i}\frac{t^{l-1}}{(l-1)!} + \ldots + \bar{c}_{li}\right)$$

$$l = 1, 2, \ldots, a_i; \quad i = 1, 2, \ldots, k; \quad |\bar{c}_{li}| < 2r_0.$$

2° *An important formula.* In general $\bar{c} \neq c$, i.e. $\bar{y}(0) \neq y(0)$. The following detailed formula will be necessary later:

$$(18_l) \qquad x_{li} = \left(c_{1i}\frac{t^{l-1}}{(l-1)!} + c_{2i}\frac{t^{l-2}}{(l-2)!} + \ldots + c_{l-1,i}t + c_{li}\right)e^{\lambda_i t} +$$

$$+ e^{\lambda_i t}\int_0^t \left(f_{1i}\frac{(t-s)^{l-1}}{(l-1)!} + f_{2i}\frac{(t-s)^{l-2}}{(l-2)!} + \ldots + f_{l-1,i}(t-s) + f_{li}\right)e^{-\lambda_i s}\,ds =$$

$$= e^{\lambda_i t}\left[\bar{c}_{1i}\frac{t^{l-1}}{(l-1)!} + \ldots + \bar{c}_{l-1,i}t + \bar{c}_{li} + \right.$$

$$+ \int_0^t \left(f_{1i}\frac{(t-s)^{l-1}}{(l-1)!} + \ldots + f_{l-1,i}(t-s) + f_{li}\right)e^{-\lambda_i s}\,ds -$$

$$- \frac{t^{l-1}}{(l-1)!}\int_0^\infty f_{1i}e^{-\lambda_i s}\,ds - \frac{t^{l-2}}{(l-2)!}\int_0^\infty f_{2i}e^{-\lambda_i s}\,ds -$$

$$\left. - \ldots - t\int_0^\infty f_{l-1,i}e^{-\lambda_i s}\,ds - \int_0^\infty f_{li}e^{-\lambda_i s}\,ds\right].$$

By a well known formula

$$(18') \qquad \frac{1}{(k-1)!}\int_0^t f(s)(t-s)^{k-1}\,ds = \int_0^t\int_0^{t_k}\int_0^{t_{k-1}}\dots\int_0^{t_2} f(t_1)\,dt_1\,dt_2\dots dt_k =$$

$$= \frac{t^{k-1}}{(k-1)!}\int_0^\infty f(s)\,ds - \int_0^t\int_0^{t_k}\int_0^{t_{k-1}}\dots\int_0^{t_3}\int_{t_2}^\infty f(t_1)\,dt_1\,dt_2\dots dt_k,$$

whence

$$\frac{1}{(k-1)!}\int_0^t f(s)(t-s)^{k-1}\,ds - \frac{t^{k-1}}{(k-1)!}\int_0^\infty f(s)\,ds =$$

$$= -\int_0^t\int_0^{t_k}\int_0^{t_{k-1}}\dots\int_0^{t_3}\int_{t_2}^\infty f(t_1)\,dt_1\dots dt_k = o(t^{k-1}).$$

Applying this to x_{li} we have

$$(19) \qquad x_{li} = e^{\lambda_i t}\left[\bar c_{1i}\frac{t^{l-1}}{(l-1)!} + \bar c_{2i}\frac{t^{l-2}}{(l-2)!} + \dots + \bar c_{l-1,i}t + \bar c_{l,i} - \right.$$

$$-\int_0^t\int_0^{t_l}\dots\int_0^{t_3}\int_{t_2}^\infty f_{1i}e^{-\lambda_i t_1}\,dt_1\dots dt_l - \int_0^t\int_0^{t_l}\dots\int_0^{t_4}\int_{t_3}^\infty f_{2i}e^{-\lambda_i t_2}\,dt_2\dots dt_l -$$

$$\left. -\dots - \int_0^t\int_{t_l}^\infty f_{l-1,i}e^{-\lambda_i t_{l-1}}\,dt_{l-1}\,dt_l - \int_t^\infty f_{li}e^{-\lambda_i t_l}\,dt_l\right] \qquad (l = 1, \dots, a_i;\ i = 1, \dots, k)$$

which is of the form (18). This formula enables us to estimate $r(t)$ on subfamilies depending on less than n parameters.

3° *Modification of the normal form* of A. It is well known that there exists a normal form of A in the blocks J_i of which the 1's under the main diagonal are replaced by an arbitrary $\gamma = \text{const} > 0$ (as small as wanted). The two normal forms are connected by a linear transformation. Now our formulae will be modified as follows: in $\exp(J_i t)$ the variable t will be substituted everywhere by γt except the factor $\exp(\lambda_i t)$. The same is valid concerning $\exp(J_i t)\cdot c$ and $\exp[J_i(t-s)]\cdot f_i$. In the last expression $t-s$ has to be replaced by $\gamma(t-s)$ except again the factor $\exp[\lambda_i(t-s)]$. Instead of (13) we assume

$$(13') \qquad \varphi_i(t, r) = (1+\gamma t)^{(a-1)(1-m)} r^{\frac{\alpha_i}{\alpha}}\theta(t, r),$$

then (15) reads now as

$$(15') \qquad r(t) < Cr_0(1+\gamma t)^{a-1}e^{\alpha t}, \qquad t \geqq 0, \quad C = 2M.$$

4° *Comparison of some expressions*

(A) If $0 > \alpha = \alpha_1 > \alpha_2$, $a = a_1 \leqq a_2$ and $\alpha_1 - \alpha_2 > \gamma(a_2 - a_1)$, then

$$e^{\alpha_1 t}(1+\gamma t)^{a_1} > e^{\alpha_2 t}(1+\gamma t)^{a_2}, \qquad t \geqq 0,$$

(Of course, this holds a fortiori provided $a_1 > a_2$) or

$$e^{(\alpha_1-\alpha_2)t} > (1+\gamma t)^{a_2-a_1}, \quad (\alpha_1-\alpha_2)t > (a_2-a_1)\log(1+\gamma t).$$

Indeed

$$(\alpha_1-\alpha_2)t > (a_2-a_1)\gamma t > (a_2-a_1)\log(1+\gamma t) \quad (\text{since } x > \log(1+x)).$$

(B) Under the same conditions we have

$$Le^{\alpha_1 t}\int_0^t\int_0^{t_{a_1}}\dots\int_0^{t_3}\int_{t_2}^{\infty}\Phi_1(t_1, r(t_1))e^{-\alpha_1 t_1}\,dt_1\dots dt_{a_1} >$$

$$> e^{\alpha_2 t}\int_0^t\int_0^{t_{a_2}}\dots\int_0^{t_3}\int_{t_2}^{\infty}\Phi_2(t_1, r(t_1))^{-\alpha_2 t_1}\,dt_1\dots dt_{a_2}$$

where L is an appropriate constant. — Denoting these integrals by S_1 and S_2 we have by (13′) and (15′)

$$S_2 = \int_0^t\int_0^{t_{a_2}}\dots\int_0^{t_3}\int_{t_2}^{\infty}\Phi_1(t_1, r(t_1))r^{\frac{\alpha_2-\alpha_1}{\alpha}}e^{-\alpha_2 t_1}dt_1\dots dt_{a_2} <$$

$$< (Cr_0)^{\frac{\alpha_2-\alpha_1}{\alpha}}\int_0^t\int_0^{t_{a_2}}\dots\int_0^{t_{a_1+2}}\int_0^{t_{a_1+1}}S_1(t_1)(1+\gamma t_1)^{\frac{\alpha_2-\alpha_1}{\alpha}}dt_{a_1+1}\dots dt_{a_2} <$$

$$< Cr_0^{\frac{\alpha_2-\alpha_1}{\alpha}}S_1(t)\frac{(1+\gamma t)^l}{N\gamma^{a_2-a_1}}, \quad l = \frac{\alpha_2-\alpha_1}{\alpha}+a_2-a_1,\ N = \left(\frac{\alpha_2-\alpha_1}{\alpha}+1\right),$$

therefore

$$e^{\alpha_2 t}S_2 < KS_1 e^{\alpha_2 t}(1+\gamma t)^l < LS_1 e^{\alpha_1 t}$$

where K and L are suitable constants.

5° By statements 4° (A—B) if in (19) $\bar{c}_{11}=0$, then there is a constant $M>0$ with

(20)
$$r < \sum_{i=1}^{k}\sum_{l=1}^{a_i}|x_{li}| < Mne^{\alpha t}\Big[2r_0(1+\gamma t)^{a-2}+$$

$$+\gamma^{a-1}\int_0^t\int_0^{t_a}\dots\int_0^{t_3}\int_{t_2}^{\infty}\Phi_1(t_1, r(t_1))e^{-\alpha t_1}\,dt_1\dots dt_a+$$

$$+\gamma^{a-2}\int_0^t\int_0^{t_a}\dots\int_0^{t_4}\int_{t_3}^{\infty}\Phi_1 e^{-\alpha t_2}\,dt_2\dots dt_a+$$

$$+\dots+\gamma\int_0^t\int_{t_a}^{\infty}\Phi_1 e^{-\alpha t_{a-1}}\,dt_{a-1}dt_a+\int_t^{\infty}\Phi_1 e^{-\alpha t_a}dt_a\Big].$$

LEMMA 2. *If* $a>1$, $\bar{c}_{11}=0$ *and* $\alpha=\alpha_1>\alpha_i$, $i\neq 1$, *then along the corresponding paths*

$$r(t) < Dr_0(1+\gamma t)^{a-2}e^{\alpha t}, \qquad t \geqq 0, \quad r_0 \leqq r^* \leqq r_1 \tag{21}$$

where $D>0$ *is a suitable constant.*

PROOF. Denote the expression in square brackets in (20) by $p(t)$. The simple integral there involved is less than r_0. Merging (adding) r_0 in the polynomial $2r_0(1+\gamma t)^{a-2}$ we get $\bar{p}(t)$ instead of $p(t)$ which already does not involve a simple integral and $p<\bar{p}$. The function $p_1=\dot{\bar{p}}$ obtained by differentiation involves a polynomial of order $a-3$ and a simple integral. Merging its majorant γr_0 again in the polynomial part of $p_1(t)$ a function $\bar{p}_1(t)$ arises with $p_1<\bar{p}_1$. Continuing this process, eventually after $a-2$ differentiations and mergings we get $p_{a-2}<\bar{p}_{a-2}$ (in general $p_{l+1}=\dot{\bar{p}}_l$) where the polynomial part of $\bar{p}_{a-2}$ is a constant followed by a double integral as follows

$$p_{a-2}(t) < Ar_0\gamma^{a-2}+\gamma^{a-1}\int_0^t\int_\tau^\infty \Phi_1(s, r(s))e^{-\alpha s}\,ds\,d\tau = \bar{p}_{a-2}(t), \quad A = 2(a-2)!. \tag{22}$$

To prove (21) it is sufficient to show that

$$\bar{p}_{a-2}(t) < Br_0\gamma^{a-2}, \qquad t \geqq 0, \quad B = 3A. \tag{23}$$

Indeed, since

$$\begin{aligned}
p(0) &< \bar{p}(0) < 2r_0+r_0 < Br_0\\
p_1(0) &< \bar{p}_1(0) < 2r_0(a-2)\gamma+\gamma r_0 = r_0\gamma(2a-1) < Br_0\gamma\\
p_2(0) &< \bar{p}_2(0) < 2r_0(a-2)(a-3)\gamma^2+r_0\gamma^2 < r_0\gamma^2 3(a-2)! < Br_0\gamma^2\\
&\vdots\\
p_{a-2}(0) &< \bar{p}_{a-2}(0) < 2r_0(a-2)!\gamma^{a-2} < Br_0\gamma^{a-2}
\end{aligned}$$

successive integrations starting from (23) give

$$\begin{aligned}
p_{a-2} &< \bar{p}_{a-2} < Br_0\gamma^{a-2}\\
p_{a-3} &< \bar{p}_{a-3} < Br_0\gamma^{a-3}(1+\gamma t)\\
&\vdots\\
p &< \bar{p} < Br_0(1+\gamma t)^{a-2}
\end{aligned}$$

The inequality (23) holds for $t=0$ and for $t>0$ small enough. If it were not true for $t\geqq 0$, then there would exist a smallest $T>0$ with

$$\bar{p}_{a-2}(T) = Br_0\gamma^{a-2}, \quad \text{and} \quad \bar{p}_{a-2}(t) < Br_0\gamma^{a-2}, \qquad t < T. \tag{24}$$

The function $\bar{p}_{a-2}(t)$ is concave, implying

$$\bar{p}_{a-2}(t) < \bar{p}_{a-2}(0)+\dot{\bar{p}}_{a-2}(0)t < Ar_0\gamma^{a-2}+r_0\gamma^{a-1}t = Ar_0\gamma^{a-2}(1+\gamma t)$$

and since $\dot{\bar{p}}_{a-2}>0$

$$\bar{p}_{a-2}(t) < \bar{p}_{a-2}(1) < Ar_0\gamma^{a-2}(1+\gamma) < Br_0\gamma^{a-2} \qquad (0 \leqq t \leqq 1).$$

Therefore it can be supposed that $T>1$. Then, as already seen,

$$p(t) < \bar{p}(t) < Br_0(1+\gamma t)^{a-2}, \qquad t < T$$

$$r(t) < Br_0(1+\gamma t)^{a-2}e^{\alpha t}, \qquad t < T.$$

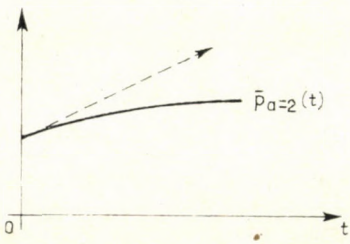


Fig. 1

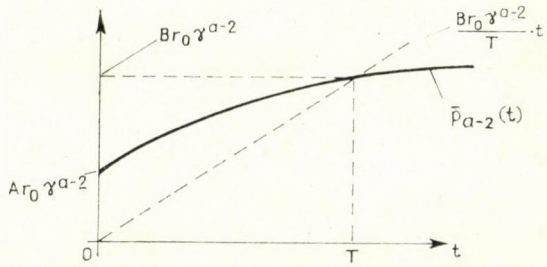


Fig. 2

On the other hand by the concavity of $\bar{p}_{a-2}(t)$

$$\bar{p}_{a-2}(t) < \frac{Br_0\gamma^{a-2}}{T}t < \frac{Br_0\gamma^{a-3}}{T}(1+\gamma t), \qquad t > T,$$

whence by integration

$$p_{a-3}(t) < \bar{p}_{a-3}(t) < \frac{Br_0\gamma^{a-4}}{T}\left[\frac{(1+\gamma t)^2}{2} - \frac{(1+\gamma T)^2}{2}\right] + Br_0\gamma^{a-3}(1+\gamma T) <$$

$$< \frac{Br_0\gamma^{a-4}}{T}(1+\gamma t)^2$$

(indeed this reduces to $\gamma^2T^2-1<(1+\gamma t)^2$). Finally, after similar steps we receive

$$(25) \qquad p(t) < \frac{Br_0}{\gamma T}(1+\gamma t)^{a-1}, \quad r(t) < \frac{Dr_0}{\gamma T}(1+\gamma t)^{a-1}e^{\alpha t}, \qquad t > T, \quad D = MnB.$$

Now (22) and (24) involve

$$\bar{p}_{a-2}(T) = Ar_0\gamma^{a-2} + \gamma^{a-1}\int_0^T s\Phi_1 e^{-\alpha s}\,ds + \gamma^{a-1}T\int_T^\infty \Phi_1 e^{-\alpha s}\,ds = Ar_0\gamma^{a-2} + S_1 + S_2.$$

Here

$$\Phi_1\big(t, r(t)\big) = (1+\gamma t)^{(a-1)(1-m)}r(t)\theta\big(t, r(t)\big)$$

implying

$$S_1 < \gamma^{a-1} \int_0^T s(1+\gamma s)^{(a-1)(1-m)} Dr_0(1+\gamma s)^{a-2} e^{\alpha s} \theta\big(s, Dr_0(1+\gamma s)^{a-2} e^{\alpha s}\big) e^{-\alpha s}\, ds <$$

$$< Dr_0 \gamma^{a-2} \int_0^T (1+\gamma s)^{a-1} \theta\big(s, Dr_0(1+\gamma s)^{a-2} e^{\alpha s}\big)\, ds,$$

namely $\gamma s < \gamma s + 1$ and $1+(a-1)(1-m)+a-2 \leqq a-1$ since $m \geqq 1$. The inequality

$$Ar_0\gamma^{a-2} + S_1 < Ar_0\gamma^{a-2} + Dr_0\gamma^{a-2} \int_0^T (1+\gamma s)^{a-1} \theta\big(s, Dr_0(1+\gamma s)^{a-2} e^{\alpha s}\big)\, ds <$$

$$< \frac{Br_0\gamma^{a-2}}{2}$$

will be satisfied provided one has

$$\int_0^\infty (1+\gamma s)^{a-1} \theta\big(s, Dr_0(1+\gamma s)^{a-2} e^{\alpha s}\big)\, ds < \frac{1}{6Mn}$$

and this inequlity holds, if r_0 is small enough. Furthermore by (25) the inequality

$$S_2 < \gamma^{a-1} T \int_T^\infty (1+\gamma s)^{(a-1)(1-m)} \frac{Dr_0}{\gamma T} (1+\gamma s)^{a-1} e^{\alpha s} \theta\left(s, \frac{Dr_0}{\gamma T}(1+\gamma s)^{a-1} e^{\alpha s}\right) e^{-as} ds <$$

$$< Dr_0\gamma^{a-2} \int_T^\infty (1+\gamma s)^{a-1} \theta\left(s, \frac{Dr_0}{\gamma}(1+\gamma s)^{a-1} e^{\alpha s}\right) ds < \frac{Br_0\gamma^{a-2}}{2} \qquad (T > 1)$$

is satisfied if

$$\int_0^\infty (1+\gamma s)^{a-1} \theta\left(s, \frac{Dr_0}{\gamma}(1+\gamma s)^{a-1} e^{\alpha s}\right) ds < \frac{1}{2Mn}$$

holds, and this is valid for r_0 small enough. Summarizing: there exists an r^* satisfying $0 < r^* \leqq r_1$ such that for $r_0 \leqq r^*$

$$\bar{p}_{a-2}(T) < Br_0\gamma^{a-2}$$

which contradicts (24) thus completing the proof of LEMMA 2.

Asymptotic form of the solution. Under the conditions of LEMMA 2, $\bar{c}_{11}=0$ implies (20) and this implies (see the estimate of S_1 above)

$$\int_t^\infty s|f_{1i}| e^{-\alpha_i s}\, ds < \infty, \quad \int_t^\infty f_{1i} e^{-\alpha_i s}\, ds = o\left(\frac{1}{t}\right), \quad \text{etc.}$$

and so the asymptotic form will change as follows. From (18_1), (18_2), (18_l) (for the sake of simplicity we put again $\gamma=1$*)

$$x_{1i}-\bar{\bar{y}}_{1i}=o(t^{-1}e^{\alpha_i t}),\quad \bar{\bar{y}}_{1i}=\bar{y}_{1i}=e^{\lambda_i t}\bar{c}_{1i},$$

$$x_{2i}=e^{\lambda_i t}\Big(\bar{c}_{1i}t+\bar{c}_{2i}-t\int_t^\infty f_{1i}e^{-\lambda_i s}\,ds+\int_t^\infty sf_{1i}e^{-\lambda_i s}\,ds-\int_t^\infty f_{2i}e^{-\lambda_i s}\,ds\Big).$$

$$\bar{\bar{c}}_{2i}=\bar{c}_{2i}-\int_0^\infty sf_{1i}e^{-\lambda_i s}\,ds=c_{2i}+\int_0^\infty f_{2i}e^{-\lambda_i s}\,ds-\int_0^\infty sf_{1i}e^{-\lambda_i s}\,ds,$$

$$x_{2i}-\bar{\bar{y}}_{2i}=o(e^{\alpha_i t}),\quad \bar{\bar{y}}_{2i}=e^{\lambda_i t}(\bar{c}_{1i}t+\bar{\bar{c}}_{2i})$$

$$\vdots$$

$$x_{li}=e^{\lambda_i t}\Big(\bar{c}_{1i}\frac{t^{l-1}}{(l-1)!}+\bar{\bar{c}}_{2i}\frac{t^{l-2}}{(l-2)!}+\ldots+\bar{\bar{c}}_{l-1,i}t+\bar{\bar{c}}_{li}+$$

(26
$$+\int_0^t\Big[f_{1i}\frac{(t-s)^{l-1}}{(l-1)!}+f_{2i}\frac{(t-s)^{l-2}}{(l-2)!}+\ldots+f_{l-1,i}(t-s)+f_{li}\Big]e^{-\lambda_i s}\,ds-$$

$$-\frac{t^{l-1}}{(l-1)!}\int_0^\infty f_{1i}e^{-\lambda_i s}\,ds-\frac{t^{l-2}}{(l-2)!}\int_0^\infty f_{2i}e^{-\lambda_i s}\,ds-$$

$$-\ldots-t\int_0^\infty f_{l-1,i}e^{-\lambda_i s}\,ds-\int_0^\infty f_{li}e^{-\lambda_i s}\,ds+$$

$$+\frac{t^{l-2}}{(l-1)!}\int_0^\infty (l-1)f_{1i}se^{-\lambda_i s}\,ds+\ldots+\int_0^\infty f_{l-1,i}se^{-\lambda_i s}\,ds$$

where

$$\bar{\bar{c}}_{ji}=\bar{c}_{ji}-\int_0^\infty f_{j-1,i}se^{-\lambda_i s}\,ds,\quad \bar{c}_{ji}=c_{ji}+\int_0^\infty f_{ji}e^{-\lambda_i s}\,ds\qquad (j=2,3,\ldots,l).$$

The sum of the three integrals in a column of (26) is of the form

$$S=\frac{1}{k!}\int_0^t\varphi(s)(t-s)^k\,ds-\frac{1}{k!}\int_0^\infty t^k\varphi(s)\,ds+\frac{1}{k!}\int_0^\infty k\varphi(s)st^{k-1}\,ds,\qquad \varphi=f_{li}e^{-\lambda_i s}.$$

According to the integral formula (18′) the sum of the first two terms can be rewritten as a $k+1$-fold integral, the third term as a k-fold integral

$$S=-\int_0^t\int_0^{t_{k+1}}\ldots\int_0^{t_3}\int_{t_2}^\infty\varphi(t_1)\,dt_1\ldots dt_k\,dt_{k+1}+\int_0^t\int_0^{t_k}\ldots\int_0^{t_3}\int_0^\infty t_1\varphi(t_1)\,dt_1\ldots dt_k.$$

* This is equivalent to a multiplication of the Φ_i's by a common constant.

This can be transformed into

$$S = \int_0^t \int_0^{t_{k+1}} \dots \int_0^{t_4} \left[-\int_0^{t_3} \int_{t_2}^{\infty} \varphi(t_1)\, dt_1\, dt_2 + \int_0^{\infty} t_1 \varphi(t_1)\, dt_1 \right] dt_3 \dots dt_k\, dt_{k+1}.$$

The expression in the square brackets is

$$-\int_0^{t_3} t_1 \varphi(t_1)\, dt_1 - t_3 \int_{t_3}^{\infty} \varphi(t_1)\, dt_1 + \int_0^{\infty} t_1 \varphi(t_1)\, dt_1 = \int_{t_3}^{\infty} (t_1 - t_3)\varphi(t_1)\, dt_1$$

and so

$$S = \int_0^t \int_0^{t_{k+1}} \dots \int_0^{t_4} \int_{t_3}^{\infty} (t_1 - t_3)\varphi(t_1)\, dt_1\, dt_3 \dots dt_k\, dt_{k+1} =$$

$$= \int_0^t \int_0^{t_k} \dots \int_0^{t_3} \int_{t_2}^{\infty} (t_1 - t_2)\varphi(t_1)\, dt_1 \dots d_{tk}$$

Hence

$$|S| < \int_0^t \int_0^{t_k} \dots \int_0^{t_3} \int_{t_2}^{\infty} t_1 |\varphi(t_1)|\, dt_1 \dots dt_k,$$

i.e. instead of a $k+1$-fold integral we get a k-fold (convergent) one, but the integrand $f_{li} e^{-\alpha_i s}$ will be replaced by $s f_{li} e^{-\alpha_i s}$.

This estimate may be applied to every column of (26). Thus we obtain an estimate of $|x_{li}|$ in which compared to (19) every j-fold integral ($j=2, 3, \dots, l$) will be replaced by a $j-1$-fold one and has an integrand multiplied by s. Therefore

$$x_{li} - \bar{\bar{y}}_{li} = o(t^{l-2} e^{\alpha_i t}), \quad \bar{\bar{y}}_{li} = e^{\lambda_i t} \left(\bar{c}_{1i} \frac{t^{l-1}}{(l-1)!} + \bar{\bar{c}}_{2i} \frac{t^{l-2}}{(l-2)!} + \dots + \bar{\bar{c}}_{li} \right)$$

or

$$x_{li} = e^{J_i t} (\bar{\bar{c}}_i + o(t^{-1}))$$

where $\bar{\bar{c}}_i = (\bar{c}_{1i}, \bar{\bar{c}}_{2i}, \dots, \bar{\bar{c}}_{a_i, i})$ $(i>1)$, but $\bar{\bar{c}}_1 = (0, \bar{\bar{c}}_{21}, \dots, \bar{\bar{c}}_{a,1})$.

COROLLARY OF LEMMA 2. Now, if $a>2$ and $\bar{c}_{11} = \bar{\bar{c}}_{21} = 0$, then Lemma 2 implies, instead of (20), that along the corresponding subfamily (for r_0 small enough)

$$r(t) < D_1 r_0 (1+\gamma t)^{a-3} e^{\alpha t}, \qquad t \geqq 0 \tag{27}$$

where D_1 is an appropriate constant. The process can be continued and finally if $\alpha = \alpha_1 > \alpha_i$ $(i \neq 1)$ and

$$\bar{c}_{11} = \bar{\bar{c}}_{21} = \dots = \tilde{c}_{a-1,1} = 0, \qquad \tilde{\bar{c}}_{a,1} \neq 0$$

where $\sim$ means $a-1$ bars, then

$$r(t) < \tilde{D} r_0 e^{\alpha t}, \qquad t \geqq 0 \tag{28}$$

with some constant $\tilde{D}$. — The new asymptotic form of x_{l1} involves only integrals of the form $\int_t^\infty$, the absolute value of which $< \int_t^\infty s^{a-1}\Phi_1 e^{-\alpha s}ds < \infty$.

LEMMA 3. *If* $\alpha=\alpha_1>\alpha_i$ $(i\neq 1)$ *and* $\alpha_2>\alpha_j$ $(j\neq 1, 2)$ *and* $(a_2-1)\gamma<\alpha-\alpha_2$, $\bar{c}_{11}=\bar{\bar{c}}_{21}=\ldots=\tilde{c}_{a-1,1}=\bar{\tilde{c}}_{a,1}=0$, *then along the corresponding subfamily*

$$r < C_1 r_0 (1+\gamma t)^{a_2-1} e^{\alpha_2 t}, \qquad t \geqq 0, \qquad r_0 \leqq r^{**} \leqq r_1$$

where C_1 *and* r^{**} *are appropriate constants.*

PROOF. Now we have

$$r(t) < K_1 r_0 (1+\gamma t)^{a_2-1} e^{\alpha_2 t} + K_2 e^{\alpha t} \int_t^\infty \gamma s^{a-1} \Phi_1(s, r(s)) e^{-\alpha s}\, ds,$$

where K_1, $K_2>0$ are suitable constants. If $r(t)e^{-\alpha t}=p(t)$, then

$$\gamma s^{a-1}\Phi_1 e^{-\alpha s} < \gamma s^{a-1}(1+\gamma s)^{(a-1)(1-m)} r e^{-\alpha s}\theta(s, r) < (1+\gamma s)^{a-1} p(s)\theta(s, pe^{\alpha s})$$

implying

$$p < K_1 r_0 (1+\gamma t)^{a_2-1} e^{(\alpha_2-\alpha)t} + K_2 \int_t^\infty p(s)(1+\gamma s)^{a-1}\theta(s, pe^{\alpha s})\, ds.$$

Here $pe^{\alpha s}=r<\tilde{B}r_0 e^{\alpha s}$. By the assumption $(a_2-1)\gamma<\alpha-\alpha_2$ the function $(1+\gamma t)^{a_2-1}e^{(\alpha_2-\alpha)t}$ is monotonic and tends to 0 as $t\to\infty$. Therefore GRONWALL's lemma (see e.g. [1]) may be applied and gives

$$r < LK_1 r_0 (1+\gamma t)^{a_2-1} e^{\alpha_2 t}, \qquad t \geqq 0, \quad L = e^M,$$

$$M = K_2 \int_0^\infty (1+\gamma s)^{a-1}\theta(s, \tilde{B}r_0 e^{\alpha s})\, ds.$$

The corresponding asymptotic form is obvious. Only the form of x_1 will change. Viz.

$$x_{l1} = e^{\alpha_1 t}\int_t^\infty s^{l-1}\Phi_1 e^{-\alpha_1 s}\, ds < e^{\alpha_2 t} I, \quad I = \int_t^\infty s^{l-1}\Phi_1 e^{-\alpha_2 s}\, ds \quad (<\infty).$$

Thus $x_{l1} = o(e^{\alpha_2 t})$, $l = 1, \ldots, a_1$. If here $\bar{\bar{c}}_{12} = 0$

$$r < D_1 r_0 (1+\gamma t)^{a_2-2} e^{\alpha_2 t}, \qquad t \geqq 0, \quad D_1 = \text{const}, \quad r_0 \leqq \hat{r} \leqq r_1,$$

etc. The proof proceeds on previous lines. — If also the last coefficient, say, $\hat{c}_{a_k,k}=0$, then $x(t)\equiv 0$.

The case of equal α*'s.* If for some $i\neq 1$ (for one or several i's)

$$(*) \qquad \alpha = \alpha_1 = \alpha_i, \quad a = a_1 > a_i,$$

then the following implications (denoted by the sign $\Rightarrow$) hold:

$$\{\bar{c}_{11} = 0\} \Rightarrow r < Cr_0(1+\gamma t)^{a-2}e^{\alpha t} \quad [1]$$

$$\vdots$$

$$\{\bar{c}_{11} = \ldots = \tilde{c}_{a_1-a_i,1} = 0\} \Rightarrow r < Cr_0(1+\gamma t)^{a_i-1}e^{\alpha t}$$

$$\{\bar{c}_{11} = \ldots = \tilde{c}_{a_1-a_i,1} = \bar{\bar{c}}_{a_1-a_2+1,1} = 0 \quad \text{and} \quad \bar{\bar{c}}_{1,i} = 0\} \Rightarrow r < Cr_0(1+\gamma t)^{a_i-2}e^{\alpha t}$$

etc. where $\bar{\bar{c}}_{1i}$ means *any* of the coefficients of the terms of order a_i-1 corresponding to the blocks J_i, i.e. from now on it is not sufficient the vanishing of some c_{l1}'s but some of the c_{l1}'s must also vanish.

The case of equal blocks. Finally if for some $i \neq 1$, $\alpha=\alpha_1=\alpha_i$, $a=a_1=a_i$, then

$$\{\bar{c}_{11} = \tilde{c}_{1i} = 0\} \Rightarrow r < Cr_0(1+\gamma t)^{a-2}e^{\alpha t}$$

and in general

$$\{\bar{c}_{j1} = \tilde{c}_{ji} = 0,\ j = 1, \ldots, l\} \Rightarrow r < Cr_0(1+\gamma t)^{a-l-1}e^{\alpha t}.$$

In this way every solution of (1) is assigned to a class with definite asymptotic properties. These classes are not empty. Viz., every class corresponds to a certain system of integral equations as (19) etc. and the existence of their solution can be easily shown by the Schauder fixed point theorem (see [1]). This solution is not necessarily unique unless some supplementary conditions assure this. — Thereby the proof of Theorem 1 is complete.

4. A new problem arises if for a certain $j \neq i$ simultaneously $\alpha_i=\alpha_j$, and $a_i=a_j \neq 1$ hold. The following example of a four dimensional system sheds light on the situation. Let

$$A = \begin{bmatrix} J_1 & \\ & J_2 \end{bmatrix}, \quad J_1 = J_2 = \begin{bmatrix} \lambda & 0 \\ 1 & \lambda \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad f = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}, \quad c = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix},$$

$$x_i = \begin{bmatrix} x_{1i} \\ x_{2i} \end{bmatrix}, \quad f_i = \begin{bmatrix} f_{1i} \\ f_{2i} \end{bmatrix}, \quad c_i = \begin{bmatrix} c_{1i} \\ c_{2i} \end{bmatrix}, \qquad i = 1, 2,$$

then

$$x_{1i} = e^{\lambda t}\Big(c_{1i} + \int_0^t f_{1i}e^{-\lambda s}\,ds\Big) \qquad (i = 1, 2)$$

$$x_{2i} = e^{\lambda t}\Big[c_{1i}t + c_{2i} + \int_0^t \big(f_{1i}(t-s) + f_{2i}\big)e^{-\lambda s}\,ds\Big].$$

If (12)—(13)—(14) hold, i.e.

(29) $|f_{ik}(t, x)| < r\theta(t, r)$, $\theta(t, 0) = 0$, $r(t) = |x(t)|$, $r < R_1$, $R_1 = \text{const}$; $i = 1,2$.

$$\int_t^\infty t\theta(t, Cte^{\lambda t})\,dt < \infty, \qquad C = \text{const} > 0,$$

[1] The constant C is not necessarily the same in all implications.

then by Lemma 1

$$r < Cr_0(1+t)e^{\alpha t}, \qquad t \geqq 0, \quad r_0 \leqq R^* \leqq R_1, \quad \alpha = \operatorname{Re}\lambda, \tag{30}$$

where C is an appropriate constant and $\int_0^\infty |f_{ik}|e^{-\alpha s}ds < r_0$ $(i, k=1, 2)$ implying

$$x_{1i} = e^{\lambda t}\left(\bar{c}_{1i} - \int_t^\infty f_{1i}e^{-\lambda s}\,ds\right)$$

$$x_{2i} = e^{\lambda t}\left(\bar{c}_{1i}t + \bar{c}_{2i} - t\int_t^\infty f_{1i}e^{-\lambda s}\,ds - \int_0^t sf_{1i}e^{-\lambda s}\,ds - \int_t^\infty f_{2i}e^{-\lambda s}\,ds\right) =$$

$$= e^{\lambda t}\left(\bar{c}_{1i}t + \bar{c}_{2i} - \int_0^t\int_\tau^\infty f_{1i}e^{-\lambda s}\,ds\,d\tau - \int_t^\infty f_{2i}e^{-\lambda s}\,ds\right)$$

$$\bar{c}_{ik} = c_{ik} + \int_0^\infty f_{ik}e^{-\lambda s}\,ds, \quad |\bar{c}_{ik}| < 2r_0 \qquad (i, k = 1, 2).$$

Suppose now $\bar{c}_{11}=0$. The question arises: is it true that

$$r_1(t) = |x_1(t)| < C_1 r_0 e^{\alpha t}, \qquad t \geqq 0, \quad C_1 = \text{const}, \tag{31}$$

i.e. is the norm of the subvector x_1 smaller than that of x?

The answer will be that assuming (29), (31) cannot be stated (at least in the framework of the present approach), while the modified condition

$$|f_{11}| < r_1\theta(t, r), \quad r_i = |x_i| \qquad (i = 1, 2) \tag{29'}$$

is sufficient to this end. An alternative way is to enlarge the exponent of t in the integral condition by 1, i.e. to assume

$$\int^\infty t^2\theta(t, Cte^{\lambda t})\,dt < \infty. \tag{29''}$$

PROOF. Assuming (29')

$$r_1 = |x_1| \leqq |x_{11}| + |x_{21}| < e^{\lambda t}\left(|\bar{c}_{21}| + \int_0^t\int_\tau^\infty |f_{11}|e^{-\lambda s}\,ds\,d\tau + \right.$$

$$\left. + \int_t^\infty |f_{11}|e^{-\lambda s}\,ds + \int_t^\infty |f_{21}|e^{-\lambda s}\,ds\right) <$$

$$< e^{\lambda t}\left(4r_0 + \int_0^t\int_\tau^\infty r_1(s)\theta(s, r(s))e^{-\alpha s}\,ds\,d\tau\right) \equiv e^{\alpha t}p(t).$$

It is sufficient to show that $p(t) < C_1 r_0$, $t \geqq 0$ holds. In the opposite case there exists a smallest $T>0$ with

$$p(T) = C_1 r_0, \quad \text{and} \quad p(t) < C_1 r_0, \quad r_1(t) < C_1 r_0 e^{\alpha t} \quad \text{for} \quad t < T \tag{32}$$

and on the lines of the proof of Lemma 2

$$p(t) < \frac{C_1 r_0}{T} t \quad \text{and} \quad r_1(t) < \frac{C_1 r_0}{T} te^{\alpha t} \quad \text{for} \quad t > T.$$

Since $0 < \dot{p} < r_0$

$$p(t) < \dot{p}(0)t + 4r_0 < r_0 t + 4r_0,$$

this involves

$$p(t) < p(1) < 5r_0 \quad \text{for} \quad 0 \leqq t \leqq 1.$$

Thus for $C_1 > 5$ we have $T \leqq 1$. Now

$$p(T) = 4r_0 + \int_0^T s r_1(s)\theta(s, r(s))e^{-\alpha s}\, ds + T\int_T^\infty r_1 \theta e^{-\alpha s}\, ds = 4r_0 + S_1 + S_2.$$

If (29) were valid, then r_1 should be replaced by r and a formal way would lead to the "estimate"

$$S_1 < Cr_0 \int_0^T s(1+s)\theta(s, Cr_0(1+s)e^{\alpha s})\, ds$$

and to a similar one concerning S_2 where the integrals are not convergent unless we assume (29″). — While assuming (29′)

$$S_1 < C_1 r_0 \int_0^T s e^{\alpha s}\theta(s, C_1\, r_0(1+s)e^{\alpha s})e^{-\alpha s}\, ds$$

and

$$4r_0 + S_1 < \frac{C_1 r_0}{4}$$

provided

$$\text{(33)} \qquad \int_0^\infty s\theta(s, C_1 r_0(1+s)e^{\alpha s})\, ds < \frac{C_1 - 16}{4C_1} = \frac{1}{4 \cdot 17}, \quad \text{if} \quad C_1 = 17.$$

Furthermore

$$S_2 < T\int_T^\infty \frac{C_1 r_0}{T} s e^{\alpha s}\theta(s, C_1 r_0(1+s)e^{\alpha s})e^{-\alpha s}\, ds =$$

$$= C_1 r_0 \int_T^\infty s\theta(s, C_1 r_0(1+s)e^{\alpha s})\, ds < \frac{C_1 r_0}{4} = \frac{17 r_0}{4}$$

provided (33) holds. However (33) holds good for r_0 small enough. Then $4r_0 + S_1 + S_2 < C_1 r_0$, $p(T) < C_1 r_0$ contradicting (32).

How can be generalized the result of this example, i.e. solved the problem raised at the beginning of this section?

LEMMA 4. *Suppose without restriction* $j = i+1$, *i.e. that the i-th and* $i+1$*-th blocks are identical or at least* $\alpha_i = \alpha_{i+1}$, $a_i = a_{i+1}$, $i \neq 1$ *(if* $\alpha_i = \alpha_{i+1}$ *for some i, then a situation equivalent to the present one will occur for certain family of solutions, namely when the polynomial parts of* x_i *and* x_{i+1} *are of equal order) and*

$$0 > \alpha = \alpha_1 > \alpha_2 > \ldots > \alpha_i = \alpha_{i+1} > \ldots > \alpha_k,$$

$$\bar{c}_{1l} = \bar{\bar{c}}_{2l} = \ldots = \tilde{c}_{a_i l} = 0, \qquad l = 1. \ldots, i-1; \; \bar{c}_{1i} = 0,$$

where the sign $\sim$ *substitutes bars in appropriate number and assume* (*beside* (12)-(14)) *either*

$$|f_{lj}| < (1+t)^{(a-1)(1-m)} \overline{r_i^{\alpha_j}}^{\alpha} \theta(t, r), \qquad (\alpha, \theta(t, r), m \text{ are as in 2—3}).$$

$$(j = i, i+1; \; l = 1, 2, \ldots, a_i)$$

or the above alternative. Then the following estimate holds

$$r < Cr_0(1+t)^{a_i - 1} e^{\alpha_i t} \quad \text{and} \quad r_i < C' r_0 (1+t)^{a_i - 2} e^{\alpha_i t}, \qquad t \geqq 0$$

where C *and* C' *are suitable constants. — The proof does not require new elements and may be omitted.*

5. Saddle-point cases. Paths tending to 0 as $t \to \infty$ will be called as "0 paths" and 0 is a saddle-point if there exist both 0 paths and not 0 paths too (i.e. paths first approaching and then tending away from 0). This can occur only when there are eigenvalues of opposite sign.

First it will be proved that along an 0 path $r(t)$ is eventually (i.e. for t large enough) monotonic.

5.1. The monotonicity of $r(t)$. Let us denote (only in this section) the eigenvalues and their real parts by λ_i and α_i $(i = 1, \ldots, n)$, respectively. Of course, among them there may be equals, too. In general, the Jordan form of the system is complex even if the original system was real. Thus $x = (x_1, \ldots, x_n)$ is a complex vector function of the real variable t $(t \geqq 0)$. — Suppose

$$\alpha_1 \geqq \alpha_2 \geqq \ldots \geqq \alpha_i > 0 > \alpha_{i+1} \geqq \ldots \geqq \alpha_n$$

and introduce the notations

$$r(t) = \sqrt{\sum_{l=1}^{n} |x_l|^2}, \quad u = \sqrt{\sum_{l=1}^{i} |x_l|^2}, \quad v = \sqrt{\sum_{l=i+1}^{n} |x_l|^2}, \quad \pm\frac{v}{r} = \cos\theta$$

supplemented by the assumption

$$0 \leqq \theta(0) \leqq \frac{\pi}{2}, \qquad \theta \in C_1.$$

Then $\pm\dfrac{u}{r} = \sin\theta$ and so $\theta(t)$ is defined obviously in a unique sense. θ can be regarded as the angle of the radius vector and of the space determined by the eigenvectors

belonging to the eigenvalues with negative real parts. — Each of the equations is of the form

$$\dot{x}_l = \delta_l x_{l-1} + \lambda_l x_l + f_l, \qquad l = 1, \ldots, n; \quad x_{-1} = 0$$

where δ_l is 0 or γ. This can be rewritten more concisely

$$\dot{z} = \delta u + \lambda z + f$$

where $u(t)$ and $z(t)$ are *scalar* complex functions of the real variable t.

We shall need the explicit form of $\dfrac{d(|z|^2)}{dt}$. This is

$$\frac{d(|z|^2)}{dt} = \frac{d(z\bar{z})}{dt} = \dot{z}\bar{z} + z\dot{\bar{z}} = (\delta u + \lambda z + f)\bar{z} + z(\delta\bar{u} + \bar{\lambda}\bar{z} + \bar{f}) =$$

$$= (\lambda + \bar{\lambda}) z\bar{z} + \delta(\bar{z}u + z\bar{u}) + (\bar{z}f + z\bar{f}) = 2\alpha|z|^2 + 2\,\mathrm{Re}\,[\bar{z}(\delta u + f)].$$

Applying this formula to r^2 and u^2 we have

$$2r\dot{r} = \frac{d(|x|^2)}{dt} = 2\sum_{l=1}^{n} [\alpha_l |x_l|^2 + \mathrm{Re}\,\bar{x}_l(\delta_l x_{l-1} + f_l)],$$

$$2u\dot{u} = 2\sum_{l=1}^{i} [\alpha_l |x_l|^2 + \mathrm{Re}\,\bar{x}_l(\delta_l x_{l-1} + f_l)],$$

$$r\dot{r} = F + R_1, \quad F = \sum_{i=1}^{n} \alpha_l |x_l|^2, \quad R_1 = \mathrm{Re}\sum_{l=1}^{n} \bar{x}_l(\delta_l x_{l-1} + f_l)$$

and in the same way we get from $\sin^2\theta = \dfrac{u^2}{r^2}$ that

$$\sin 2\theta \cdot \dot{\theta} = 2\frac{u\dot{u}r^2 - u^2 r\dot{r}}{r^4} =$$

$$= 2\frac{r^2 \sum_{l=1}^{i} [\alpha_l |x_l|^2 + \mathrm{Re}\,\bar{x}_l(\delta_l x_{l-1} + f_l)] - u^2 \sum_{l=1}^{n} [\alpha_l |x_l|^2 + \mathrm{Re}\,\bar{x}_l(\delta_l x_{l-1} + f_l)]}{r^4} = G + R_2,$$

where

$$G = 2\frac{\left(\sum_{l=1}^{i} \alpha_l |x_l|^2\right) \sum_{l=i+1}^{n} |x_l|^2 - \sum_{l=1}^{i} |x_l|^2 \sum_{l=i+1}^{n} \alpha_l |x_l|^2}{r^4} =$$

$$= 2\frac{\cos^2\theta \sum_{l=1}^{i} \alpha_l |x_l|^2 + \sin^2\theta \sum_{l=i+1}^{n} (-\alpha_l)|x_l|^2}{r^2},$$

$$R_2 = 2\frac{r^2\,\mathrm{Re}\sum_{l=1}^{i} \bar{x}_l(\delta_l x_{l-1} + f_l) - u^2\,\mathrm{Re}\sum_{l=1}^{n} \bar{x}_l(\delta_l x_{l-1} + f_l)}{r^4}.$$

Let us consider the following four domains of the (r, θ) plane

$$\left.\begin{array}{ll} D_1\colon & -\beta < \theta < \beta \quad \text{and} \quad \pi-\beta < \theta < \pi+\beta \\ D_2\colon & \dfrac{\pi}{2}-\beta < \theta < \dfrac{\pi}{2}+\beta \quad \text{and} \quad \dfrac{3\pi}{2}-\beta < \theta < 3\dfrac{\pi}{2}+\beta \\ D\colon & \beta \leqq \theta < \dfrac{\pi}{2}-\beta \\ D'\colon & C-(D_1 \cup D_2 \cup D) \end{array}\right| \begin{array}{l} r \leqq r_2 \leqq r_1 \\ \\ \beta < \dfrac{\pi}{4} \end{array}$$

where

$$C\colon \quad r \leqq r_2$$

LEMMA M. *Along an* 0 *path* $r(t)$ *is eventually monotonic provided* $L(r)\to 0$ *as* $r\to 0$. *Here* $L(r)=\max\left(H(r), K(r)\right)$ *and* $H(r)=\sup\limits_{t\geqq 0} \bar{\theta}(t, r)$, $K(r)=\sup\limits_{t\geqq 0} \theta(t, r)$.[2]

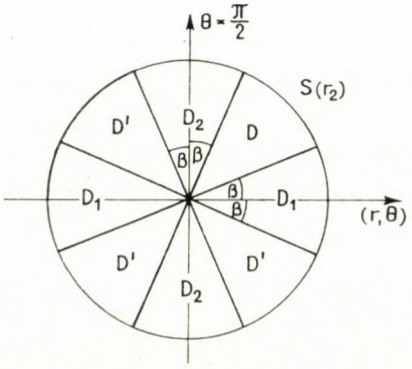


Fig. 3

PROOF. β and (to this) r_2 and γ can be chosen so that

1° a path having a point in D will reach in a finite time the circle $S(r_2)$ either in D or enters (the upper part of) D_2 and remains there, furthermore a path staying eventually in D_2 will reach here $S(r_2)$.

In fact

$$G > 2\frac{\cos^2\theta\cdot\alpha_i\cdot u^2+\sin^2\theta\cdot(-\alpha_{i+1})v^2}{r^2} = \frac{1}{2}\sin^2 2\theta\cdot(\alpha_i-\alpha_{i+1}) > 0.$$

Thus in D (for here we have $\sin 2\theta > \sin 2\beta$)

$$G > \frac{\alpha_i-\alpha_{i+1}}{2}\sin^2 2\beta > 0,$$

involving

$$\dot{\theta} > \frac{\alpha_i-\alpha_{i+1}}{2}\sin 2\beta + R_2', \quad R_2' = \frac{R_2}{\sin 2\beta}.$$

[2] See the definition of $\bar{\theta}(t, r)$ and $\theta(t, r)$ in section 5.2. under (37).

In 3^0 it will be proved that $|R_2'| < \frac{\alpha_i - \alpha_{i+1}}{4} \sin 2\beta$ and so in D

$$\dot{\theta} > \frac{\alpha_i - \alpha_{i+1}}{4} \sin 2\beta > 0$$

proving the first part of 1°. — On the other hand, in D_2

$$F > \alpha_i u^2 + \alpha_n v^2 = r^2(\alpha_i \sin^2\theta + \alpha_n \cos^2\theta) > \frac{1}{2}\alpha_i r^2$$

provided β is small anough, involving

$$\frac{\dot{r}}{r} > \frac{\alpha_i}{2} + \frac{R_1}{r^2};$$

Furthermore for $r \leqq r_2$ $\frac{|R_1|}{r^2} < \frac{\alpha_i}{4}$ (see 3°), therefore in D_2

$$\frac{\dot{r}}{r} > \frac{\alpha_i}{4} > 0,$$

which implies the second part of 1°. — A similar statement holds for D' and the corresponding (upper or lower) part of D_2.

In consequence, an 0 path can stay eventually only in D_1.

2° along a path $r(t)$ is monotonically decreasing as long as it stays in D_1. Namely

$$F < \alpha_1 u^2 + \alpha_{i+1} v^2 = r^2(\alpha_1 \sin^2\theta + \alpha_{i+1}\cos^2\theta)$$

and in D_1

$$F < r^2(\alpha_1 \sin^2\beta + \alpha_{i+1}\cos^2\beta) < \frac{1}{2} r^2 \alpha_{i+1} < 0, \quad \frac{\dot{r}}{r} < \frac{1}{2}\alpha_{i+1} + \frac{R_1}{r^2}$$

provided β is small enough and therefore in the same domain

$$\frac{\dot{r}}{r} < \frac{1}{4}\alpha_{i+1} < 0$$

since here $|R_1|/r^2 < |\alpha_{i+1}|/r^2$ (see 3°), which proves 2°.

3°. *Estimates of* R_1 *and* R_2. Obviously

$$|R_1| \leqq \sqrt{\sum_{l=1}^{n} |x_l|^2} \sqrt{\sum_{l=1}^{n} (\delta_l |x_{l-1}| + |f_l|)^2} < r(\gamma r + L(r) r)\sqrt{n} \quad \text{or} \quad \frac{|R_1|}{r^2} \leqq (\gamma + L(r))\sqrt{n}.$$

So for γ and r_2 small enough

$$\frac{|R_1|}{r^2} < \frac{1}{4} \min(\alpha_i, |\alpha_{i+1}|)$$

as stated above. — Analogously

$$|R_2| < 4(\gamma + L(r))\sqrt{n} \quad \text{and} \quad \frac{|R_2|}{\sin 2\beta} < \frac{\alpha_i - \alpha_{i+1}}{4} \sin 2\beta$$

provided γ and r_2 are sufficiently small.

Lemma M is the consequence of the assertions 1°—2°—3°.

5.2. Estimate of $r(t)$ along an 0 path. Suppose now

$$\alpha_1 \geqq \alpha_2 \geqq \ldots \geqq \alpha_i > 0 > \alpha_{i+1} \geqq \alpha_{i+2} \geqq \ldots \geqq \alpha_k \tag{34}$$

and that there are 0 paths the existence of which can be easily proved. — In the present case we use again instead of (3) that form of J where the 1's of the subdiagonal are replaced by $\gamma = \text{const} > 0$.

According (8) these paths satisfy

$$\bar{c}_j = c_j + \int_0^\infty e^{-J_j s} f_j \, ds = 0, \qquad (j = 1, 2, \ldots, i) \tag{35}$$

and so

$$x_j(t) = -\int_t^\infty e^{J_j(t-s)} f_j \, ds \qquad (j = 1, 2, \ldots, i). \tag{36}$$

Let the following conditions be imposed on f_{lj} for $r < r_1$

$$|f_{lj}| \leqq \begin{cases} r\bar{\theta}(t, r), & (j = 1, \ldots, i) \\ \Phi_j(t, r) = (1+\gamma t)^{(a-1)(1-m)} r^{\frac{\alpha_j}{\alpha}} \theta(t, r), & (l = 1, \ldots, a_j;\ j = i+1, \ldots, k) \end{cases} \tag{37}$$

$$\alpha = \alpha_{i+1}, \quad a = \max_{\alpha_j = \alpha} a_j, \quad m = \max_{j>i} \frac{\alpha_j}{\alpha} \geqq 1, \quad \gamma < |\alpha|,$$

where $\bar{\theta}(t, r)$ and $\theta(t, r)$ are continuous, nondecreasing in r, $\bar{\theta}(t, 0) = \theta(t, 0) = 0$ and

a) $L(r) \to 0$ as $r \to 0$ ($H(r)$, $K(r)$, $L(r)$ have the same meaning as in the proof of Lemma M) (38)

b) $\int^\infty s^{a-1} \theta(s, Cs^{a-1} e^{\alpha s}) \, ds < \infty, \qquad C = \text{const} > 0.$

LEMMA 5. *Along every 0 path the estimate*

$$r < Cr_0(1+\gamma t)^{a-1} e^{\alpha t}, \qquad t \geqq 0, \quad r_0 \leqq r^* \leqq r_2 \tag{39}$$

holds with appropriate constants C and r^*. (r_2 has been defined in the proof of Lemma M.) Moreover $|x_j| = o(r(t))$ as $t \to \infty$ for $j = 1, \ldots, i$.

PROOF. Introducing the notations

$$\varrho_1 = |x_1| + \ldots + |x_i|, \quad \varrho_2 = |x_{i+1}| + \ldots + |x_n|$$

we have by (18_1)

$$\varrho_1 < \sum_{j=1}^{i} e^{\alpha_j t} \int_t^{\infty} [1+\gamma(s-t)]^{a_j-1} r(s)\bar{\theta}\big(s, r(s)\big) e^{-\alpha_j s}\, ds = R_1(t)$$

$$\varrho_2 < Mr_0(1+\gamma t)^{a-1} e^{\alpha t} + \sum_{j=i+1}^{k} (1+\gamma t)^{a_j-1} e^{\alpha_j t} \int_0^t \Phi_j\big(s, r(s)\big) e^{-\alpha_j s}\, ds = R_2(t),$$

where $M>0$ is an appropriate constant. Obviously $r \leqq \varrho_1+\varrho_2 < R_1+R_2$. Now by Lemma M $r(t)$ is monotonic along an 0 path for $r \leqq r_2$, therefore

$$r(s)\bar{\theta}\big(s, r(s)\big) < r(t)\bar{\theta}\big(s, r(t)\big) < r(t)\bar{\theta}(s, r_0) \leqq r(t)H(r_0).$$

Given $0<\varepsilon<1$ there exists an r^* satisfying $0<r^* \leqq r_2$ such that for $r_0 \leqq r^*$ we have $H(r_0)<\varepsilon/N$ involving

$$R_1(t) < H(r_0)Nr(t) < \varepsilon r(t), \quad N = \sum_{j=1}^{i} e^{\alpha_j t} \int_t^{\infty} [1+\gamma(s-t)]^{a_j-1} e^{-\alpha_j s}\, ds = \text{const},$$

where N depends only upon the α's, a's and γ. From this we have

$$r(t) < \frac{1}{1-\varepsilon} R_2(t), \tag{40}$$

and in fact inequality (40) agrees with the one leading to the corresponding of (39), i.e. (15).

THEOREM 2. *Along an* 0 *path* $r(t)$ *tends to* 0 *at least in the order given by* (39). *Under appropriate conditions concerning some soefficients* c_j $(j = i+1, \ldots, k)$ *the* 0 *paths satisfy conditions like*

$$r < Cr_0(1+\gamma t)^{a_j-l-1} e^{\alpha_j t}, \qquad (j = i+1, \ldots, k; \quad l = 1, \ldots, a_j-1).$$

REFERENCES

[1] BIHARI, I. and ELBERT, Á.: Pertubation Theory of Three-dimensional Real Linear Autonomous Systems, *Period. Math. Hung.* (to appear).
[2] HARTMAN, PH. and WINTNER, A.: Asymptotic Integrations of Ordinary Nonlinear Diff. Equations, *Amer. J. Math.* **77** (1955), 692—724.
[3] HARTMAN, PH.: *Ordinary Differential Equations,* 1964, New York.
[4] BRAUER, F. and WONG, J. S. W.: On the Asymptotic Relationships between Solutions of Two Systems of Ordinary Diff. Equations. *Journal of Diff. Equations,* **6** (1969), 527—534.

Mathematical Institute of the Hungarian Academy of Seciences, Budapest

(Received June 20, 1972)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 269—271.

ON EXTENSIONS OF A THEOREM OF S. N. BERNSTEIN

by

K. SCHERER

The theorem of S. N. BERNSTEIN in question states that, given a sequence $\{M_n\}_{n=1}^{\infty}$ of n-dimensional subspaces M_n of a Banach space X satisfying

$$M_n \subset M_{n+1}, \tag{1}$$

and given an arbitrary sequence $E_1 \geqq E_2 \geqq \ldots > 0$, $\lim_{n\to\infty} E_n = 0$, there exists an element $f \in X$ such that $E_n(f) \equiv \inf_{q_n \in M_n} \|f - q_n\| = E_n$ for each $n = 1, 2, 3, \ldots$ (cf. [2, pp. 330], [5, pp. 264]).

Our first observation is that assumption "n-dimensional" can be weakened to "finite-dimensional". This is a consequence of the following simple

LEMMA. *Let $\{M_n\}_{n=1}^{\infty}$ be a sequence of proximinal subspaces of a linear normed space X. Then to each $f \in X$ and $A \geqq E_{n+1}(f)$ there exists $g \in M_{n+1}$ and η such that*

$$E_n(f + \eta g) = A. \tag{2}$$

PROOF. Let g_0 be the element of best approximation to f from M_{n+1}. Then one has

$$E_{n+1}(f) = E_{n+1}(f - g_0) \leqq E_n(f - g_0) \leqq \|f - g_0\| = E_{n+1}(f).$$

Thus in case $E_n(g_0) > 0$ equation (2) follows in view of $\lim_{|\sigma|\to\infty} E_n(f + \sigma g) = \infty$ and the continuity of this function with respect to σ. In case $E_n(g_0) = 0$ one has $E_n(f) = E_{n+1}(f)$ so that one may take any $g \in M_{n+1}$ with $E_n(g) > 0$ in order to establish relation (2).

With the help of this lemma an extension of BERNSTEIN's theorem may be established as in [2]. One first proves the assertion for nonincreasing sequences having only a finite number of nonvanishing terms and then for arbitrary ones tending to zero, using the fact that the unit sphere of a finite-dimensional subspace is compact.

A good illustration of this may be given in the case of spline approximation. Let $S_p(n, \Delta)$ be the class of spline function $s(x)$ which are polynomials of degree m on each segment (x_i, x_{i+1}) of the partition $\Delta = \{a = x_0 < x_1 < x_n = b\}$ of the interval $[a, b]$, and for which $s^{m-1}(x)$ belongs to $C[a, b]$. (In case $m = 0$ we define $S_p(0, \Delta)$ to be the class of step functions with jumps at the knots of Δ). Since $S_p(m, \Delta)$ is finite dimensional an immediate consequence is

COROLLARY 1. Let $\{\Delta_n\}_{n=1}^{\infty}$ be a nested sequence of partitions of the interval $[a, b]$, i.e., $\Delta_n \subset \Delta_{n+1}$, and let $E_n(f) = \inf_{s \in S_p(m, \Delta_n)} \sup_{a \leqq x \leqq b} |f(x) - s(x)|$ be the best approximation of $f \in C[a, b]$ by the elements of $S_p(m, \Delta_n)$. Then, given an arbitrary sequence $E_1 \geqq E_2 \geqq \ldots > 0$, $\lim_{n \to \infty} E_n = 0$, there exists a function $f \in C[a, b]$ such that $E_n(f) = E_n$ for each $n = 1, 2, 3, \ldots$.

Hence the degree of spline approximation may be arbitrarily high in such cases. An actual example of this is the best approximation by Walsh polynomials where smoothness properties implying arbitrarily high degree of approximation are known (cf. BUTZER—WAGNER [1]). In contrast to this one has the fact that the best approximation $E_n(f)$ to $f \in C[a, b]$ from $S_p(m, \Delta_N)$ with equidistant partition $\Delta_N = \{x_i : x_{i+1} - x_i = (b-a)/N\}$ cannot satisfy $E_N(f) = o(N^{-m-1})$, $N \to \infty$, unless f is a polynomial of degree m on $[a, b]$, i.e., $E_N(f)$ is zero for all N (see e.g. [3]). The reason is that the relation $S_p(m, \Delta_N) \subset S_p(m, \Delta_{N+1})$ does not hold for these subspaces of splines. This shows that in general condition (1) cannot be weakened.

Another question is whether the assumption "finite-dimensional" can still be weakened. TJUREMSKIH [6] has shown that it may be dropped if X is a Hilbert space. We give a partial answer of this problem for reflexive spaces.

THEOREM. *Let $\{M_t\}_{0 < t \leqq 1}$ be a family of closed linear subspaces of a reflexive Banach space X satisfying the strict inclusion relation $M_t \subset M_s$ for $s < t$ and $M_t = \bigcap_{s < t} M_s$. Then, given any nonincreasing function $E(t)$ on $(0, 1]$, $\lim_{t \to 0+} E(t) = 0$, there exists an element $f \in X$ such that $E_t(f) \equiv \inf_{g \in M_t} \|f - g\| \leqq E(t)$ for $t \in (0, 1]$.*

PROOF. Since X is reflexive every subspace M_t is proximinal. Hence by the foregoing for each fixed natural number n there exists $f_n \in X$ such that $E_{k2^{-n}}(f_n) = E(k2^{-n})$, $1 \leqq k \leqq 2^n$. In addition one may assume $\|f_n\| = E(1)$. Then by the reflexivity of X there exists a subsequence $\{f_{n_k}\}_{k=1}^{\infty}$ converging weakly to an element $f \in X$. Furthermore, for a fixed pair l, m with $1 \leqq l \leqq 2^m$ one can find (Hahn—Banach theorem) a continuous linear functional f^* such that $\|f^*\| = 1$, $f^*(g) = 0$ for $g \in M_{l2^{-m}}$ and $f^*(f) = E_{l2^{-m}}(f)$. Denoting by g_k an element of best approximation to f_{n_k} from $M_{l2^{-m}}$ it follows

$$E_{l2^{-m}}(f) = \lim_{k \to \infty} f^*(f_{n_k}) \leqq \overline{\lim_{k \to \infty}} f^*(f_{n_k} - g_k) \leqq \overline{\lim_{k \to \infty}} E_{l2^{-m}}(f_{n_k}) = E(l2^{-m}).$$

Now let $\{r_j = l_j 2^{-n_j}\}_{j=1}^{\infty}$ be a sequence converging to $t \in (0, 1]$ from the left and let $\{h_{j_k}\}_{k=1}^{\infty}$ be a sequence of elements of best approximation to f from M_{r_j}. Since this sequence is bounded there exists a subsequence $\{h_{j_k}\}_{k=1}^{\infty}$ converging weakly to an element h_t where $h_t \in \bigcap_k M_{r_{j_k}} = M_t$. Then let g^* be a continuous linear functional such that $\|g^*\| = 1$ and $g^*(f - h_t) = \|f - h_t\|$. It follows

$$E_t(f) \leqq \|f - h_t\| = g^*(f - h_t) \leqq \overline{\lim_{k \to \infty}} g^*(f - h_{j_k}) =$$

$$\leqq \overline{\lim_{k \to \infty}} E_{r_{j_k}}(f) \leqq \overline{\lim_{k \to \infty}} E(r_{j_k}) \leqq E(t),$$

proving the theorem.

The same inequality holds for sequences $\{M_n\}_{n=1}^{\infty}$ of subspaces of X satisfying (1) and nonincreasing sequences $\{E_n\}_{n=1}^{\infty}$. Together with a result of TJUREMSKIH [7] establishing the opposite inequality this yields

COROLLARY 2. Let $\{M_n\}_{n=1}^{\infty}$ be a sequence of closed linear subspaces of a reflexive Banach space X satisfying (1) with $\bigcup_{n=1}^{\infty} M_n$ being dense in X. Then there exist elements $f \in X$ such .hat $E_n(f)$ decreases arbitrarily slowly or arbitrarily rapidly to zero.

Another consequence is the following analogue of BERNSTEIN's theorem in the case of best approximation by entire functions of exponential type.

COROLLARY 3. Let M_t be the subspace B_t^p of entire functions of exponential type $\leqq t^{-1}$, $t \in (0, 1]$, whose restrictions on the real axis belong to the Lebesgue space $L_p(-\infty, \infty)$, $1 < p < \infty$. Then, given any nonincreasing function $E(t)$ on $(0, 1]$, $\lim_{t \to 0+} E(t) = 0$ there exists an element $f \in L_p(-\infty, \infty)$ such that

$$E_t^p(f) \equiv \inf_{q \in B_t^p} \|f - q\| \leqq E(t) \quad \text{for} \quad t \in (0, 1].$$

There are sharper results with equality holding in Corollary 2 for $p = 2$ and the space $C(-\infty, \infty)$ of functions which are bounded and uniformly continuous on the whole real axis, see e.g. IBRAGIMOV [4, p. 293]. However, in the latter case the proof given there seems to yield only the above inequality since it uses not the uniform convergence on $(-\infty, \infty)$.

REFERENCES

[1] BUTZER, P. L.—WAGNER, H. J.: Approximation by Walsh polynomials and the concept of a derivative. In: *Proceedings of the Symposium on Applications of Walsh Functions*, Washington 1972 (in print).

[2] DAVIS, P. J.: *Interpolation and Approximation*, Blaisdell Publ. Co., New York 1963.

[3] GAIER. D.: Saturation by Spline-Approximation and Quadratur, *Numer. Math.* **16**. (1970), 129—140.

[4] IBRAGIMOV, I. I.: *Extremal Properties of Entire Functions of Finite Order* (Russ.), Izdat. Akad. Nauk Azerbeĭdžan. SSR Baku 1962.

[5] SINGER, I.: *Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces*, Springer, Berlin—Heidelberg—New York 1970.

[6] TJUREMSKIH, I. S.: The (B)-property of Hilbert spaces (Russ.), *Kalinin. Gos. Ped. Inst. Učen. Zap.* **39** (1964), 53—64.

[7] TJUREMSKIH, I. S.: A certain theorem of S. N. Bernstein (Russ.), *Kalinin. Gos. Ped. Inst. Učen. Zap.* **52** (1967), 123—129.

Lehrstuhl A für Mathematik, Technological University of Aachen, 51 Aachen, Templergraben 55

(Received October 17, 1972)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 273—284.

ON A THEOREM OF LYAPUNOFF

by

L. GERENCSÉR

In 1947 LYAPUNOFF proved the following theorem:

If $\boldsymbol{\mu}$ is a non-atomic finite-dimensional vector-measure then the range of $\boldsymbol{\mu}$ is a closed convex set in the n-dimensional Eucledian-space.

He also gave examples that the theorem is not true for l_p — valued measures. $(1 \leqq p < \infty)$. Later P. R. HALMOS and D. BLACKWELL gave simpler proofs of the theorem. The latter author also showed an application of this result for game theory. Recently it is used again in control theory. (See OLECH.) An original approach to the problem is given in HARTMANN's paper.

We consider here a related problem for stochastic set functions. We repeat here the notions that were defined by A. PRÉKOPA in his pioneer work [10]. A stochastic set function can be described by the set of distribution functions. The question is wether this set is closed or not. We shall show for important examples that both cases are possible. While investigating this problem I found and proved Lyapunoff's result. It seems that my proof is simpler than other proofs, though it has one point in common with BLACKWELL's proof. (Theorem 1.)

At this place I express my thanks to Prof. A. PRÉKOPA for his help and for formulating me the problem. I also thank here for the advices of Prof. W. A. WOYCYNSKI, who helped me in finding the original papers.

1. Preliminary results

Let E^n denote the n-dimensional Eucledian space. Let be given an E^n valued measure $\boldsymbol{\mu}$ and consider its range

$$G = \{\boldsymbol{\mu}(A) : A \in \mathscr{R}\} \tag{1.1}$$

where $\mathscr{R}$ is the σ-algebra of the measure space. The entire space will be denoted by H.

We shall consider here only non-atomic measures. The result that G is closed for all kind of measures can be proved by decomposition into purely atomic and non-atomic components. However we omit details here, and turn to the statements and their proofs which lead to the final result.

THEOREM 1. *Let the non-atomic probability measures $\mu_1, \ldots, \mu_n$ be given on a σ-algebra $\mathscr{R}$ of subsets of a set H. Then there exists a subalgebra $\mathscr{R}_0 \subset \mathscr{R}$, containing H, on which the measures μ_i are still non-atomic and indentical with each other.*

In the proof we need the following

LEMMA: *Let μ be a real-valued non-atomic measure defined on a σ-algebra $\mathcal{R}$ of subsets of a set H. If μ takes on values a and b, then it takes on every value between a and b too.* (see HAHN.)

PROOF OF THEOREM 1: We give a σ-algebra $\mathcal{R}_0 \subset \mathcal{R}$ containing H, on which $\mu_1 \equiv \mu_2$ and the measures $\mu_1, \ldots, \mu_n$ remain non-atomic. Then we can proceed by induction.

Let us introduce the following notations: $\eta = \mu_1 - \mu_2$ and $\mu = \mu_1 + \ldots \mu_n$. We shall give a sequence of decompositions of H: $H = A^k_{i_1 \ldots i_k}$ ($i_k = 0$ or 1) such that

$$\eta(A^k_{i_1 \ldots i_k}) = 0 \tag{1.2}$$

$$\mu(A^k_{i_1 \ldots i_k}) \leqq n \left(\frac{2}{3}\right)^k \tag{1.3}$$

The sets $A^k_{i_1 \ldots i_k}$ will generate the desired $\mathcal{R}_0$ σ-algebra.

At the first step we proceed as follows. Consider a subdivision of H:

$$H = H_1 \cup H_2 \cup H_3$$

where $\mu(H_i) = \frac{1}{3}\mu(H) = \frac{1}{3}n$ ($i = 1, 2, 3$). μ is non-atomic (see HAHN), therefore such a subdivision is possible according to the Lemma. The signs of the values $\eta(H_i)$ can not all be the same, since their sum $\eta(H)$ equals 0. Say

$$\eta(H_1) \geqq 0, \quad \eta(H_2) \leqq 0 \quad \text{and} \quad \eta(H_3) \geqq 0.$$

For the same reason as before we have $\eta(H_1) + \eta(H_2) \leqq 0$ which can also be written as

$$\eta(H_2) \leqq -\eta(H_1) \leqq 0 \tag{1.4}$$

Applying the Lemma to the measure-space built from subsets of H_2, we get a measurable set H_2':

$$H_2' \subset H_2, \quad \text{for which} \quad \eta(H_2') = -\eta(H_1). \tag{1.5}$$

Denote the set $H_2 \setminus H_2'$ by H_2'' and consider the sets

$$A^1_0 = H_1 \cup H_2' \quad \text{and} \quad A^1_1 = H_2'' \cup H_3$$

The decomposition $H = A^1_0 \cup A^1_1$ is such that

$$\eta(A^1_0) = \eta(A^1_1) = 0$$

and

$$\mu(A^1_0) = \mu(H_1) + \mu(H_2') \leqq \mu(H_1) + \mu(H_2) = \frac{2}{3}n$$

Similarly $\mu(A^1_1) \leqq \frac{2}{3}n$.

Since $\eta(A^1_i) = 0$ ($i = 0, 1$) we can repeat the previous procedure taking A^1_i instead of H and so on. In the k-th step we get the sets $A^k_{i_1 \ldots i_k}$ ($i_k = 0$ or 1) for which the relations (1.2) and (1.3) hold.

The sets $A^k_{i_1 \dots i_k}$ $(k=1, 2 \dots)$ generate an algebra $\mathscr{R}_1$, on which η is identically 0, that is to say μ_1 equals μ_2 identically. If $\mathscr{R}_0$ denotes the σ-algebra generated by $\mathscr{R}_1$ we have $\mu_1 \equiv \mu_2$ on $\mathscr{R}_0$, too. The only thing that we still have to prove is that the $\mu_j - s$ are non-atomic on $\mathscr{R}_0$.

Let a B taken from $\mathscr{R}$, for which $\mu_j(B)=m>0$. B is the union of the sets $B \cap A^k_{i_1 \dots i_k}$ (k fixed) and every set in this decomposition has a μ_j-measure less than $n\left(\frac{2}{3}\right)^k$ which is less than m, if k is large enough. This means, that B is not a μ_j atom, and the theorem is proved.

THEOREM 2. *Let $\mu=(\mu_1, \dots, \mu_n)$ be a non-atomic vector-measure. Then its range*

$$G = \{\mu(A) : A \in \mathscr{R}\}$$

is a cone. (By a cone we mean a set of vectors E such that $v \in E$ implies $\lambda_v \in E$ for $C \leqq \lambda \leqq 1$.)

PROOF: Consider a set H_1 and prove that $c \cdot \boldsymbol{\mu}(H_1)$ belongs to E if $0 \leqq c \leqq 1$. Clearly it is enough to consider the case when $H_1 = H$.

Apply the previous theorem to the $2n$ measures

$$\frac{\mu_1^+}{\mu_1^+(H)}, \quad \frac{\mu_1^-}{\mu_1^-(H)}, \dots, \frac{\mu_n^+}{\mu_n^+(H)}, \quad \frac{\mu_n^-}{\mu_n^-(H)} \tag{1.6}$$

where $\mu_i^{\pm}$ denote the positiv and the negativ variation of the measure μ_i respectively. If say $\mu_1^+(H)$ equals 0, then the measure μ_1^+ is identically 0 and we shall omit it.

The measures in (1.6) are identical with a non-atomic probability measure η on some σ-algebra $\mathscr{R}_0$. According to the intermediate-value theorem for every c between 0 and 1 there is a set A with $\eta(A)=c$. Write out this in the way:

$$\mu_i^{\pm}(A) = c\mu_i^{\pm}(H) \tag{1.7}$$

Adding the $\pm$ terms for every i, we get

$$\mu_i(A) = c\mu_i(H)$$

that is

$$\boldsymbol{\mu}(A) = c \cdot \boldsymbol{\mu}(H) \tag{1.8}$$

which proves the theorem.

Remark: Before ending this section we also give a different interpretation to the result of Theorem 1. Suppose we are given a non-atomic probability measure P and integrable random variables $\xi_1, \dots \xi_n$. Apply Theorem 1 to the measures

$$\begin{aligned} \mu_i^+(A) &= \int_A \xi_i^+ \, d\mathsf{P} \Big/ \int_H \xi_i^+ \, d\mathsf{P} \\ \mu_i^-(A) &= \int_A \xi_i^- \, d\mathsf{P} \Big/ \int_H \xi_i^- \, d\mathsf{P} \end{aligned} \tag{1.9}$$

and the measure P itself. For a set $A \in \mathscr{R}_0$ we have:

$$\mu_i^{\pm}(A) = \frac{\int\limits_A \xi_i^{\pm}\, d\mathsf{P}}{\int\limits_H \xi_i^{\pm}\, d\mathsf{P}} = \mathsf{P}(A) \tag{1.10}$$

which gives after multiplying and adding the $\pm$ terms:

$$\int\limits_A \xi_i\, d\mathsf{P} = \mathsf{P}(A) \int\limits_H \xi_i\, d\mathsf{P}.$$

Divide by $\mathsf{P}(A)$, then we have

$$\mathsf{M}(\xi_i|A) = \int\limits_H \xi_i\, d\mathsf{P} = \mathsf{M}(\xi_i).$$

This is true for all sets $A \in \mathscr{R}_0$, hence

$$\mathsf{M}(\xi_i|\mathscr{R}_0) = \mathsf{M}(\xi_i) = \text{const.} \tag{1.11}$$

If we take a random variable η, which expands the σ-algebra $\mathscr{R}_0$, then we can also write

$$\mathsf{M}(\xi_i|\eta) = \mathsf{M}(\xi_i) \quad \text{for} \quad i = 1, 2, \ldots . \tag{1.12}$$

The most obvious way to fulfill this equality is to take a random variable η, which is independent of the $\xi_i - s$. But this is not always possible. Here is the power of Theorem 1.

2. Lyapunoff's theorem

We can now give the proof of Lyapunoff's results.

THEOREM 3. *Let $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ be a non-atomic vector measure on a σ-algebra $\mathscr{R}$ of subsets of a set H. Then the range of $\boldsymbol{\mu}$:*

$$E = \{\boldsymbol{\mu}(A) : A \in \mathscr{R}\}$$

is a closed convex set in E^n.

PROOF. We shall find integral representations for the set E, from which the assertions will follow.

First consider the step functions:

$$h(x) = c_k \quad \text{if} \quad x \in H_k \tag{2.1}$$

where the measurable sets H_k constitute a decomposition of the set H, and the numbers c_k lie between 0 and 1. The family of all these $h(x) - s$ will be denoted by S.

First representation of E:

$$E = \left\{\int\limits_H h(x)\, d\boldsymbol{\mu} : h(x) \in S\right\} \tag{2.2}$$

PROOF: Clearly E is contained in the right-hand side set. Thus we have only to prove that for any $h(x) \in S$ there exists a set A, such that

$$\int_H h(x)\, d\boldsymbol{\mu} = \boldsymbol{\mu}(A) \tag{2.3}$$

Apply Theorem 2. for the sets H_k, and numbers c_k. We find sets A_k, for which the equalities

$$\boldsymbol{\mu}(A_k) = c_k \boldsymbol{\mu}(H_k) \tag{2.4}$$

hold. Sum up these for $k=1, 2, \ldots, p$ to get

$$\boldsymbol{\mu}(A) = \int_H h(x)\, d\boldsymbol{\mu}$$

where

$$A = \bigcup_{k=1}^{p} A_k \quad \text{q.e.d.}$$

Since S is a convex set of functions the integrals of these functions, that is the vectors of E constitute a convex set, too. To prove that E is closed we give another representation of E.

Denote by T the set of all measurable real functions $f(x)$ defined on H, for which $0 \leqq f(x) \leqq 1$ for all $x \in H$.

Second representation of E

$$E = \left\{ \int_H f(x)\, d\boldsymbol{\mu} : f(x) \in T \right\}. \tag{2.5}$$

PROOF: Suppose that for some function

$$f \in T \quad \text{we have} \quad \mathbf{v} = \int_H f(x)\, d\boldsymbol{\mu} \notin E. \tag{2.6}$$

Denote the sets $\{x : f(x)=1\}$, $\{x : f(x)=0\}$ and $\{x : 0<f(x)<1\}$ by A, B and C respectively Obviously

$$\mathbf{v} = \int_H f(x)\, d\boldsymbol{\mu} = \int_A + \int_B + \int_C = \boldsymbol{\mu}(A) + \int_C f(x)\, d\boldsymbol{\mu}.$$

There exists no such subset D of C, for which

$$\int_C f(x)\, d\boldsymbol{\mu} = \boldsymbol{\mu}(D) \tag{2.7}$$

Otherwise we would obtain that

$$\mathbf{v} = \boldsymbol{\mu}(A) + \boldsymbol{\mu}(D) = \boldsymbol{\mu}(A \cup D) \tag{2.8}$$

a contradiction.

Thus the set C may play the same role as H did before with the difference that the function $f(x)$ in question is nowhere 0 or 1. To simplify notations we suppose, that $C=H$.

Next we choose a hyperplane, which separates $\mathbf{v}$ from C. Since E is convex and $\mathbf{v} \notin E$, this is possible, i.e. there are numbers $d_1, \ldots, d_n$ such that

$$\sum_{i=1}^{n} d_i \int_H f \, d\mu_i \geqq \sum_{i=1}^{n} d_i \int_H h \, d\mu_i \tag{2.9}$$

holds for every $h \in S$. Here we used the first representation of E.

We introduce the notation θ for the measure $\Sigma d_i \mu_i$:

$$\theta = \sum d_i \mu_i \tag{2.10}$$

Then we can write (2.9) as

$$\int_H f \, d\theta \geqq \int_H h \, d\theta. \tag{2.11}$$

Since the values of $f(x)$ lie between 0 and 1, the left-hand side is at most $\theta^+(H)$, the positive variation of θ on the set H. The L.U.B. of the right hand side is also $\theta^+(H)$, thus

$$\int_H f \, d\theta = \theta^+(H). \tag{2.12}$$

Denote the Hahn-decomposition of the space H by $H = H^+ \cup H^-$. An easy calculation shows, that equation (2.12) implies that f is equal to the characteristic function of the set H^+ except perhaps a set of measure 0. Measure $\bar{\theta}$ means the total variation of θ. But $f(x)$ is nowhere 0 or 1, hence $\bar{\theta}(H) = 0$, which implies that

$$\theta(A) = 0 \quad \text{for all} \quad A \in \mathscr{R}. \tag{2.13}$$

This means a linear relation between the measures μ_i:

$$d_1 \mu_1 + \ldots + d_n \mu_n \equiv 0. \tag{2.14}$$

Expressing one of these measures through the rest we can proceed by induction.

Using the just proved representation we can prove that E is closed. Suppose, that $\mathbf{m} = (m_1, \ldots, m_n)$ is a point of accumulation of E:

$$\int_H h_k(x) \, d\mu_i \to m_i \tag{2.14}$$

for some sequence of functions $h_n(x) \in S$. The level sets of these functions generate a σ-algebra, which we denote by $\mathscr{R}_1$. Consider the measure

$$\eta = \bar{\mu}_1 + \ldots + \bar{\mu}_n \tag{2.15}$$

restricted to this algebra.

The measures μ_i are absolutely continuous with respect to the measure η on $\mathscr{R}_1$. Writing $g_i(x)$ for the Radon-Nykodim derivates we get from (2.14)

$$\int_H h_k(x) g_i(x) \, d\eta \to m_i \tag{2.16}$$

The space $L^1(H, \mathcal{R}_1, \eta)$ is a separable Banach-space. Its dual-space is $L^\infty(H, \mathcal{R}_1, \eta)$.

The functions $h_k(x)$ form a bounded set in $L^\infty(H, \mathcal{R}_1, \eta)$ so using diagonal procedure we easily get a weakly convergent subsequence:

$$h_{k_s} \to f.$$

It is also easy to see that $0 \leqq f(x) \leqq 1$, that is $f \in T$. The weak convergence implies

$$\int_H h_{k_s}(x) g_i(x)\, d\eta \to \int_H f(x) g_i(x)\, d\eta \tag{2.19}$$

which compared with (2.14) gives:

$$m_i = \int_H f(x) g_i(x)\, d\eta \tag{2.20}$$

or

$$\mathbf{m} = \int_H f(x)\, d\mu \tag{2.21}$$

the desired result.

3. Counter-examples for infinite-dimensional spaces

In this paragraph we show, that Theorem 3. is not valid for infinite dimensional measures. We give the example on the σ-algebra of Borel-sets of the interval [0, 1].

Consider a sequence of measures given on $\mathfrak{B}$: $\mu_1, \mu_2 \ldots$. If A is a Borel-set we write shortly $\boldsymbol{\mu}(A)$ for the infinite vector $(\mu_1(A), \mu_2(A) \ldots)$. We consider the set

$$E = \{\boldsymbol{\mu}(A) : A \in \mathfrak{B}\} \tag{3.1}$$

as a subset of the Tychonoff-product of the real line by itself.

Suppose, that the measure μ_n has a density function $g_n(x)$ with respect to the Lebesgue-measure that is

$$\mu_n(A) = \int_A g_n(x)\, dx \tag{3.2}$$

furthermore that the finite linear combinations of the functions $g_n(x)$ form a dense subset of L^1 [0, 1]. (That is they form a closed system.)

In this case the set E is neither convex nor closed.

PROOF: Suppose indirectly, that E is closed. Take a sequence of Borel-sets A_j, such that their characteristic functions $X_{A_j}(x)$ converge weakly to $\frac{1}{2}$. (For example A_j may be the set of those numbers whose n-th diadic is 1.)

Weak convergence means, that

$$\int_0^1 x_{A_j}(x) g(x)\, dx \to \int_0^1 \frac{1}{2} g(x)\, dx \tag{3.3}$$

for every $g(x) \in L^1(0, 1)$. Specially, if we take $g(x) = g_n(x)$, we get

$$\mu_n(A_j) \to \frac{1}{2}\mu_n(I) \quad \text{when} \quad j \to \infty \tag{3.4}$$

where I denotes the whole interval [0, 1]. This can also be written in the form

$$\boldsymbol{\mu}(A_j) \to \frac{1}{2}\boldsymbol{\mu}(I) \tag{3.5}$$

Since E is closed by hypothesis, we must have a set A such that

$$\boldsymbol{\mu}(A) = \frac{1}{2}\boldsymbol{\mu}(I) \tag{3.6}$$

Write this out in detail:

$$\int_0^1 \chi_A(x) g_n(x)\, dx = \int_0^1 \frac{1}{2} g_n(x)\, dx \tag{3.7}$$

for all n.

This means that the functions $\chi_A(x)$ and $\frac{1}{2}$ considered as linear functionals on $L^1[0, 1]$ have the same value on a closed system. Hence $\chi_A(x) = \frac{1}{2}$ a.e. what is clearly impossible.

Now it is clear that E is not convex, moreover it is not a cone, because in this case equality (3.6) could be fulfilled. Consider now l_p-valued measures. To begin with we have the following

Remark: If a sequence of non-negativ measures $\mu_1, \mu_2, \ldots$ is such that the sequence $(\mu_1(I), \mu_2(I)\ldots)$ belongs to l_p $(l \leqq p < \infty$, I denotes the whole interval), then

$$\boldsymbol{\mu}(A_j) \to \mathbf{m} \tag{3.8}$$

in the product topology implies

$$\boldsymbol{\mu}(A_j) \to \mathbf{m} \quad \text{in} \quad l_p. \tag{3.9}$$

The proof is easy, we shall omit it. If we use the above example with the necessary modifications we get the stronger result:

There are l_p-valued measures whose range is not closed.

4. Stochastic set functions

First we define stochastic set functions according to Prékopa and repeat the illustrative examples mentioned in [10].

Definition: Let the σ-algebra $\mathcal{R}$ consist of subsets of a space H. To every set $A \in \mathcal{R}$ corresponds a random variable $\xi(A) = \xi(\omega, A)$. We suppose the following:

a) If $A_1, \ldots A_n$ are disjoint sets, then the random variables $\xi(A_1), \ldots \xi(A_r)$ are completely independent.

b) If $A_1, \dots A_r \dots$ is a sequence of disjoint sets then

$$\xi(A) = \sum_{n=1}^{\infty} \xi(A_n)$$

with probability 1, where $A = \bigcup_{1}^{\infty} A_n$. In this case we call the set function $\xi(A)$ a completely additive stochastic set function.

Let be given a completely additive stochastic set function. The distribution function of the random variable $\xi(A)$ will be denoted by $F(A) = F(x, A)$. We shall investigate the set

$$\mathcal{F} = \{F(x, A) : A \in \mathcal{R}\} \tag{4.1}$$

from a topological point of view.

For this purpose we mention here the definition of the usual topology introduced in the space of distribution functions.

Definition: We say that a sequence of distribution functions F_n converges to a distribution function F, shortly $F_n \to F$, if $F_n(x) \to F(x)$ for every x where F is continous.

We can formulate this in the language of characteristic functions, too.

Definition: Denote the characteristic functions of the distribution functions by φ_n, φ. We say, that $F_n \to F$ if $\varphi_n(t) \to \varphi(t)$ for all t.

The two definitions are equivalent (see GNEDENKO). A central result of paper [10] is:

THEOREM 4.: *The set $\mathcal{F}$ defined in* (4.1) *is conditionally compact.*

However $\mathcal{F}$ is not necessarily closed, as we shall see.

If we are given a priori a set characteristic functions $\varphi(A)$, which satisfy (4.2) then using a theorem of Kolmogoroff we can build the random variables $\xi(A)$ of a completely additive stochastic set function, for which the characteristic functions are just the $\xi(A)$-*s*. Hence we give the characteristic functions only in our examples.

Cauchy-type set function

The characteristic function is

$$\varphi(t, A) = e^{-\mu(A)|t|} \tag{4.3}$$

where μ is a finite non-negative measure.

Gauss-type set function

Its characteristic function is

$$\varphi(t, A) = e^{it\mu(A) - \frac{1}{2}t^2\sigma^2(A)} \tag{4.4}$$

where μ, σ^2 are finite measures, σ^2 is non-negativ.

A direct example for this is the Ito-integral of a one-dimensional Wiener-process $w(t)$ $0 \leqq t \leqq 1$. For a Borel-set A we define

$$\xi(A) = \int_A dw(t) \tag{4.5}$$

an Ito integral. Using the arguments of GIKHMAN and SKOROHOD (see [3] p. 491) it is not difficult to show, that $\xi(A)$ is completely additive.

Poisson-type set functions

the characteristic functions are of the form

$$\varphi(t, A) = \exp \mu(A)(e^{it} - 1) \tag{4.6}$$

where μ is a finite non-negativ measure.

Composed Poisson-type set functions

are defined in the following way: let $\mu_1, \mu_2 \ldots$ be a sequence of finite non-negativ measures, and suppose that their sum:

$$\mu(A) = \sum_{n=1}^{\infty} \mu_n(A)$$

is also a finite measure. The characteristic function of $\xi(A)$ should be then

$$\varphi(t, A) = \exp \sum_{n=1}^{\infty} \mu_n(A)(e^{it\gamma_n} - 1) \tag{4.7}$$

where the γ_n−s are given numbers.

The measures $\mu_n(A)$ are called parameters. The random variable $\xi(A)$ for a single A is conceivable in the form $\Sigma \gamma_n \xi_n$, where the $\xi_n - s$ are independent random variables of Poisson-type with parameters $\mu_n(A)$.

In our examples the set of characteristic (or distribution) functions are in a homeomorphic correspondence with the set of values of a finite dimensional measure, except perhaps the last example. Applying Theorem 3. we get the following

THEOREM 5.: *Let the completely additive stochastic set function be of Cauchy-type, Gauss-type or Poisson-type depending on a finite number of parameters. Then the set of distribution functions $\mathcal{F}$ is a closed set.*

We mention here a related result without proof. In case of a Gauss-type set function the random variables $\xi(\omega, A)$ belong to the Hilbert-space of quadratic integrable functions of the variable ω. Denote this space by L^2. Then we state:

The set of random variables $\xi(\omega, A)$ form a closed set in L^2.

Now we give an example, that in case of composed Poisson-type set function the set of distribution functions is not necessarily closed. Let H be the interval

[0, 1] $\mathcal{R}$ the σ-algebra of Borel sets, and put $\gamma_n = 2\pi_n$. Suppose that F is closed, or equivalently, that

$$\Phi = \{\varphi(t, A) : A \in \mathcal{R}\}, \tag{4.8}$$

the set formed by the characteristic functions is closed. We show that in this case the range of the infinite dimensional measure $\boldsymbol{\mu}(A) = (\mu_1(A), \mu_2(A)\ldots)$, i.e.

$$E = \{\boldsymbol{\mu}(A) : A \in \mathcal{R}\} \tag{4.9}$$

is also closed in the product-topology.

Namely let $A_j \in \mathcal{R}$ be a sequence of sets, such that

$$\mu_n(A_j) \to m_n \quad \text{when} \quad i \to \infty \tag{4.10}$$

or in vector-form: $\boldsymbol{\mu}(A_j) \to \mathbf{m}$. According to the remark in paragraph 3. We also have

$$\boldsymbol{\mu}(A_j) \to \mathbf{m} \quad \text{in} \quad l_1. \tag{4.11}$$

It follows easily from this, that the functions

$$k_j(t) = \sum_{n=1}^{\infty} \mu_n(A_j)(e^{2\pi i n t} - 1) \tag{4.12}$$

converge uniformly in [0, 1] to a continuous function $k(t)$. We also have

$$\exp k_j(t) \to \exp k(t) \tag{4.13}$$

uniformly in [0, 1].

Since the set of characteristic functions is closed, we conclude that there exists a set A such that

$$\varphi(t, A) = \exp k(t). \tag{4.14}$$

Hence $k(t)$ can be written in the form

$$k(t) = \sum_{n=1}^{\infty} \mu_n(A)(e^{2\pi i n t} - 1). \tag{4.15}$$

Since $k_j(t)$ converges uniformly to $k(t)$ it holds the more, that

$$k_j(t) \to k(t) \quad \text{in} \quad L^2(0, 1). \tag{4.16}$$

Clearly the numbers $\mu_n(A_j)$, $\mu_n(A)$ are the Fourier-coefficients of the corresponding functions. It follows from (4.16) that the sequence of Fourier-coefficients converge in l_2, which implies also

$$\mu_n(A_j) \to \mu_n(A) = m_n \qquad (j \to \infty) \tag{4.17}$$

for all n. Thus the point $(m_1, m_2 \ldots)$ belongs to E, which proves that E is closed, as asserted.

Choose a sequence of measures μ_n according to the counterexample of paragraph 3. To use them in a composed Poisson-type set function in addition we have to suppose that the density functions $g_n(x)$ are non-negative and $\sum_{n=1}^{\infty} g_n(x)$ is integrable.

Now the set E is not closed, hence neither Φ is closed. In this way we get:

There are composed Poisson-type set functions with an infinite number of parameters, such that the set of distribution functions in not closed.

BIBLIOGRAPHY

[1] BLACKWELL, D.: The range of certain vector integrals, *Proc. of the Amer. Math. Soc.* **2** (1951), 390—395.
[2] FELLER, W.: *An Introduction to Probability Theory and its Applications,* Vol. II. Wiley.
[3] GIKHMAN, I. I.—SKOROHOD, A. V.: *Introduction to the theory of stochastic processes,* (in Russian) Nauka, Moscow 1965.
[4] GNEDENKO, B. V.—KOLMOGOROV, A. N.: *Limit distributions of sums of independent random variables,* (In Russian) Moscow—Leningrad 1949.
[5] HAHN, H.—ROSENTHAL, A.: *Set functions,* New Mexico Press 1949.
[6] HALMOS, P. R.: The range of a measure function. *Bull. Amer. Math. Soc.* **54** (1948), 416—421.
[7] HARTMANN, Ph.: On limits of Riemann Sums and Range of Integrals. *Math. Annalen,* **176** (1968), 208—226.
[8] LYAPUNOFF, A. A.: Sur les fonctions-vecteurs completement additives. *Izv. Akad. Nauk. USSR. Ser. Mat.* **8** (1940), 465—478.
[9] OLECH, C.: Extremal solutions of a control system, *J. Diff. Egons.* **2** (1965), 74—101.
[10] PRÉKOPA, A.: On stochastic set functions I. *Acta Math. Ac. Sci. Hung.* **7** (1956).
[11] YOSIDA, K.: *Functional analysis,* Springer 1956.

Computing Center of the Hungarian Academy of Sciences

(Received February 1, 1971)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 285—305.

INVESTIGATIONS ON WEIGHTED APPROXIMATION BY POLYNOMIALS

by

G. FREUD

1. Introduction

In our present paper we continue our research concerning weighted polynomial approximation on the real axis. We shall formulate and prove the analogoues of JACKSON's and ZYGMUND's theorems for the weights

$$v_{\alpha\beta}(x) = (1+|x|^{\alpha})^{\beta/2\alpha} e^{-1/2|x|^{\alpha}} \qquad (\alpha \geqq 2). \tag{1.1}$$

The case $\alpha=2$, $\beta \geqq 0$ was treated earlier: $\beta=0$ in [1], [2], [3]; and $\beta \geqq 0$, $\alpha=2$ in [4], [5]. For $\alpha=2k$, $\beta=0$ ($k=1, 2, \ldots$) this results where already proved in our series of papers [6], [7], [8]. The essentially new elements needed to solve the more general case treated here are the assymptotic polynomial representations of the functions $v_{\alpha\beta}$ given in part 2 of this paper and the "asymptotic envelope" representation in Theorem 3.2.

Let $\|g(\cdot)\|_p$ denote the $L_p(-\infty, \infty)$-norm of $g(\cdot) \in L_p$. For a function f satisfying $v_{\alpha\beta} f \in L_p$ we define the weighted L_p-modulus of continuity by the following formula:

$$\omega(L_p, v_{\alpha\beta}; f, \delta) \overset{\text{def}}{=} \sup_{0 \leqq t \leqq \delta} \|f(x+t)v_{\alpha\beta}(x+t) - f(x)v_{\alpha\beta}(x)\|_p + \tag{1.2}$$

$$+ \|[\tau(\delta^{\frac{1}{\alpha-1}} x)]^{\alpha-1} f(x) v_{\alpha\beta}(x)\|_p$$

where norms $\|\cdot\|$ refer to the expression "$\cdot$" as a function of the variable x, and

$$\tau(x) = \begin{cases} |x| & (|x| \leqq 1) \\ 1 & (|x| > 1) \end{cases}.$$

Let $\mathfrak{p}_n$ denote the set of polynomials of degree n at most and for an f satisfying $f v_{\alpha\beta} \in L_p$ let

$$\mathscr{E}_n^{(p)}(v_{\alpha\beta}; f) = \inf_{q \in \mathfrak{p}_n} \|(f-q)v_{\alpha\beta}\|_p. \tag{1.3}$$

Our main result reads as follows:

THEOREM 1.1. *For an arbitrary $1 \leqq p \leqq \infty$ and an arbitrary nonnegative integer r and $0 < \varrho < 1$ the estimate*

$$\mathscr{E}_n^{(p)}(v_{\alpha\beta}, f) = O(n^{-(r+\varrho)(1-\gamma)}) \qquad \left(\gamma = \frac{1}{\alpha}\right) \tag{1.4}$$

holds if and only if f has an r-th order derivative[1,2] $f^{(r)}$ *satisfying* $f^{(r)}v_{\alpha\beta} \in L_p$ *and we have*

$$\omega(L_p, v_{\alpha\beta}; f^{(r)}, \delta) = O(\delta^\varrho) \qquad (\delta \to 0). \tag{1.5}$$

Remark. For $p=\infty$ our condition (1.5) is fulfilled if and only if the following two conditions a) and b) are satisfied simultanously:

a) We have uniformly with respect to x

$$v_{\alpha\beta}(x+h)f(x+h) - v_{\alpha\beta}(x)f(x) = O(|h|^\varrho) \qquad (h \to 0). \tag{1.6}$$

b) We have

$$v_{\alpha\beta}(x)f(x) = O(|x|^{-(\alpha-1)\varrho}) \qquad (|x| \to \infty). \tag{1.7}$$

The proof of this equivalence is elementary.

In the present paper we prove that (1.5) implies (1.4). (Theorem 5.3.)

The converse theorem will be proved later. We anounced Theorem 1.1 earlyer as a welfounded conjecture at different occasions[3].

At the end of our present paper we define the weighted Zygmund modulus $\omega_2(L_p, v_{\alpha\beta}; f, \delta)$ of an f satisfying $fv_{\alpha\beta} \in L_p$. We are going to show that (1.4) with $\varrho=1$ is valid if and only if $\omega_2(L_p, v_{\alpha\beta}; f^{(r)}, \delta) = O(\delta)$.

2. Polynomial asymptotics

We denote in the whole paper by $c_1, c_2, \ldots$ positive numbers depending only on the parameters α and β.

LEMMA 2.1. *If ν is an integer, $\nu - 1 \geqq \varrho > 0$ then the binomial coefficients* $C_\varrho^\nu = \dfrac{\varrho(\varrho-1)\ldots(\varrho-\nu+1)}{\nu!}$ *satisfy*

$$|C_\varrho^\nu| \leqq c_1 \varrho^\varrho (\varrho+1)^{1/2} \nu^{-\varrho-1}. \tag{2.1}$$

PROOF. (2.1) is trivial if ϱ is an integer; else

$$C_\varrho^\nu = \frac{1}{\nu!}\varrho(\varrho-1)\ldots(\varrho-[\varrho])(-1)^{\nu-[\varrho]-1}\cdot \tag{2.2}$$

$$\cdot([\varrho]+1-\varrho)([\varrho]+2-\varrho)\ldots(\nu-1-\varrho) = \frac{(-1)^{\nu-[\varrho]-1}}{\nu!}\frac{\Gamma(\varrho+1)\Gamma(\nu-\varrho)}{\Gamma(\varrho-[\varrho])\Gamma(1-\varrho+[\varrho])} =$$

$$= (-1)^{\nu-[\varrho]-1}\frac{\sin\pi(\varrho-[\varrho])}{\pi}\Gamma(\varrho+1)\frac{\Gamma(\nu-\varrho)}{\nu!},$$

[1] For $r=0$ that should be interpreted so that $f^{(0)} \equiv f$ satisfies (1.5).

[2] For $p<\infty$ we mean by that that f is a.e. equal to an r-times iterated integral function of $f^{(r)}$.

[3] E.g. In our report on the Conference on Approximation Theory, Poznan 1972 and in our report at the III. Congress of Bulgarian Mathematiciens, Varna 1972.

the last part follows from the identity $\Gamma(s)\Gamma(1-s) = \pi \sin^{-1}\pi s$. From Stirling's formula we infer

$$c_2 e^{-u} u^u (u+1)^{1/2} \leqq \Gamma(u) \leqq c_3 e^{-u} u^u (u+1)^{1/2} \qquad (u \geqq 0). \tag{2.3}$$

Combining (2.2) and (2.3) we get

$$|C_\varrho^\nu| < c_4(\varrho+1)^{\varrho+(1/2)} e^{-\varrho} \nu^{-\varrho} \left(1-\frac{\varrho}{\nu}\right)^{\nu-\varrho-(1/2)} < c_1(\varrho+1)^{\varrho+(1/2)} \nu^{-\varrho}. \blacksquare$$

LEMMA 2.2. *For every $\varrho > 0$ and every integer $n \geqq \varrho+1$ there exists an $s_n(\varrho, t) \in \mathfrak{P}_n$ satisfying*

$$|t^\varrho - s_n(\varrho, t)| \leqq c_1 \varrho^{\varrho-1}(\varrho+1)^{1/2} n^{-\varrho} e^{-nt} \qquad (0 \leqq t \leqq 1). \tag{2.4}$$

PROOF. If ϱ is an integer then the statement is trivial. If ϱ is not an integer we approximate t^ϱ by the first $n+1$ terms of the binomial series

$$[1-(1-t)]^\varrho = \sum_{\nu=0}^{\infty} C_\varrho^\nu (-1)^\nu (1-t)^\nu.$$

The error is by (2.1) less then

$$\sum_{\nu=n+1}^{\infty} c_1 \varrho^\varrho (\varrho+1)^{1/2} \nu^{-\varrho-1}(1-t)^n <$$

$$< c_1 \varrho^\varrho (\varrho+1)^{1/2} \int_n^\infty \frac{du}{u^{\varrho+1}} (1-t)^n < c_1 \varrho^{\varrho-1}(\varrho+1)^{1/2} n^{-\varrho} e^{-nt}. \blacksquare$$

Remark. Introducing in (2.4) the new variable $u = At$ $(A > 1)$ and multiplying both sides by A^ϱ we get

$$|u^\varrho - A^\varrho s_n(\varrho, A^{-1}u)| \leqq c_1 \varrho^{\varrho-1}(\varrho+1)^{1/2} A^\varrho n^{-\varrho} e^{-nA^{-1}u} \qquad (0 \leqq u \leqq A). \tag{2.5}$$

LEMMA 2.3. *We have*

$$c_5 e^x \leqq \sum_{\nu=0}^{n} \frac{x^\nu}{\nu!} \leqq c_6 e^x \qquad \left(-\frac{n}{4} \leqq x \leqq \frac{n}{4}\right). \tag{2.6}$$

PROOF. For $x \geqq 0$ (2.6) is a consequence of the relation

$$0 \leqq 1 - e^{-x} \sum_{\nu=0}^{n} \frac{x^\nu}{\nu!} = \frac{1}{n!} \int_0^x t^n e^{-t}\, dt \leqq \frac{1}{n^n e^{-n}} \int_0^x t^n\, dt = \frac{x^{n+1}}{(n+1)n^n e^{-n}} \leqq \left(\frac{ex}{n}\right)^{n+1}.$$

For $-\frac{n}{4} < x < 0$ (2.6) follows from

$$\left| e^{-x}\left(e^x - \sum_{\nu=0}^{n} \frac{x^\nu}{\nu!}\right)\right| \leqq e^{-x} \frac{|x|^{n+1}}{(n+1)!} \leqq e^{|x|} \frac{|x|^{n+1}}{(n+1)^{n+1} e^{-(n+1)}} =$$

$$= \left\{\frac{|x|}{n+1} \exp\left(1 + \frac{|x|}{n+1}\right)\right\}^{n+1} < (e^{5/4}/4)^{n+1} < e^{5/4}/4 < 1. \blacksquare$$

LEMMA 2.4. *For every* $\alpha \geqq 1$ *there exist two polynomial sequences* $\{r_n(\alpha, x) \in \mathfrak{p}_n\}$ *and* $\{R_n(\alpha, x) \in \mathfrak{p}_n\}$ *satisfying*

$$c_7 e^{-x^\alpha/2} \leqq r_n(\alpha, x) \leqq c_8 e^{-x^\alpha/2} \qquad (0 \leqq x \leqq c_9 n^{1/\alpha}) \tag{2.7}$$

resp.

$$c_{10} e^{x^\alpha/2} \leqq R_n(\alpha, x) \leqq c_{11} e^{x^\alpha/2} \qquad (0 \leqq x \leqq c_{12} n^{1/\alpha}). \tag{2.8}$$

PROOF. For $\alpha = 1$ this is a consequence of Lemma 2.3. Let $\alpha > 1$. By Lemma 2.3 we have for arbitrary fixed $c_{13} > 0$

$$c_5 \leqq e^{x^\alpha/2} \sum_{\nu=0}^{c_{13} n} (-1)^\nu \frac{x^{\nu\alpha}}{2^\nu \nu!} \leqq c_6 \qquad \left[0 \leqq x \leqq \left(\frac{c_{13} n - 1}{4}\right)^{1/\alpha}\right]. \tag{2.9}$$

Inserting $A = n^{1/\alpha}$ in (2.5) let

$$r_n(\alpha, x) = \sum_{\nu=0}^{c_{13} n} \frac{(-1)^\nu}{2^\nu \nu!} n^\nu s_n(\nu\alpha, n^{-1/\alpha} x)$$

where we are going to dispose on c_{13} later. By (2.5) setting $\gamma = \frac{1}{\alpha}$

$$\left| r_n(\alpha, x) - \sum_{\nu=0}^{c_{13} n} (-1)^\nu \frac{x^{\nu\alpha}}{2^\nu \nu!} \right| \leqq$$

$$\leqq c_1 e^{-n^{1-\gamma} x} \sum_{\nu=1}^{c_{13} n} (\nu\alpha)^{\nu\alpha - 1} (\nu\alpha + 1)^{1/2} \frac{n^{-(1-\gamma)\nu\alpha}}{2^\nu \cdot \nu!} = c_1 e^{-n^{1-\gamma} x} \sum_{\nu=1}^{c_{13} n} K_\nu \tag{2.10}$$

Fixing c_{13} sufficiently small we get

$$\frac{K_{\nu+1}}{K_\nu} < \frac{1}{2(\nu+1)} [(\nu+1)\alpha]^\alpha \left(1 + \frac{1}{\nu}\right)^{\nu\alpha} n^{-(\alpha-1)} <$$

$$< \frac{e^\alpha \alpha^\alpha}{2} \left(\frac{\nu+1}{n}\right)^{\alpha-1} < \frac{1}{2} \qquad (1 \leqq \nu \leqq c_{13} n)$$

consequently

$$\sum_{\nu=1}^{c_{13} n} K_\nu < 2K_1 = c_{14} n^{-(\alpha-1)}. \tag{2.11}$$

By (2.10) and (2.11) we have for $0 < x < c_{15} n^{1/\alpha}$ and $n \geqq c_{16}$ with a suitably great choice of c_{16} taking $\gamma = \frac{1}{\alpha}$

$$e^{x^\alpha/2} \left| r_n(\alpha, x) - \sum_{\nu=0}^{c_{13} n} (-1)^\nu \frac{x^{\nu\alpha}}{2^\nu \nu!} \right| \leqq \tag{2.12}$$

$$\leqq c_{17} c_{16}^{-(\alpha-1)} \exp\left\{-n^{1-\gamma} x + \frac{x^\alpha}{2}\right\} \leqq c_{17} c_{16}^{-(\alpha-1)} < \frac{1}{2} c_5$$

i.e. by (2.9) we have for $n \geqq c_{16}$

$$\frac{1}{2}c_5 \leqq e^{x^\alpha/2} r_n(\alpha, x) \leqq c_6 + \frac{1}{2}c_5 \qquad (0 < x < c_{15} n^{1/\alpha}).$$

This proves (2.7) for $n \geqq c_{16}$; for $1 \leqq n < c_{16}$ we can take $r_n(\alpha, x) = 1$ and then adjust the constants c_7 and c_8.

The proof of (2.8) is similar but simpler. We set

$$R_n(\alpha, x) = \sum_{\nu=0}^{c_{13}n} \frac{1}{2^\nu \nu!} n^\nu s_n(\nu\alpha; n^{-1/\alpha} x).$$

Thus

$$\left| R_n(\alpha, x) - \sum_{\nu=0}^{c_{13}n} \frac{x^{\nu\alpha}}{2^\nu \nu!} \right| \leqq c_1 e^{-n^{1-\gamma} x} \sum_{\nu=1}^{c_{13}n} K_\nu \qquad \left(\gamma = \frac{1}{\alpha}\right) \tag{2.13}$$

provided that $n \geqq c_{17}$. In this case (2.8) follows from (2.9) and (2.13). For $1 \leqq n \leqq c_{17}$ we take — as before — $R_n(\alpha, x) = 1$ and adjust the constants c_{10} and c_{11}. ∎

LEMMA 2.5. *For every real β there exists a sequence $\{q_n(\beta, x)\}$ satisfying*

$$c_{18}(1+x)^\beta \leqq q_n(\beta, x) \leqq c_{19}(1+x)^\beta \qquad (0 \leqq x \leqq c_{20} n) \tag{2.14}$$

PROOF. The statement is trivial for $\beta = 0$. For $\beta > 0$ let

$$q_n(\beta, x) = 1 + (c_{20} n)^\beta s_n[\beta, (c_{20} n)^{-1} x]. \tag{2.15}$$

We infer from (2.5) that for sufficiently small but fixed c_{20} we have

$$|q_n(\beta, x) - (1 + x^\beta)| \leqq \frac{1}{2} \qquad (0 \leqq x \leqq c_{20} n). \tag{2.16}$$

(2.16) implies that (2.14) holds.

For $\beta < 0$ Lemma 2.5 was proved by G. P. NÉVAI [11]. ∎

THEOREM 2.6. *For every $\alpha \geqq 2$ and every real β there exist two polynonial sequences $\{p_\nu(\alpha, \beta; x) \in \mathfrak{p}_\nu\}$ and $\{P_\nu(\alpha, \beta; x) \in \mathfrak{p}_\nu\}$ satisfying*

$$c_{21} \leqq v_{\alpha\beta}^{-1}(x) p_\nu(\alpha, \beta; x) \leqq c_{22} \qquad (|x| \leqq c_{23} \nu^{1/\alpha}) \tag{2.17}$$

and

$$c_{24} \leqq v_{\alpha\beta}(x) P_\nu(\alpha, \beta; x) \leqq c_{25} \qquad (|x| \leqq c_{26} \nu^{1/\alpha}). \tag{2.18}$$

PROOF. Let

$$p_\nu(\alpha, \beta; x) = q_{[\nu/4]}\left(\frac{\alpha}{2}, \frac{\beta}{2}, x^2\right] r_{[\nu/4]}\left(\frac{\alpha}{2}, x^2\right)$$

and apply Lemma 2.4 and Lemma 2.5. ∎

Remark. We are going to make use only of the first part (i.e. (2.17)) of Theorem 2.6. We included (2.18) resp. (2.8) in view of later applications and considering the fact that no additional effort was needed in proving them.

3. Bohr type inequalities and L_1 — approximation

We start by a polynomial construction differing from those in the previous paragraph. The essential content of our subsequent Theorem 3.2 is that we construct a set of polynomials of fixed degree so that the weight function is an assymptotic envelope of this set. Let $\alpha>1$ and $A_0>0$ be fixed.

The function

$$f_a(x) = x^a(1-x)^{1-a}e^{-A_0x^\alpha} \qquad (0 \leqq a \leqq 1) \tag{3.1}$$

has a single maximum ξ in (0, 1) and we have

$$a=\varphi_\alpha(\xi) = \xi + A_0\alpha\xi^\alpha(1-\xi). \tag{3.2}$$

Let us take $d_0<{}^1/_2$ — in dependence on A_0 — so small that

$$\frac{1}{2} \leqq \varphi'_\alpha(t) \leqq 2 \qquad (0 \leqq t \leqq d_0) \tag{3.3}$$

this choice is possible since $\varphi'_\alpha \in C$ and $\varphi'_\alpha(0)=1$.

Since $\varphi(0)=0$ we have $a<1$ provided that $0\leqq\xi\leqq d_0<{}^1/_2$.

For an arbitrary $\xi\in[0, d_0]$ let $a=\varphi_\alpha(\xi)$ and $a'=n^{-1}[na]$. Then

$$a' = \frac{p}{n}, \quad 1-a' = \frac{q}{n} \tag{3.4}$$

where $p=[na]$, $q = n-[na]$ are nonnegative integers not exceeding n.

We set

$$\tau_n(\alpha, A_0, \xi; x) = e^{A_0n\xi^\alpha}\left(\frac{x}{\xi}\right)^p\left(\frac{1-x}{1-\xi}\right)^q \in \mathfrak{p}_n \tag{3.5}$$

where p and q are defined by (3.4).

LEMMA 3.1. *The polynomials* $\tau_n(\alpha, A_0, \xi; x)\in\mathfrak{p}_n$ *satisfy for* $\xi\in[0, d_0]$, $\alpha>1$

$$0 \leqq \tau_n(\alpha, A_0, \xi; x) \leqq d_1e^{A_0nx^\alpha} \qquad (0 \leqq x \leqq 1) \tag{3.6}$$

and

$$\tau_n(\alpha, A_0, \xi; \xi) = e^{A_0n\xi^\alpha} \tag{3.7}$$

where $d_0, d_1, \ldots$ *are positive numbers depending on* α *and* A_0 *but independent of* n *and* ξ.

PROOF. We have by construction

$$\tau_n(\alpha, A_0, \xi; x) = \frac{[f_{a'}(x)]^n}{[f_{a'}(\xi)]^n}e^{A_0nx^\alpha} \tag{3.8}$$

where $a'=\varphi_\alpha(\xi')$. By (3.3) and by $a-n^{-1}\leqq a'\leqq a$

$$0 \leqq \xi' \leqq \xi \leqq \xi' + \frac{2}{n} \tag{3.9}$$

consequently observing that $0 \leqq \xi' \leqq \xi \leqq d_0 \leqq 1/2$ we have by (3.9)

$$\frac{f_{a'}(\xi')}{f_{a'}(\xi)} = \left(\frac{\xi'}{\xi}\right)^a \left(\frac{1-\xi'}{1-\xi}\right)^{1-a} e^{-A_0(\xi'^\alpha - \xi^\alpha)} \leqq \frac{1-\xi'}{1-\xi} e^{A_0(\xi^\alpha - \xi'^\alpha)} \leqq 1 + \frac{d_2}{n} \tag{3.10}$$

where d_2 is dependent on A_0 and α but not depending on n or ξ. We infer from (3.8) and (3.10) that

$$\tau_n(\alpha, A, \xi, x) \leqq \left(1 + \frac{d_2}{n}\right)^n e^{Anx^\alpha} \leqq e^{d_2} e^{Anx^\alpha}. \quad \blacksquare$$

THEOREM 3.2. *For every fixed $A_1 > 0$, $\alpha \geqq 2$ and every $\xi \in (-d_3 n^{1/\alpha}, d_3 n^{1/\alpha})$ there exists a polynomial sequence $\{t_n(\alpha, \beta, A_1; \xi; x) \in \mathfrak{p}_n$ satisfying*

$$0 < t_n(\alpha, \beta, A_1; \xi; x) \leqq d_4 [v_{\alpha\beta}(x)]^{-1} \qquad (|x| \leqq A_1 n^{1/\alpha}) \tag{3.11}$$

and

$$t_n(\alpha, \beta, A_1; \xi; \xi) \geqq d_5 [v_{\alpha\beta}(\xi)]^{-1} \qquad (|\xi| \leqq d_3 n^{1/\alpha}) \tag{3.12}$$

where d_3, d_4 and d_5 do not depend either on n or on ξ.

PROOF. By Lemma 3.1 we infer that for $n \geqq 4$, $\alpha > 2$

$$\tau_n^*(x) = \tau_{[n/4]}\left(\frac{\alpha}{2}, A_0, \cdot \left(\frac{1}{2} A_0 n\right)^{-2/\alpha} \xi^2, \left(\frac{1}{2} A_0 n\right)^{-2/\alpha} x^2\right) \leqq$$

$$\leqq d_1 e^{|x|^\alpha} \qquad \left(|x| \leqq \left(\frac{1}{8} A_0 n\right)^{1/\alpha}\right),$$

$$\tau_n^* \in \mathfrak{p}_{n/2} \quad \text{and} \quad \tau_n^*(\xi) = e^{|\xi|^\alpha}.$$

Defining A_0 by $\left(\frac{1}{8} A_0\right)^{1/\alpha} = A_1$, we se by Lemma 2.5 that (3.11) holds taking $\alpha > 2$, $n \geqq c_{27}$, $t_n(x) = q_{[n/4]}(\beta/2, x^2)\tau_n^*(x)$; for $n < c_{27}$ we take $t_n \equiv 1$ and modify d_4 and d_5 if necessary.

If $\alpha = 2$, take

$$\tau_n^*(x) = \sum_{\nu=0}^{[n/4]} \frac{x^{2\nu}}{2^\nu \nu!}, \quad t_n = q_{[n/4]}\left(\frac{\beta}{2}, x^2\right)\tau_n^*(x).$$

Then clearly $e^{x^2/2} \geqq \tau_n^*(x)$, $\tau_n^* \in \mathfrak{p}_{n/2}$ so that by Lemma 2.3 and Lemma 2.5 $t_n(x)$, chosen as above, satisfies (3.11).

By $p_n(v_{\alpha\beta}^2; x) = \gamma_n(v_{\alpha\beta}^2)x^n + \ldots \in \mathfrak{p}_n$, $\gamma_n(v_{\alpha\beta}^2) > 0$ we denote the orthonormal polynomials with respect to $v_{\alpha\beta}^2$ (see (1.1)) i.e.

$$\int_{-\infty}^{\infty} p_n(v_{\alpha\beta}^2; x) p_m(v_{\alpha\beta}^2; x)\, dx = \delta_{mn}.$$

We are going to estimate first the Christoffel function

$$\lambda_n(v_{\alpha\beta}^2; x) = \left\{\sum_{\nu=0}^{n-1} [p_\nu(v_{\alpha\beta}^2; x)]^2\right\}^{-1}. \tag{3.13}$$

We refer to the minimum property

$$\lambda_n(v_{\alpha\beta}^2; x) = \min_{\substack{P\in\mathfrak{P}_{n-1}\\ P(x)=1}} \int_{-\infty}^{\infty} [P(t)]^2 v_{\alpha\beta}^2(t)\,dt \tag{3.14}$$

see e.g. [9], Theorem I. 4.1.

LEMMA 3.3. *We have for every* $P\in\mathfrak{P}_{n-1}$, *every real* β *and every* $\alpha>1$ *for a suitably large* $c_{29}>1$

$$\int_{-\infty}^{\infty} [P(t)]^2 v_{\alpha\beta}^2(t)\,dt \leqq (1+c_{30}e^{-c_{28}n}) \int_{-c_{29}n^{1/\alpha}}^{c_{29}n^{1/\alpha}} [P(t)]^2 v_{\alpha\beta}^2(t)\,dt. \tag{3.15}$$

PROOF. For $\beta\geqq 0$ 3.15 is a consequence of the inequality

$$\int_{c_{29}n^{1/\alpha}}^{\infty} [P(t)]^2 v_{\alpha\beta}^2(t)\,dt \leqq e^{-2c_{28}n} \int_{-n^{1/\alpha}}^{n^{1/\alpha}} [P(t)]^2 v_{\alpha\beta}^2(t)\,dt \qquad (P\in\mathfrak{P}_n)$$

wich was proved (even in a slightly more precise form) by G. P. NÉVAI [12], Lemma 3.2.

For $\beta<0$ we have by the the same inequality applied for "$\beta=0$"

$$\int_{c_{29}n^{1/\alpha}}^{\infty} [P(t)]^2 v_{\alpha\beta}^2(t)\,dt \leqq \int_{c_{29}n^{1/\alpha}}^{\infty} [P(t)]^2 v_{\alpha 0}^2(t)\,dt \leqq e^{-2c_{28}n} \int_{-n^{1/\alpha}}^{n^{1/\alpha}} [P(t)]^2 v_{\alpha 0}^2(t)\,dt \leqq$$

$$\leqq (1+n^{2/\alpha})^{-\beta} e^{-2c_{28}n} \int_{-n^{1/\alpha}}^{n^{1/\alpha}} [P(t)]^2 v_{\alpha\beta}^2(t)\,dt \leqq$$

$$\leqq c_{30}e^{-c_{28}n} \int_{-n^{1/\alpha}}^{n^{1/\alpha}} [P(t)]^2 v_{\alpha\beta}^2(t)\,dt. \quad \blacksquare$$

The results of the rest of this section 3 where proved for $\beta=0$ and $\alpha=2, 4, 6, \ldots$ in [6] resp. for $\beta>0$ and $\alpha=2, 4, 6 \ldots$ by G. P. NÉVAI [11].

LEMMA 3.4. *We have for every real* β *and* $\alpha\geqq 2$

$$\lambda_n(v_{\alpha\beta}^2; \xi) \leqq c_{31}n^{-1+\gamma}v_{\alpha\beta}^2(\xi) \qquad \left(\gamma=\frac{1}{\alpha}\right). \tag{3.16}$$

PROOF. Let

$$P(t) = d_4^{-1}t_{[n/2]}(\alpha, \beta, c_{29}; \xi; t)v_{\alpha\beta}(\xi)p(t) \tag{3.17}$$

where $|\xi|\leqq d_3(c_{29})n^{1/\alpha}$ and $p\in\mathfrak{P}_{n/2-1}$ satisfies $p(\xi)=1$. Then by Theorem 3.2

$$P^2(t)v_{\alpha\beta}^2(t) \leqq c_{33}p^2(t) \qquad (|x|\leqq c_{29}n^{1/\alpha}) \tag{3.18}$$

and

$$P(\xi) \geqq c_{34}v_{\alpha\beta}^{-1}(\xi). \tag{3.19}$$

From Lemma 3.3 and (3.18)

$$\int_{-\infty}^{\infty} [P(t)]^2 v_{\alpha\beta}^2(t)\,dt \leqq c_{36} \int_{-c_{29}n^{1/\alpha}}^{c_{29}n^{1/\alpha}} [P(t)]^2 v_{\alpha\beta}^2(t)\,dt \leqq \tag{3.20}$$

$$\leqq c_{36}c_{33} \int_{-c_{29}n^{1/\alpha}}^{c_{29}n^{1/\alpha}} p^2(t)\,dt.$$

By (3.14), (3.19) and (3.20) we have for every $P(t)$ satisfying our conditions

$$\lambda_n(v_{\alpha\beta}^2;\,\xi) = \int_{-\infty}^{\infty} \left[\frac{P(x)}{P(\xi)}\right]^2 v_{\alpha\beta}^2(x)\,dx \leqq$$

$$\leqq c_{37} v_{\alpha\beta}^2(\xi) \int_{-c_{29}n^{1/\alpha}}^{c_{29}n^{1/\alpha}} p^2(t)\,dt \qquad (|\xi| \leqq c_{35}n^{1/\alpha})$$

i.e. applying the transformation $t = c_{29}n^{1/\alpha}u$

$$\lambda_n(v_{\alpha\beta}^2;\,\xi) \leqq c_{37} v_{\alpha\beta}^2(\xi) \min_{\substack{p \in \mathcal{P}_{(n/2)-1} \\ p(\xi)=1}} \int_{-c_{29}n^{1/\alpha}}^{c_{29}n^{1/\alpha}} p^2(t)\,dt =$$

$$= c_{38} v_{\alpha\beta}^2(\xi) n^{1/\alpha} \min_{\substack{p \in \mathcal{P}_{(n/2)-1} \\ p(c_{29}^{-1}n^{-1/\alpha}\xi)=1}} \int_{-1}^{1} p^2(u)\,du.$$

The "min" expression on the right side is a Christoffel function with respect to the weight $w_1(x)=1$ with support $[-1, 1]$. Consequently it is smaller than $c_{39}n^{-1}$ for $|c_{29}^{-1}n^{-\gamma}\xi| \leqq 1/2$ (see e.g. [9], Lemma III. 3.1); this gives the desired estimate (3.16). ∎

LEMMA 3.5. *Let*

$$\Gamma_\xi(x) = \begin{cases} 0 & (x \leqq \xi) \\ 1 & (x > \xi) \end{cases} \tag{3.21}$$

then

$$\mathscr{E}_n^{(1)}(v_{\alpha\beta};\,\Gamma_\xi) \leqq c_{39}n^{-1+\gamma} v_{\alpha\beta}(\xi) \qquad \left(\gamma = \frac{1}{\alpha}\right). \tag{3.22}$$

PROOF. Taking the elementary inequalities

$$c_{40} v_{\alpha\beta}(2^{1/\alpha}x) \leqq v_{\alpha\beta}^2(x) \leqq c_{41} v_{\alpha\beta}(2^{1/\alpha}x)$$

in consideration, (3.22) is equivalent to

$$\mathscr{E}_n^{(1)}(v_{\alpha\beta}^2;\,\Gamma_\xi) \leqq c_{42}n^{-1+\gamma} v_{\alpha\beta}^2(\xi) \qquad \left(\gamma = \frac{1}{\alpha}\right). \tag{3.23}$$

Approximating Γ_ξ by the Markov—Stieltjes polynomials (see e.g. [9] § I. 5) we infer that

$$\mathscr{E}_n^{(1)}(v_{\alpha\beta}^2; \Gamma_\xi) \leqq \lambda_n(v_{\alpha\beta}^2; \xi);$$

by Lemma 3.4 this proves (3.23) for $|\xi| \leqq c_{32} n^{1/\alpha}$. For $\xi > c_{32} n^{1/\alpha}$ we approximate Γ_ξ by $s_0 \equiv 0$ resp. for $\xi < -c_{32} n^{1/\alpha}$ by $s_1 \equiv 1$. In both cases we have, taking $\gamma = \frac{1}{\alpha}$

$$\int_{-\infty}^{\infty} |\Gamma_\xi(x) - s_i(x)| v_{\alpha\beta}^2(x)\, dx \leqq \int_{|\xi|}^{\infty} v_{\alpha\beta}^2(x) dx \leqq c_{43} |\xi|^{-\alpha+1} v_{\alpha\beta}^2(\xi) \leqq c_{44} n^{-1+\gamma} v_{\alpha\beta}^2(\xi). \quad \blacksquare$$

We are now prepared to prove our fundamental Bohr-type inequality:

THEOREM 3.6. *Let* $\|g v_{\alpha\beta}\|_\infty \leqq 1$ *and for every* $p \in \mathfrak{P}_n$ *let* $\gamma = \frac{1}{\alpha}$

$$\int_{-\infty}^{\infty} g(x) p(x) v_{\alpha\beta}^2(x)\, dx = 0 \qquad (p \in \mathfrak{P}_n) \tag{3.24}$$

be satisfied then

$$|G(\xi)| = \left| \int_{\xi}^{\infty} g(x) v_{\alpha\beta}^2(x)\, dx \right| \leqq c_{45} n^{-1+\gamma} v_{\alpha\beta}(\xi) \qquad (-\infty < \xi < \infty). \tag{3.25}$$

PROOF. We have for an arbitrary $p \in \mathfrak{P}_n$

$$\left| \int_{\xi}^{\infty} g(x) v_{\alpha\beta}^2(x)\, dx \right| = \left| \int_{-\infty}^{\infty} \Gamma_\xi(x) g(x) v_{\alpha\beta}^2(x)\, dx \right|$$

so that

$$\left| \int_{\xi}^{\infty} g(x) v_{\alpha\beta}^2(x)\, dx \right| \leqq \mathscr{E}_n^{(1)}(v_{\alpha\beta}; \Gamma_\xi) \|g v_{\alpha\beta}\|_\infty. \tag{3.26}$$

(3.26) and (3.22) imply (3.25). $\blacksquare$

THEOREM 3.7. *Let* $F(x)$ *be of bounded variation in every finite intervall then*

$$\mathscr{E}_n^{(1)}(v_{\alpha\beta}; F) \leqq c_{46} n^{-1+\gamma} \int_{-\infty}^{\infty} v_{\alpha\beta}(\xi) |dF(\xi)| \qquad \left(\gamma = \frac{1}{\alpha}\right). \tag{3.27}$$

PROOF. Let the integral on the right side of (3.27) be finite. Then we infer from the inequality

$$|F(x)| \leqq |F(\pm\Omega)| + \left| \int_{\Omega}^{|x|} v_{\alpha\beta}^{-1}(t) v_{\alpha\beta}(t) |dF(\pm t)| \right| \leqq$$

$$\leqq |F(\pm\Omega)| + v_{\alpha\beta}^{-1}(x) \left| \int_{\Omega}^{|x|} v_{\alpha\beta}(t) |dF(\pm t)| \right|$$

that

$$\lim_{|x| \to \infty} F(x) v_{\alpha\beta}(x) = 0. \tag{3.28}$$

By the duality principle of S. M. NIKOLSKI [13] (see also e.g. [4]) we have

$$\mathscr{E}_n^{(1)}(v_{\alpha\beta}; F) = \sup_{g \in \Delta} \int_{-\infty}^{\infty} F(x) g(x) v_{\alpha\beta}^2(x)\, dx \tag{3.29}$$

where Δ is the set of functions g satisfying the conditions of Theorem 3.4. We set

$$G(x) = \int_x^{\infty} g(t) v_{\alpha\beta}^2(t)\, dt.$$

By partial integration we obtain using (3.28) and Theorem 3.4

$$\int_{-\infty}^{\infty} F(x) g(x) v_{\alpha\beta}^2(x)\, dx = \int_{-\infty}^{\infty} G(\xi)\, dF(\xi) \leqq \tag{3.30}$$

$$\leqq c_{45} n^{-1+1/\alpha} \int_{-\infty}^{\infty} v_{\alpha\beta}(\xi) |dF(\xi)|;$$

(3.27) is a consequence of (3.29) and (3.30). ∎

4. Study and application of the orthogonal polynomial expansion

Let us denote by $x_{1n}(v_{\alpha\beta}^2)$ the greatest zero of the orthogonal polynomial $p_n(v_{\alpha\beta}^2; x)$ with respect to the weight $v_{\alpha\beta}$ and let $\gamma_n(v_{\alpha\beta}^2)$ be the coefficient of x^n in $p_n(v_{\alpha\beta}^2; x)$. By a classical result of P. L. TCHEBYCHEV we have (see e.g. [14] § 7.72)

$$x_{1n}(v_{\alpha\beta}^2) = \max_{p \in \mathfrak{P}_{n-1}} \frac{\int_{-\infty}^{\infty} x p^2(x) v_{\alpha\beta}^2(x)\, dx}{\int_{-\infty}^{\infty} p^2(x) v_{\alpha\beta}^2(x)\, dx} \tag{4.1}$$

and

$$\max_{\nu \leqq n} \frac{\gamma_{\nu-1}(v_{\alpha\beta}^2)}{\gamma_\nu(v_{\alpha\beta}^2)} \leqq x_{1,n+1}(v_{\alpha\beta}) \leqq 2 \max_{\nu \leqq n} \frac{\gamma_{\nu-1}(v_{\alpha\beta}^2)}{\gamma_\nu(v_{\alpha\beta}^2)} \tag{4.2}$$

(see [10]).

LEMMA 4.1. *We have*

$$x_{1n}(v_{\alpha\beta}^2) \leqq c_{46} n^{1/\alpha} \tag{4.3}$$

and

$$\frac{\gamma_{n-1}(v_{\alpha\beta}^2)}{\gamma_n(v_{\alpha\beta}^2)} \leqq c_{47} n^{1/\alpha}. \tag{4.4}$$

For similar results see G. P. NÉVAI [12].

PROOF. We have by Lemma 3.3 for $p \in \mathfrak{P}_{n-1}$

$$\int_{-\infty}^{\infty} x^2 p^2(x) v_{\alpha\beta}^2(x)\,dx \leqq c_{48} \int_{-c_{29}n^{1/\alpha}}^{c_{29}n^{1/\alpha}} x^2 p^2(x) v_{\alpha\beta}^2(x)\,dx \leqq$$

$$\leqq c_{49} c_{29}^2 n^{2/\alpha} \int_{-\infty}^{\infty} p^2(x) v_{\alpha\beta}^2(x)\,dx$$

so that

$$\int_{-\infty}^{\infty} |x p^2(x) v_{\alpha\beta}^2(x)\,dx \leqq \left\{ \int_{-\infty}^{\infty} x^2 p^2(x) v_{\alpha\beta}^2(x)\,dx \times \right. \tag{4.5}$$

$$\left. \times \int_{-\infty}^{\infty} p^2(x) v_{\alpha\beta}^2(x)\,dx \right\}^{1/2} \leqq c_{48}^{1/2} c_{29} n^{1/\alpha} \int_{-\infty}^{\infty} p^2(x) v_{\alpha\beta}^2(x)\,dx.$$

Now (4.3) is a consequence of (4.1) and (4.5) and (4.4) is implied by (4.2) and (4.3). ∎

LEMMA 4.2. *Let* $w(x) \geqq 0$ $(-1 \leqq x \leqq 1)$ *and* $\log w(\cos\theta) \in L(0, \pi)$. *If the real polynomial*

$$p(x) = \eta_n x^n + q(x) \qquad (q \in \mathfrak{P}_{n-1})$$

satisfies

$$\int_{-1}^{1} p^2(x) w(x)\,dx \leqq 1$$

then

$$\eta_n \leqq \pi^{-1/2} 2^n \exp\left\{ -\frac{1}{2\pi} \int_0^{\pi} \log\left[w(\cos\theta) \sin\theta \right] d\theta \right\}. \tag{4.6}$$

PROOF. We have

$$1 \geqq \int_{-1}^{1} p^2(x) w(x)\,dx = \frac{1}{2} \int_{-\pi}^{\pi} p^2(\cos\theta) w(\cos\theta) |\sin\theta|\,d\theta. \tag{4.7}$$

By the representation theorem of G. SZEGŐ (see [14], chapter X.) there exists a function $D(z) \in H_2$ so that the boundary values satisfy a.e. $|D(e^{ie})|^2 = w(\cos\theta)|\sin\theta|$ and

$$D^2(0) = \exp\left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \log\left[w(\cos\theta) |\sin\theta| \right] d\theta \right\}. \tag{4.8}$$

In this way the boundary values of the function $\Phi(z) = z^n p\left[\frac{1}{2}(z + z^{-1})\right] D(z) \in H_2$ satisfy for a.e. θ

$$|\Phi(e^{i\theta})|^2 = p^2(\cos\theta) w(\cos\theta) |\sin\theta| \tag{4.9}$$

and we have

$$\Phi(0) = 2^{-n}\eta_n D(0). \tag{4.10}$$

We infer from $\Phi \in H_2$, (4.10), (4.9) and (4.7)

$$[2^{-n}\eta_n D(0)]^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |\Phi(e^{i\theta})|^2\, d\theta = \tag{4.11}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} p^2(\cos\theta)\, w(\cos\theta)|\sin\theta|\, d\theta \leqq \frac{1}{\pi}.$$

Now (4.6) is a consequence of (4.8) and (4.11). ∎

LEMMA 4.3. *We have*

$$|p_n(v_{\alpha\beta}^2; x)|v_{\alpha\beta}(x) \leqq c_{49}e^{-c_{50}n} \qquad (|x| \geqq c_{51}n^{1/\alpha}). \tag{4.12}$$

PROOF. Since $v_{\alpha\beta}$ is even, $p_n(v_{\alpha\beta}^2)$ is even resp. odd at the same time as n is even or odd. We may assume $|x|$ greater then $x_{1n}(v_{\alpha\beta}^2)$; this is by Lemma 4.1 certainly the case for $|x| > c_{46}n^{1/\alpha}$ so we take $c_{51} \geqq c_{46}$. Thus denoting by $x_{1n} > x_{2n} > \ldots$ the zeros of $p_n(v_{\alpha\beta}^2)$ we have

$$|p_n(v_{\alpha\beta}^2; x)| = \gamma_n(v_{\alpha\beta}^2)\left|x^{n-2[n/2]}\prod_{k=1}^{[n/2]}(x^2 - x_{kn}^2)\right| \leqq \gamma_n(v_{\alpha\beta}^2)|x|^n.$$

Further we have

$$1 = \int_{-\infty}^{\infty} p_n^2(v_{\alpha\beta}^2; x)v_{\alpha\beta}^2(x)\, dx \geqq \tag{4.13}$$

$$\geqq \int_{-n^{1/\alpha}}^{n^{1/\alpha}} p_n^2(v_{\alpha\beta}^2; x)v_{\alpha\beta}^2(x)\, dx \geqq n^{-|\beta|/\alpha}\int_{-n^{1/\alpha}}^{n^{1/\alpha}} p_n^2(v_{\alpha\beta}^2; x)e^{-|x|^\alpha}\, dx =$$

$$= \int_{-1}^{1} [n^{(-|\beta|+1)/2\alpha}p_n(v_{\alpha\beta}^2; n^{1/\alpha}t)]^2 e^{-n|t|^\alpha}\, dt.$$

The coefficient of t^n in the polynomial $n^{(-|\beta|+1)/2\alpha}\, p_n(v_{\alpha\beta}^2; n^{1/\alpha}t)$ is $\eta_n = n^{(-|\beta|+1)/2\alpha} \cdot n^{n/\alpha}\gamma_n(v_{\alpha\alpha}^2)$. Consequently by Lemma 4.2.

$$n^{(-|\beta|+1)/2\alpha}n^{n/\alpha}\gamma_n(v_{\alpha\beta}^2) \leqq$$

$$\leqq \pi^{-1/2}2^n \exp\left\{\frac{n}{2\pi}\int_0^{\pi} |\cos\theta|^\alpha\, d\theta - \frac{1}{2\pi}\int_0^{\pi} \log(\sin\theta)\, d\theta\right\}$$

i.e.

$$\gamma_n(v_{\alpha\beta}^2) \leqq c_{52}e^{c_{53}n}n^{-n/\alpha}. \tag{4.14}$$

Let us take c_{54} so large that for every natural n, $x^n e^{-x^\alpha/2}$ is decreasing for $|x|>c_{54}n^{1/\alpha}$ then by (4.13) and (4.14) we have for $|x|>c_{54}n^{1/\alpha}$

$$v_{\alpha\beta}(x)|p_n(v_{\alpha\beta}^2;x)| \leqq c_{52}e^{c_{53}n}n^{-n/\alpha}|x|^n c_{55}n^{|\beta|/\alpha}e^{-|x|^\alpha/2} \leqq$$

$$\leqq c_{56}c^{c_{57}n}n^{-n/\alpha}\max_{x>c_{54}n^{1/\alpha}}(x^n e^{-x^\alpha/2}) = c_{56}e^{c_{57}n}n^{-n/\alpha}(c_{54}n^{1/\alpha})^n e^{-(1/2)c_{54}^{1/\alpha}n} =$$

$$= c_{16}\exp\left\{\left(-\frac{1}{2}c_{54}^{1/\alpha}+\log c_{54}+c_{57}\right)n\right\}.$$

Taking c_{54} great enaugh the factor of n in the last exponent becomes negative so that (4.12) holds. ∎

THEOREM 4.4. *We have for $\alpha>2$ and every real β taking $\gamma=\frac{1}{\alpha}$*

$$\lambda_n^{-1}(v_{\alpha\beta}^2;x) \leqq c_{58}n^{1-\gamma}v_{\alpha\beta}^{-2}(x) \qquad (-\infty<x<\infty). \tag{4.15}$$

PROOF. By (3.14) and Theorem 2.6 we have for $|x|\leqq c_{23}n^{1/\alpha}$

$$\lambda_n(v_{\alpha\beta}^2;x) \geqq \min_{P\in\mathfrak{p}_{n-1},\,P(x)=1}\int_{-c_{23}n^{1/\alpha}}^{c_{23}n^{1/\alpha}} P^2(t)v_{\alpha\beta}^2(t)\,dt \geqq$$

$$\geqq c_{21}^{-2}\min_{P\in\mathfrak{p}_{n-1},\,P(x)=1}\int_{-c_{23}n^{1/\alpha}}^{c_{23}n^{1/\alpha}}[P(t)p_{vn}(\alpha,\beta;t)]^2\,dt.$$

Since $P(x)=1$ and — by (2.19) — $p_nP\in\mathfrak{p}_{2n-1}$ attains for $t=x$ a value not greater then $c_{22}v_{\alpha\beta}(x)$ we have

$$\lambda_n(v_{\alpha\beta}^2;x) \geqq c_{21}^{-2}[P(x)p_n(\alpha,\beta;x)]^2\times \tag{4.16}$$

$$\times\min_{P\in\mathfrak{p}_{n-1},\,P(x)=1}\int_{-c_{23}n^{1/\alpha}}^{c_{23}n^{1/\alpha}}\left[\frac{P(t)p_n(\alpha,\beta;t)}{P(x)p_n(\alpha,\beta;x)}\right]^2 dt \geqq$$

$$\geqq c_{12}^{-2}c_{22}^2v_{\alpha\beta}^2(x)\min_{q\in\mathfrak{p}_{2n-1},\,q(x)=1}\int_{-c_{23}n^{1/\alpha}}^{c_{23}n^{1/\alpha}}q^2(t)\,dt =$$

$$= c_{21}^{-2}c_{22}^2c_{23}v_{\alpha\beta}^2(x)n^{1/\alpha}\min_{\substack{r\in\mathfrak{p}_{2n-1}\\ r(x/c_{23}n^{1/\alpha})=1}}\int_{-1}^{1}r^2(t)\,dt = c_{59}n^{1/\alpha}\lambda_{2n}(w_1;x/c_{23}n^{1/\alpha})$$

where $w_1(x)=1$ for $x\in[-1,1]$ and $w_1(x)=0$ for $x\in[-1,1]$. By [9], Theorem III. 3.3 we get from (4.16)

$$\lambda_n(v_{\alpha\beta}^2;x) \geqq c_{60}v_{\alpha\beta}^2(x)n^{-1+\gamma} \qquad \left(|x|\leqq\frac{1}{2}c_{23}n^{1/\alpha},\gamma=\frac{1}{\alpha}\right). \tag{4.17}$$

Since for fixed x, $\lambda_n(v_{\alpha\beta}^2; x)$ is a decreesing function of n (see (3.13)) we infer from (4.17) that for every $D>1$ with $\gamma = \frac{1}{\alpha}$

(4.18)
$$\lambda_n^{-1}(v_{\alpha\beta}^2; x) \leqq \lambda_{[Dn]}^{-1}(v_{\alpha\beta}^2; x) \leqq c_{60} D^{1-\gamma} n^{1-(1/\alpha)} v_{\alpha\beta}^2(x) \qquad \left(|x| \leqq \frac{1}{2} c_{23} D^{1/\alpha} n^{1/\alpha}\right).$$

Fixing D so great that $\frac{1}{2} c_{23} D^{1/\alpha} > c_{51}$ we see by (4.17) that (4.14) holds for $|x| \leqq c_{51} n^{1/\alpha}$. If $|x| > c_{51} n^{1/\alpha}$ we infer by (3.13) and (4.12) that (4.14) is valid in this case too. ∎

Let $f(x)$ be a measurable function defined on $(-\infty, \infty)$ for which the orthogonal expansion

(4.19)
$$f(x) \sim \sum_{\nu=0}^{\infty} a_\nu(v_{\alpha\beta}^2; f) p_\nu(v_{\alpha\beta}^2; x),$$

$$a_\nu(v_{\alpha\beta}^2; f) = \int_{-\infty}^{\infty} f(t) p_\nu(v_{\alpha\beta}^2; t) v_{\alpha\beta}^2(t)\, dt$$

is well defined; this is certainly the case if $f v_{\alpha\beta} \in L_p$ for some $1 \leqq p \leqq \infty$ (these will be the cases we need). We consider the partial sums

(4.20)
$$s_m(v_{\alpha\beta}^2; f; x) = \sum_{\nu=0}^{m-1} a_\nu(v_{\alpha\beta}^2; f) p_\nu(v_{\alpha\beta}^2; x)$$

of (4.18).

THEOREM 4.5. *We have for every f safisfying $v_{\alpha\beta} f \in L_\infty$*

(4.21)
$$v_{\alpha\beta}(x) n^{-1} \sum_{m=1}^{n} |s_m(v_{\alpha\beta}^2; f; x)| \leqq c_{61} \|v_{\alpha\beta} f\|_\infty.$$

PROOF. Since the idea is exactly the same as in [9], § IV. 3 and [7] Lemma 6 we give here a short scatch. Let x be fixed, $\delta_n = n^{-1+1/\alpha}$, $f_n(t) = f(t)$ for $|t-x| \leqq \delta_n$ $f_n(t) = 0$ for $|x-t| > \delta_n$ and $F_n(t) = (t-x)^{-1}[f(t) - f_n(t)]$. By the Christoffel—Darboux formula

(4.22)
$$K_m(v_{\alpha\beta}^2; x, t) = \sum_{\nu=0}^{m-1} p_\nu(v_{\alpha\beta}^2; x) p_\nu(v_{\alpha\beta}^2; t) =$$

$$= \frac{\gamma_{m-1}(v_{\alpha\beta}^2)}{\gamma_m(v_{\alpha\beta}^2)} \frac{p_{m-1}(v_{\alpha\beta}^2; x) p_m(v_{\alpha\beta}^2; t) - p_m(v_{\alpha\beta}^2; x) p_{m-1}(v_{\alpha\beta}^2; t)}{t-x}.$$

We have by Schwarz's inequality for every $m \leqq n$ in applying (4.15)

(4.23)
$$|s_m(v_{\alpha\beta}^2; f_n; x)| = \left| \int_{x-\delta_n}^{x+\delta_n} f(t) K_m(v_{\alpha\beta}^2; x, t) v_{\alpha\beta}^2(t)\, dt \right| \leqq$$

$$\leqq \left\{ \int_{x-\delta_n}^{x+\delta_n} f^2(t) v_{\alpha\beta}^2(t)\, dt \int_{-\infty}^{\infty} K_m^2(v_{\alpha\beta}^2; x, t) v_{\alpha\beta}^2(t)\, dt \right\}^{1/2} \leqq$$

$$\leqq 2^{1/2} \delta_n^{1/2} \|f v_{\alpha\beta}\|_\infty \lambda_m^{-1/2}(v_{\alpha\beta}^2; x) \leqq 2^{1/2} c_{58} v_{\alpha\beta}^{-1}(x) \|f v_{\alpha\beta}\|_\infty \qquad (m = 1, 2, \ldots, n).$$

By (4.21) we have

$$s_m(v_{\alpha\beta}^2; f-f_n; x) = \tag{4.24}$$

$$= \frac{\gamma_{m-1}(v_{\alpha\beta}^2)}{\gamma_m(v_{\alpha\beta}^2)} [p_{m-1}(v_{\alpha\beta}^2; x) a_m(v_{\alpha\beta}^2; F_n) - p_m(v_{\alpha\beta}^2; x) a_{m-1}(v_{\alpha\beta}^2; F_n)].$$

From (4.24), (4.4), (3.13) and (4.15) we infer again by Schwarz's inequality

$$\sum_{m=1}^{n} |s_m(v_{\alpha\beta}^2; f-f_n; x)| \leqq 2c_{47} n^{1/\alpha} \times \tag{4.25}$$

$$\times \left\{ c_{58} n^{1-(1/\alpha)} v_{\alpha\beta}^{-2}(x) \sum_{\nu=0}^{n} a_\nu^2(v_{\alpha\beta}^2; F_n) \right\}^{1/2}.$$

By Bessel's inequality

$$\sum_{\nu=0}^{n} a_\nu^2(v_{\alpha\beta}^2; F_n) \leqq \left(\int_{-\infty}^{x-\delta_n} + \int_{x+\delta_n}^{\infty} \right) \frac{F_n^2(t) v_{\alpha\beta}^2(t)}{(x-t)^2} dt \leqq \tag{4.26}$$

$$\leqq [\|F_n v_{\alpha\beta}\|_\infty]^2 2 \int_{\delta_n}^{\infty} \frac{dt}{t^2} = 2\delta_n^{-1} [\|F_n v_{\alpha\beta}\|_\infty]^2.$$

Now (4.21) follows from (4.23), (4.25) and (4.26). ∎

LEMMA 4.6. *The Fejér means*

$$\sigma_n(v_{\alpha\beta}^2; f; x) = \frac{1}{n} \sum_{m=1}^{[n} s_m(v_{\alpha\beta}^2; f; x) \tag{4.27}$$

of the orthogonal expansion (4.19) *satisfy for every fixed* $\alpha > 2$ *and real* β, *for every* $1 \leqq p \leqq \infty$

$$\|v_{\alpha\beta} \sigma_n(v_{\alpha\beta}^2; f)\|_p \leqq c_{62} \|v_{\alpha\beta} f\|_p. \tag{4.28}$$

PROOF. For $p = \infty$ this follows from Theorem 4.5. For the cases $1 \leqq p \leqq \infty$ it is deduced from the special case $p = \infty$ by the argument applied in our earlyer papers (see e.g. Theorem 3 in [7]).

LEMMA 4.7. *Let* F *be a primitive function of* f *and for fixed* $\alpha > 2$ *and real* β *let* $v_{\alpha\beta} f \in L_p$ *then*

$$\mathscr{E}_n^p(v_{\alpha\beta}; F) \leqq c_{63} n^{-1+1/\alpha} \|v_{\alpha\beta} f\|_p \qquad (n = 1, 2, \ldots). \tag{4.29}$$

PROOF. For $p = 1$ this is a consequence of Theorem 3.7. For $1 < p \leqq \infty$ we apply the same argument as in [1], [4], [7] so we can restrict us to a proof scatch. First we consider the case $p = \infty$.

We set

$$v_{2n}(v_{\alpha\beta}^2; f) = 2\sigma_{2n}(v_{\alpha\beta}^2; f) - \sigma_n(v_{\alpha\beta}^2; f).$$

Then (4.28) implies

$$\|v_{2n}(v_{\alpha\beta}^2; f) v_{\alpha\beta}\|_p \leqq 3c_{62} \|f v_{\alpha\beta}\|_p \qquad (1 \leqq p \leqq \infty). \tag{4.30}$$

Since $(f-v_{2n})v_{\alpha\beta}$ is by construction orthogonal to every $qv_{\alpha\beta}$ with $q\in\mathfrak{P}_n$ we have for every $q\in\mathfrak{P}_n$

$$F(x)-F(0)-\int_0^x v_{2n}(v_{\alpha\beta}^2;f,t)\,dt = \tag{4.31}$$

$$= \int_{-\infty}^{\infty}[f(t)-v_{2n}(v_{\alpha\beta}^2;f;t)][\Phi_x(t)-q(t)]v_{\alpha\beta}^2(t)\,dt$$

where $\Phi_x(t)=v_{\alpha\beta}^{-2}(t)$ for $t\in[0,x]$ and $\Phi_x(t)=0$ for $t\notin[0,x]$. We infer from (4.31) applying Theorem 3.7 and (4.30), taking $\gamma=\dfrac{1}{\alpha}$

$$\left|F(x)-F(0)-\int_0^x v_{2n}(v_{\alpha\beta}^2;f;t)dt\right| \leqq$$

$$\leqq \|[f-v_{2n}(v_{\alpha\beta}^2;f)]v_{\alpha\beta}\|_\infty \inf_{q\in\mathfrak{P}_n}\|(\Phi_x-q)v_{\alpha\beta}\| \leqq$$

$$\leqq c_{64}\|fv_{\alpha\beta}\|_\infty n^{-1+\gamma}\int_{-\infty}^{\infty} v_{\alpha\beta}(\xi)|d\Phi_x(\xi)| \leqq c_{65}n^{-1+\gamma}\|fv_{\alpha\beta}\|_\infty v_{\alpha\beta}^{-1}(x).$$

This proves (4.29) for $p=\infty$. After it is established for $p=1$ and $p=\infty$ we obtain (4.29) for the remaining cases $1<p<\infty$ by the Riesz—Thorin interpolation theorem. (See e.g. [4].) ∎

LEMMA 4.8. *We have under the conditions of Lemma* 4.7

$$\mathscr{E}_n^{(p)}(v_{\alpha\beta},F) \leqq c_{63}n^{-1+\gamma}\mathscr{E}_{n-1}(v_{\alpha\beta},f) \qquad \left(1\leqq p\leqq\infty;\ n=1,2,\ldots;\gamma=\frac{1}{\alpha}\right). \tag{4.32}$$

PROOF. Let $q\in\mathfrak{P}_{n-1}$, $Q(x)=\int q(t)dt$ then by (4.29)

$$\mathscr{E}_n^{(p)}(v_{\alpha\beta};F) = \mathscr{E}_n^{(p)}(v_{\alpha\beta};F-Q) \leqq c_{63}n^{-1+\gamma}\|v_{\alpha\beta}(f-q)\|_p. \tag{4.33}$$

Since (4.33) is valid for every $q\in\mathfrak{P}_{n-1}$ it follows (4.32). ∎

5. Proof of the direct approximation theorems

Let $v_{\alpha\beta}f\in L_p$ for some $1\leqq p\leqq\infty$ and let

$$\Phi_n(x)=\begin{cases}v_{\alpha\beta}(x)f(x) & (|x|\leqq n^{1/\alpha})\\ 0 & (|x|>n^{1/\alpha})\end{cases}. \tag{5.1}$$

We define $h_n=n^{-1+\gamma}$ $(\gamma=1/\alpha)$ and

$$f_n(x)=v_{\alpha\beta}^{-1}(x)h_n^{-1}\int_x^{x+h_n}\Phi_n(t)\,dt. \tag{5.2}$$

LEMMA 5.1. *We have for arbitrary* $\alpha>1$ *and real* β

$$\|v_{\alpha\beta}(f-f_n)\|_p \leqq \omega(L_p, v_{\alpha\beta}; f, h_n) \tag{5.3}$$

and

$$\|v_{\alpha\beta}f_n'\|_p \leqq c_{64}h_n^{-1}\omega(L_p, v_{\alpha\beta}; f, h_n). \tag{5.4}$$

PROOF. Here and in what follows we repeatedly apply Minkowski's inequality:

$$\|(f-f_n)v_{\alpha\beta}\|_p = \left\| h_n^{-1}\int_x^{x+h_n} [\Phi_n(t)-v_{\alpha\beta}(x)f(x)]\,dt \right\|_p \leqq \tag{5.5}$$

$$\leqq \left\| h_n^{-1}\int_x^{x+h_n} [v_{\alpha\beta}(t)f(t)-v_{\alpha\beta}(x)f(x)] \right\|_p +$$

$$+\left\| h_n^{-1}\int_x^{x+h_n} [\Phi_n(t)-v_{\alpha\beta}(t)f(t)]\,dt \right\|_p \leqq \sup_{0\leqq h\leqq h_n} \|v_{\alpha\beta}(x+h)f(x+h)-v_{\alpha\beta}(x)f(x)\|_p +$$

$$+\|\Phi_n(x)-v_{\alpha\beta}(x)f(x)\|_p.$$

Now by (5.1) $\Phi_n-v_{\alpha\beta}f$ vanishes for $|x|\leqq n^{1/\alpha}$ and is equal to $-v_{\alpha\beta}f$ for $|x|>n^{1/\alpha}$. Since we have by definition $\tau(h_n^{1/(\alpha-1)}x)=\tau(n^{-1/\alpha}x)=1$ for $|x|\geqq n^{1/\alpha}$ we see that

$$\|\Phi_n-v_{\alpha\beta}f\|_p \leqq \|[\tau(h_n^{1/(\alpha-1)}x)]^{\alpha-1}v_{\alpha\beta}(x)f(x)\|_p. \tag{5.6}$$

The first part of our lemma, i.e. (5.3) is a consequence of (5.5) and (5.6). We get from (5.2) by differentiation

$$v_{\alpha\beta}f_n' = \frac{|x|^\alpha}{x}\left[\frac{\alpha}{2}-\frac{\beta}{2}(1+|x|^\alpha)^{-1}\right] h_n^{-1}\int_x^{x+h_n}\Phi_n(t)\,dt +$$

$$+h_n^{-1}[\Phi_n(x+h_n)-\Phi_n(x)].$$

The first term vanishes if $|x|>n^{1/\alpha}$ and $|x+h_n|>n^{1/\alpha}$ and this is certainly the case for $|x|>2n^{1/\alpha}$; now if $|x|<2n^{1/\alpha}$ we have

$$\frac{|x|^\alpha}{x}\left[\frac{\alpha}{2}-\frac{\beta}{2}(1+|x|^\alpha)^{-1}\right] \leqq c_{65}|x|^{\alpha-1} \leqq c_{66}h_n^{-1}$$

i.e.

$$\|v_{\alpha\beta}f_n'\|_p \leqq c_{65}\||x|^{\alpha-1}\Phi_n(x)\|_p + c_{66}h_n^{-1}\left\| h_n^{-1}\int_0^{h_n}\Phi_n(x+t)-\Phi_n(x)\right\|_p + \tag{5.7}$$

$$+h_n^{-1}\|\Phi_n(x+h)-\Phi_n(x)\|_p \leqq$$

$$\leqq c_{65}\||x|^{\alpha-1}\Phi_n(x)\|_p + (2c_{66}+2)h_n^{-1}\|\Phi_n(x)-v_{\alpha\beta}(x)f(x)\|_p +$$

$$+(c_{66}+1)h_n^{-1}\sup_{0\leqq t\leqq h_n}\|v_{\alpha\beta}(x+t)f(x+t)-v_{\alpha\beta}(x)f(x)\|_p.$$

Now the second factor of $|x|^{\alpha-1}\Phi_n(x)$ vanishes for $|x|>n^{1/\alpha}$ and for $|x|\leqq n^{1/\alpha}$ the first factor is equal to $h_n^{-1}[\tau(h_\alpha^{1/(\alpha-1)}x)]^{\alpha-1}$ so that

$$|x|^{\alpha-1}\Phi_n(x)| = h_n^{-1}[\tau(h_n^{1/(1-\alpha)}x)]^{\alpha-1}|\Phi_n(x)| \leqq \tag{5.8}$$

$$\leqq h_n^{-1}[\tau(h_n^{1/(\alpha-1)}x)]^{\alpha-1}v_{\alpha\beta}(x)|f(x)|.$$

Now (5.4) is implied by (5.7), (5.8) and (5.6). ∎

We define now — as announced in the introduction — the generalized Zygmund modulus in the following way:

$$\omega_2(L_p, v_{\alpha\beta}; f, \delta) = \sup_{0\leqq t\leqq\delta}\left\| v_{\alpha\beta}\left(x+\frac{t}{2}\right)f\left(x+\frac{t}{2}\right)- \right. \tag{5.9}$$

$$\left. -2v_{\alpha\beta}(x)f(x)+v_{\alpha\beta}\left(x-\frac{t}{2}\right)f\left(x-\frac{t}{2}\right)\right\|_p +$$

$$+\sup_{0\leqq t\leqq\delta}\|[\tau(\delta^{1/(\alpha-1)}x)]^{\alpha-1}[v_{\alpha\beta}(x+t)f(x+t)-v_{\alpha\beta}(x)f(x)\|_p +$$

$$+\|[\tau(\delta^{1/(\alpha-1)}x)]^{2(\alpha-1)}v_{\alpha\beta}(x)f(x)\|_p+\delta^2\|x^{\alpha-2}v_{\alpha\beta}(x)f(x)\|_{L_p[-1,1]}.$$

Here $\|\cdot\|_{L_p[-1,1]}$ denotes the norm in $L_p[-1, 1]$.

We consider the modified auxiliary function

$$f_n^*(x) = v_{\alpha\beta}^{-1}(x)h_n^{-2}\int_{x-h_n/2}^{x+h_n/2}\int_{x-h_n/2}^{x+h_n/2}\Phi_n\left(\frac{u+v}{2}\right)du\,dv. \tag{5.10}$$

LEMMA 5.2. *We have for* $\alpha\geqq 2$ *and every real* β

$$\|v_{\alpha\beta}(f-f_n^*)\|_p \leqq \omega_2(L_p, v_{\alpha\beta}; f, h_n) \tag{5.11}$$

and

$$\|v_{\alpha\beta}f_n^{*\prime\prime}\|_p \leqq c_{67}h_n^{-2}\omega_2(A_p, v_{\alpha\beta}; f, h_n). \tag{5.12}$$

PROOF. In order to prove (5.11) we rewrite (5.10) into the form

$$v_{\alpha\beta}(x)f_n^*(x) = 4h_n^{-2}\int_0^{h_n/2}\left(\frac{h_n}{2}-t\right)[\Phi_n(x+t)+\Phi_n(x-t)]\,dt$$

and infer

$$\|v_{\alpha\beta}(f-f_n^*)\|_p \leqq \|v_{\alpha\beta}(f-\Phi_n)\|_p + \tag{5.13}$$

$$+\sup_{0\leqq t\leqq h_n/2}\|v_{\alpha\beta}(x+t)f(x+t)-2v_{\alpha\beta}(x)f(x)+v_{\alpha\beta}(x-t)f(x-t)\|_p.$$

Since $f(x)-\Phi_n(x)=0$ for $|x|\leqq n^{1/\alpha}$ and $\tau(h_n^{1/(\alpha-1)}x)=1$ for $|x|\geqq n^{1/\alpha}$ we have

$$\|v_{\alpha\beta}(f-\Phi_n)\|_p \leqq \|[\tau(h_n^{1/(\alpha-1)}x)]^{2(\alpha-1)}v_{\alpha\beta}(x)f(x)\|_p. \tag{5.14}$$

Combining (5.13) and (5.14) we see that (5.11) holds.

To prove (5.12) let us observe that

$$(5.15)\qquad (v_{\alpha\beta}f_n^*)'(x) = 2h_n^{-2}\int_{-h_n/2}^{h_n/2}\left[\Phi_n\left(x+\frac{t+h_n/2}{2}\right)+\Phi_n\left(x+\frac{t-h_n/2}{2}\right)\right]dt,$$

$$(5.16)\qquad (v_{\alpha\beta}f_n^*)''(x) = h_n^{-2}[\Phi_n(x+h_n/2)-2\Phi_n(x)+\Phi_n(x-h_n/2)]$$

and by elementary calculation

$$(5.17)\qquad |(v_{\alpha\beta}[v_{\alpha\beta}^{-1}]')(x)| \leqq c_{68}|x|^{\alpha-1}$$

$$(5.18)\qquad |(v_{\alpha\beta}[v_{\alpha\beta}^{-1}]'')(x)| \leqq c_{69}|x|^{2\alpha-2}+c_{70}|x|^{\alpha-2}.$$

We infer

$$(5.19)\qquad |v_{\alpha\beta}f_n^{*\prime\prime}| \leqq |v_{\alpha\beta}(v_{\alpha\beta}^{-1})''v_{\alpha\beta}f_n^*|+2|v_{\alpha\beta}(v_{\alpha\beta}^{-1})'(v_{\alpha\beta}f_n^*)'|+|(v_{\alpha\beta}f_n^*)''| \leqq$$

$$\leqq (c_{69}|x|^{2\alpha-2}+c_{70}|x|^{\alpha-2})|v_{\alpha\beta}f_n^*|+2c_{68}|x|^{\alpha-1}|(v_{\alpha\beta}f_n^*)'|+|(v_{\alpha\beta}f_n^*)''|.$$

To estimate the L_p-norm of the first term in (5.19) let us observe that contribution of the interval $|x|\leqq 1$ is dominated by

$$2c_{70}\|x^{\alpha-2}v_{\alpha\beta}(x)f(x)\|_{L_p[-1,1]}.$$

For $1\leqq|x|\leqq n^{1/\alpha}$ we estimate this term applying

$$c_{69}|x|^{2\alpha-2}+c_{70}|x|^{\alpha-2} \leqq (c_{69}+c_{70})|x|^{2\alpha-2} \leqq h_n^{-2}[\tau(h_n^{1/(\alpha-1)}x)]^{2(\alpha-1)} \qquad (1 \leqq |x| \leqq n^{1/\alpha})$$

and for $|x|\geqq n^{1/\alpha}$ we have $f_n^*=0$. The estimate of the last term follows from (5.16). To estimate the last but one term in (5.19) let us observe that by (5.15), $(v_{\alpha\beta}f_n^*)'$ vanishes for $|x|\geqq 2n^{1/\alpha}\geqq n^{1/\alpha}+h_n/2$ and for $|x|\leqq 2n^{1/\alpha}$ we have

$$|x|^{\alpha-1} \leqq c_{71}h_n^{-1}[\tau(h_n^{1/(\alpha-1)}x)]^{\alpha-1}.$$

The further technical details of the estimation may be left to the reader. ∎

We are now ready to formulate and prove the main theorem of this paper:

THEOREM 5.3. *Let $\alpha>2$, β arbitrary real, $1\leqq p\leqq\infty$ and let f have (in the L_p-sense) an r-th order derivative $f^{(r)}$ satisfying*[4] $v_{\alpha\beta}f^{(r)}\in L_p$ *then for* $\gamma=\dfrac{1}{\alpha}$ *we have*

$$(5.20)\qquad \mathscr{E}_n^{(p)}(v_{\alpha\beta};f) \leqq c_{72}e^{c_{73}r}n^{-r(1-(1/\alpha))}\omega(L_p, v_{\alpha\beta};f, n^{-1+(1/\alpha)}) \qquad (n>r+2)$$

and

$$(5.21)\qquad \mathscr{E}_n^{(p)}(v_{\alpha\beta};f) \leqq c_{73}e^{c_{74}r}n^{-r(1-(1/\alpha))}\omega_2(L_p, v_{\alpha\beta};f, n^{-1+(1/\alpha)}) \qquad (n>r+2).$$

[4] Compare footnotes [1] and [2], section 1.

PROOF. By Lemma 4.8 it is sufficient to consider the case $r=0$. Combining Lemma 5.1 and Lemma 4.7

$$\mathscr{E}_n^{(p)}(v_{\alpha\beta};f) \leqq \|(f-f_n)v_{\alpha\beta}\|_p + \mathscr{E}_n^{(p)}(v_{\alpha\beta};f_n) \leqq$$

$$\leqq \|(f-f_n)v_{\alpha\beta}\|_p + c_{63}n^{-1+\gamma}\|v_{\alpha\beta}f_n'\|_p \leqq (1+c_{63}c_{64})\omega(L_p, v_{\alpha\beta};f, n^{-1+\gamma})$$

so that (5.20) is proved.

Combining Lemma 5.2, Lemma 4.8 and Lemma 4.7

$$\mathscr{E}_n^{(p)}(v_{\alpha\beta};f) \leqq \|(f-f_n^*)v_{\alpha\beta}\|_p + \mathscr{E}_{n-1}^{(p)}(v_{\alpha\beta};f_n^*) \leqq$$

$$\leqq \|(f-f_n^*)v_{\alpha\beta}\| + c_{63}^2 n^{-1+\gamma}(n-1)^{-1+\gamma}\|v_{\alpha\beta}f_n^{*\prime\prime}\|_p \leqq c_{73}\omega_2(L_p, v_{\alpha\beta};f, n^{-1+\gamma})$$

and this proves (5.21). ∎

REFERENCES

[1] FREUD, G.: On approximation with weight by polynomials on the real line *Dokladii Ak. Nauk' SSSR* **191** (1970) 293—294. English translation: *Soviet Math. Dokladii* **11** (1970)' 370—371.

[2] FREUD, G.: On an inequality of Markov type. *Dokladii Ak. Nauk SSSR* **197** (1971) 790—793. English translation: *Soviet Math. Dokladii* **12** (1971) 570—573.

[3] FREUD, G.: On approximation with the weight $e^{-x^2/2}$ by polynomials. *Dokladii Ak. Nauk SSSR* **201** (1971) 1292—1294.

[4] FREUD, G.: *A contribution to the problem of weighted polynomial approximation.* ISNM vol. 20. „Linear operators and approximation theory" Conference report (Oberwolfach, 1970). Editors: P. J. Butzer, J. P. Kahane and B. Sz.-Nagy. Birkhäuser Verl. Basel 1972.

[5] FREUD, G.: Ond direct and converse theorems in the theory of weighted polynomial approximation. *Math. Zeitschr.* **134** (1972) 123—134.

[6] FREUD, G.: On weighted L_1-approximation by polymials. *Studia Math.* **46** (1973).

[7] FREUD, G.: On polynomial approximation with the weight $\exp\{-1/2\, x^{2k}\}$. *Acta Math. Acad. Sci. Hung.*, **24** (1973), 363—371.

[8] FREUD, G.: On converse theorems of weighted polynomial approximation. *Acta Math. Acad. Sci. Hung.*, **24** (1973), 389—397.

[9] FREUD, G.: *Orthogonale Polynome.* Birkhäuser Verl. Basel 1969. English translation by I. Földes, Pergamon Press 1971.

[10] FREUD, G.: On the greatest zero of an orthogonal polynomial. I. *Acta Sci. Math.* (Szeged), in print.

[11] NÉVAI, G. P.: Some properties of polynomials orthogonal to the weight $(1+x^{2k})^{\alpha}e^{-x^{2k}}$ and their application to approximation theory. (In Russian.) *Dokladii Ak. Nauk SSSR*, **24** (1973), 335—342.

[12] NÉVAI, G. P.: Polynomials orthogonal on the real axis to the weight $|x|^{\alpha}e^{-|x|^{\beta}}$. (In Russian.) *Acta Math. Ac. Sci. Hung.*, in print.

[13] NIKOLSKI, S. M.: Approximation of functions by trigonometric polynomials in the mean. (In Russian.) *Isvestia Ak. Nauk SSSR*, ser. matem. **10** (1946) 207—256.

[14] SZEGŐ, G.: *Orthogonal Polynomials.* Amer. Math. Soc. Colloquium Publ. vol. XXIII, revised edition. New York 1959.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received December 21, 1972)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 307—308.

A NOTE ON THE HAMILTONIAN INDEX OF GRAPHS

S. F. KAPOOR*
and
M. J. STEWART

For any connected graph G, which is not a path, the *hamiltonian index* $h(G)$ is defined to be the smallest non-negative integer n such that the nth iterated line-graph $L^n(G)$ is hamiltonian. In [1] it was shown that (i) if G is a connected graph of order p and G is not a path, then $h(G) \leqq p-3$, and (ii) the bound $p-3$ is sharp. The concept of the hamiltonian index was studied further in [2] where $h(T)$ was determined for all trees T which are not paths, and some improved bounds were established for connected graphs containing cycles.

It is known that every 2-connected nonhamiltonian graph contains a subgraph homeomorphic from the complete bipartite graph $K(2, 3)$. Furthermore, all graphs homeomorphic from $K(2, 3)$, or more generally, homeomorphic from $K(2, n)$, $n \geqq 3$, are nonhamiltonian. Consequently, for such graphs G, $h(G) > 0$. In this note the hamiltonian index of all graphs which are homeomorphic from $K(2, n)$ is completely determined. Some further results dealing with homeomorphisms between graphs and associated concepts appear in [4]. We will follow the terminology of [3].

In [1], a graph G is called *sequential* if its q edges, $q \geqq 3$, can be ordered as $e_1, e_2, \ldots, e_q, e_{q+1} = e_1$ such that consecutive edges in the sequence are adjacent. The following results from [1] are noted for later use.

THEOREM A. The line-graph $L(G)$ of a graph G is hamiltonian if and only if G is sequential.

THEOREM B. If G is a graph containing a cycle C such that every edge of G is incident with at least one vertex of C, then G is sequential.

THEOREM. *Let G be any graph homeomorphic from $K(2, n)$, $n \geqq 3$. Let u and v be the vertices of G having degree n and let $d(u, v)$ be the distance between u and v. Then*

$$h(G) = \begin{cases} 1 & \text{if } n \text{ is even,} \\ d(u, v) - 1 & \text{otherwise.} \end{cases}$$

PROOF. If n is even and G is homeomorphic from $K(2, n)$, then G is non-hamiltonian. But we observe that in this case G is sequential, and by Theorem A, $h(G) = 1$.

Now let $n \geqq 3$ be odd. We will use induction on $d(u, v) \geqq 2$. For $i = 1, 2, \ldots, n$, let the edges incident with u and v be respectively labeled x_i, y_i such that there exist

* Research partially supported by a Faculty Research Fellowship from Western Michigan University.

$u-v$ paths $P_i=u, x_i, \ldots, y_i, v$ for each i. Moreover, let the length of the path P_1 be $d(u, v)$. The line-graph $L(G)$ consists of two copies B_1, B_2 of the complete graph K_n with vertex sets $V(B_1)=\{x_1, x_2, \ldots, x_n\}$ and $V(B_2)=\{y_1, y_2, \ldots, y_n\}$ such that there exists a x_i-y_i path P_i' in $L(G)$ for each i. We note that the length of P_i' is one less than that of P_i.

If $d(u, v)=2$, then x_1 and y_1 are adjacent in $L(G)$. But then $h(G)=1=d(u, v)-1$ since $L(G)$ contains a hamiltonian cycle obtained by proceeding along the following sequence of vertices and paths:

$$x_1, x_2, P_2', y_2, y_1, y_3, P_3', x_3, x_4, \ldots, y_n, P_n', x_n, x_1.$$

Assume that if G is any graph homeomorphic from $K(2, n)$ with $n(\geqq 3)$ odd and $2\leqq d(u, v)<k$, then $h(G)=d(u, v)-1$. Let G be a homeomorph of $K(2, n)$ such that $d(u, v)=k\geqq 3$. Following the notation used above, let G^* denote the graph obtained from G by contracting exactly one edge in the path P_1. Then $2\leqq d(u, v)=k-1$ in G^*, and by the induction hypothesis we obtain $h(G^*)=k-2$ so that $L^{k-2}(G^*)$ is a hamiltonian graph.

The graph $L^{k-2}(G^*)$ consists of isomorphic subgraphs B_1 and B_2 (not necessarily complete) such that n vertices in B_1 are joined to the corresponding vertices in B_2 by n vertex and edge disjoint paths P_i^*, with the length of P_1^* being 1. Let e denote the edge in P_1^*. If this edge e in $L^{k-2}(G^*)$ is replaced by a path P_1' of length 2, we obtain the graph $L^{k-2}(G)$. Denote the other corresponding paths in $L^{k-2}(G)$ by P_i', and note that $P_i^*=P_i'$ for $2\leqq i\leqq n$.

We observe that no hamiltonian cycle in $L^{k-2}(G^*)$ contains the edge e since such a cycle would have to contain all the other paths $P_2^*, P_3^*, \ldots, P_n^*$ joining the corresponding vertices from B_1 and B_2, and this is seen to be impossible as n is odd, and P_i^* has length at least 2 for $2\leqq i\leqq n$. Similarly, we conclude that $L^{k-2}(G)$ is nonhamiltonian because any hamiltonian cycle in $L^{k-2}(G)$ must necessarily contain the n paths $P_1', P_2', \ldots, P_n'$ joining B_1 and B_2, and this is again impossible since n is odd, and for each $i\geqq 1$, P_i' has length at least 2.

Let C be a hamiltonian cycle in $L^{k-2}(G^*)$. Then the edge e does not lie on C. Replacing e by the path P_1' and obtaining $L^{k-2}(G)$, we note that the graph $L^{k-2}(G)$ contains a cycle C such that every edge of this graph is incident with at least one vertex of C. In view of Theorem B, the graph $L^{k-2}(G)$ is sequential, and by Theorem A, $L(L^{k-2}(G))=L^{k-1}(G)$ is hamiltonian. Thus $h(G)=k-1$.

REFERENCES

[1] CHARTRAND, G.: On hamiltonian line-graphs, *Trans. Amer. Math. Soc.* **134** (1968). 559—566.
[2] CHARTRAND, G. and WALL, C. E.: On the hamiltonian index of a graph. *Studia Sci. Math. Hungar.*, to appear.
[3] HARARY, F.: *Graph Theory,* Addison-Wesley, Reading, Mass., 1969.
[4] STEWART, M. J.: *Graphs and their Subdivisions,* Ph. D. thesis, Michigan State University, 1972.

Western Michigan University, Lansing Community College

(Received January 10, 1973)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 309—322

STOCHASTIC APPROXIMATION FOR FINDING LOCAL MAXIMA OF PROBABILITY DENSITIES

by

J. FRITZ

1. Introduction

Statistical analysis of large amounts of continuous data generally reduces to a more or less direct exploration of the underlying probability density. The most straightforward and radical approach to a total solution of such problems could be based on a good estimation of the density function; in the last few years several parametric and nonparametric methods of density estimation have been developed and discussed in details, see [10], [15] and [1]—[6] with further references.

Unfortunately, the fidelity of a parametric density estimation will be satisfactory enough only if some previous information is available on the form of the density function. On the other hand, the nonparametric estimates — the empirical density functions — have very nice asymptotic properties, but they are multidimensional numerical functions requiring large amounts of memory to compute and to store; thus their practical applicability is limited, too. Therefore, instead of a total solution of the problem, we must content ourselves with the extraction of certain characteristic features of the density function, the suitable choice of which depends on the given task. From the point of view of cluster analysis, the locations of dominant local maxima of the density function seem to be the most significant ones among them, as it is pointed out by K. S. FU [10] and D. WISHART [11]. Let us remark, that they propose cluster-seeking algorithms based on mode estimation techniques. The term 'mode' is used as an abbreviation for 'location of a local maximum of a probability density'.

In the unimodal case the most natural approach to mode estimation is to locate the absolute maximum of a corresponding empirical density function; E. PARZEN [2] proves consistency and asymptotic normality for such estimates. H. CHERNOFF [7] analyses a special version of this method, which can be based also on order statistics, as well as the estimates of U. GRENANDER [8] and of J. H. VENTER [9]; these estimates are consistent, too. Some of the above and similar methods can be extended to multimodal and multidimensional distributions (see [10], [11]), but the generalization is not immediate.

A common disadvantage of such mode estimations is the rapid increase in memory and computational requirements implied by a relatively moderate increase in sample size. In order to avoid these difficulties, we ought to use iterative algorithms. The aim of this paper is to introduce a stochastic approximation procedure that converges with probability one to one of the stationary points of the underlying density function. Despite the fact that the algorithm prefers local maxima rather then local minima or saddle points, its limit is not necessarily a mode. Of course, using randomly selected or roughly estimated initial values, any of the dominant modes can be located. Some questions of practical applicability and of cluster analysis are discussed, as well.

2. Multimodal stochastic approximation

Since we are dealing with multimodal density functions, we need stochastic approximation theorems for regression functions having several roots. The study of this problem has been initiated by V. FÁBIÁN [13], some further results and references can be found in [10], [15], [16] and [17]. Unfortunately, the known results in this direction are far from being as comprehensive as in usual stochastic approximation theory. Therefore, to obtain a satisfactory solution to the delicate problem of mode estimation, we must get rid of many restrictive assumptions on the regression function and on the procedure itself. Under some general conditions, a relatively simple proof is given for the convergence of the stochastic gradient algorithm in the case of a multimodal criterion function.

Let $f(x)$ be a differentiable function on the k-dimensional Euclidean space R, its gradient vector at the point $y \in R$ will be denoted by $U(y)$, while (x, y) and $\|y\|$ denote the scalar product and norm in R. Local maxima of the criterion function f can be determined by means of a gradient algorithm

$$y_{n+1} = y_n + \varrho_n U(y_n), \qquad \varrho_n \geqq 0, \quad n = 1, 2, \ldots;$$

provided that U is known. If the gradient vector $U(y)$ is not explicitely given, then we still have the possibility to use a stochastic gradient algorithm

$$Y_{n+1} = Y_n + \varrho_n Z_{n+1}, \qquad n = 1, 2, \ldots, \tag{1}$$

where Z_{n+1} is an estimate of $U(Y_n)$ good enough, if any.

More precisely, let $\mathscr{A}_1, \mathscr{A}_2, \ldots, \mathscr{A}_n, \ldots$ be an increasing sequence of σ-algebras in the overall probability space $(\Omega, \mathscr{A}, \mathsf{P})$, and suppose that the nonnegative, bounded random variables ϱ_n, and the k-dimensional random vectors Z_n are measurable on $\mathscr{A}_n$ and have finite second moments for each positive integer n. Then, given an arbitrary $\mathscr{A}_1$-measurable initial value $Y_1 \in R$ with $\mathsf{E}\|Y_1\|^2 < +\infty$, the iteration rule (1) produces a sequence of random vectors $Y_n \in R$, $n = 1, 2, \ldots$ such that Y_n is $\mathscr{A}_n$-measurable and $\mathsf{E}\|Y_n\|^2 < +\infty$. We shall investigate the convergence properties of this process under the following conditions on the form of the criterion function and on the procedure described in (1). Throughout this paper, E_n denotes conditional expectation given the σ-algebra $\mathscr{A}_n$, i.e. $\mathsf{E}_n X = \mathsf{E}(X|\mathscr{A}_n)$, and a $\sum$ without specifications means summation over the set of natural numbers, i.e. $\sum u_n = u_1 + u_2 + \ldots + u_n + \ldots$.

In assumptions A)—F), μ, M, r_0 denote finite, positive constants, $L(r)$ and $K(r)$ are positive, finite valued functions of $r > 0$, and y_0 is a fixed point of R; further $1 < \mu \leqq 2$.

A) $\sum \varrho_n = +\infty$ a.s., $\quad \sum \varrho_n^\mu < +\infty$ a.s.

B) $\sum \varrho_n (1 + \|Y_n - y_0\|)^{-1} \|\mathsf{E}_n Z_{n+1} - U(Y_n)\| < +\infty$ a.s.

C) $\sum \varrho_n^\mu (1 + \|Y_n - y_0\|^2)^{-1} \mathsf{E}_n \|Z_{n+1} - \mathsf{E}_n Z_{n+1}\|^2 < +\infty$ a.s.

D) $\|U(y)\|^2 \leqq M(1 + \|y - y_0\|^2)$ for each $y \in R$.

E) $(y - y_0, U(y)) \leqq 0$ if $\|y - y_0\| \geqq r_0$.

F) $f(x) \geqq f(y) + (x - y, U(y)) - L(r)\|x - y\|^\mu - K(r)\|x - y\|^2$ if $\|x - y_0\| \leqq r$, $\|y - y_0\| \leqq r$; for each $r > 0$.

These assumptions are very similar to the usual ones in stochastic approximation theory, especially if $\mu=2$. It is very important that Z_{n+1} need not be an unbiased estimate of $U(Y_n)$, since no unbiased estimation can be given for a probability density and its derivatives. This set of conditions can be weakened a little bit without any essential modification of the proof, but it seems to be general enough.

Let us remark, that F) may be replaced by the more natural, but stronger F′), as follows.

F′) For each $y \in R$ there exist positive numbers ε_y and L_y such that

$$\|U(x)-U(y)\| \leqq L_y \|x-y\|^{\mu-1} \quad \text{if} \quad \|x-y\| \leqq \varepsilon_y,$$

where L_y and $(\varepsilon_y)^{-1}$ remain bounded if y varies over a bounded subset of R.

Really, for any $r>0$ set

$$L(r) = \sup\{L_y; \|y-y_0\| \leqq r\},$$

$$\varepsilon(r) = \inf\{\varepsilon_y; \|y-y_0\| \leqq r\},$$

$$K(r) = \frac{1}{\varepsilon(r)}\sqrt{2M(1+r^2)},$$

then, in view of D) we have $L(r)<+\infty$, $K(r)<+\infty$, and

$$\|U(x)-U(y)\| \leqq L(r)\|x-y\|^{\mu-1}+K(r)\|x-y\|,$$

provided that $\|x-y_0\| \leqq r$, $\|y-y_0\| \leqq r$. On the other hand, since f is differentiable'

$$f(x) = f(y)+(x-y, U(y+\delta(x-y))) \tag{3}$$

holds for each couple x, y with a $\delta=\delta(x, y)$ between 0 and 1, whence F) follows by the Cauchy inequality and by (2); that is D) and F′) imply F).

As a consequence of E), we point out that $f(x)$ is bounded above, and takes on its maximum at a point z with $\|z-y\| \leqq r_0$; therefore the set of stationary points of f is not empty, as $U(z)=0$. This assertion follows immediately from the relation

$$f(x) \leqq \max\{f(y); \|y-y_0\| = r_0\} \quad \text{if} \quad \|x-y_0\| > r_0. \tag{4}$$

To prove (4), it suffices to put $y = y_0+(x-y_0)r_0\|x-y_0\|^{-1}$ in (3) and apply E).

The study of process (1) will be based on the following lemmas.

LEMMA 1. *If V_n is measurable on $\mathcal{A}_n$, $\mathsf{E}(V_n^2)<+\infty$, and*

$$\sum \mathsf{E}_n(V_{n+1}-\mathsf{E}_n V_{n+1})^2 < +\infty \ \ a.s.,$$

$$\lim_m \sum_{n=1}^{m} \mathsf{E}_n V_{n+1} \ \textit{finite a.s.},$$

then $S_m = \sum_{n=1}^{m} V_{n+1}$ converges also to a finite limit with probability one.

This statement is a special case of Proposition IV. 6.2. in J. NEVEU [19], while our second lemma is proved using the ideas of its proof in [19].

LEMMA 2. *Let* $V_n \geqq 0$ *be measurable on* $\mathscr{A}_n$, $V_1 = 0$, $\mathsf{E}(V_n) < +\infty$, *and* $\sum \mathsf{E}_n V_{n+1} < +\infty$ *a.s.; then* $\sum V_n < +\infty$ *a.s.*

PROOF. Given a positive integer N and a constant $a > 0$, consider the stopping rule $t(N) = \min\{m, N\}$, where $m = +\infty$ on the set

$$\Omega_a = \{\sum \mathsf{E}_n V_{n+1} \leqq a\},$$

and m is the first integer for which

$$\sum_{n=1}^{m} \mathsf{E}_n V_{n+1} > a,$$

in the opposite case. Then we have

$$\mathsf{E}\left(\sum_{n=1}^{t(N)} V_n\right) \leqq a \quad \text{for each} \quad N,$$

see the Lemma and the proof of Proposition IV. 6.2. in [19], whence

$$\lim_N \sum_{n=1}^{t(N)} V_n < +\infty \text{ a.s.}$$

follows by the Fatou—Lebesgue Theorem. However,

$$\lim_N \sum_{n=1}^{t(N)} V_n = \sum V_n \quad \text{on} \quad \Omega_a,$$

thus the assertion follows if we let a go to infinity. ◇

Lemma 3. is a slight modification of a lemma, due to J. H. VENTER [9]; for convenience we reproduce here the easy proof.

LEMMA 3. *Suppose that the sequence* $w_1, w_2, \ldots, w_n, \ldots$ *of nonnegative numbers satisfies*

$$w_{n+1} \leqq \max\{a, (1+b_n)w_n\} \quad \textit{for} \quad n \geqq n_0,$$

where $a \geqq 0$, $b_n \geqq -1$ *if* $n \geqq n_0$, *further*

$$\sup_n \left| \sum_{i=1}^{n} b_i \right| = c < +\infty;$$

then the sequence w_n *is bounded.*

PROOF. Let $n > m \geqq n_0$, then we obtain by induction that

$$w_n \leqq \max\{B_m^n w_m, \max_{m < j \leqq n} B_j^n a\},$$

where

$$B_j^n = \prod_{i=j}^{n-1} (1+b_i) \quad \text{if} \quad j < n, \quad \text{and} \quad B_n^n = 1.$$

However, as

$$\sum_{i=j}^{n-1} b_i \leqq 2c \quad \text{for every} \quad j \quad \text{and} \quad n,$$

further $1+u \leqq e^u$, we have

$$B_j^n \leqq \exp \sum_{i=j}^{n-1} b_i \leqq e^{2c}, \qquad j \leqq n,$$

therefore

$$w_n \leqq \max \{w_m e^{2c}, a e^{2c}\},$$

which proves the statement. <>

As a consequence, we prove

LEMMA 4. *Under assumptions* A)—E), *the process* Y_n *defined by* (1) *is bounded with probability one, i.e.* $\sup_n \|Y_n\| < +\infty$ *a.s.*

PROOF. From (1) we obtain

$$\|Y_{n+1}-y_0\|^2 = \|Y_n-y_0\|^2 + 2\varrho_n(Y_n-y_0, Z_{n+1}) + \varrho_n^2\|Z_{n+1}\|^2,$$

that is

$$1+\|Y_{n+1}-y_0\|^2 = 1+\|Y_n-y_0\|^2 + 2\varrho_n\big(Y_n-y_0, Z_{n+1}-U(Y_n)\big) + \\ + \varrho_n^2\|Z_{n+1}\|^2 + 2\varrho_n\big(Y_n-y_0, U(Y_n)\big).$$

With the notations $w_n = 1+\|Y_n-y_0\|^2$ and

$$b_n = 2\varrho_n w_n^{-1}\big(Y_n-y_0, Z_{n+1}-U(Y_n)\big) + \varrho_n^2 w_n^{-1}\|Z_{n+1}\|^2,$$

this equality becomes

$$w_{n+1} = (1+b_n)w_n + 2\varrho_n\big(Y_n-y_0, U(Y_n)\big). \tag{5}$$

We wish to apply Lemma 3.; first we prove that $\sum_{i=1}^{n} b_i$ converges with probability one to a finite limit, this certainly implies the corresponding assumptions of Lemma 3., at least almost surely. Really, from the Cauchy inequality we have

$$\|Z_{n+1}\|^2 \leqq 3^2\|Z_{n+1}-\mathsf{E}_n Z_{n+1}\|^2 + 3^2\|\mathsf{E}_n Z_{n+1}-U(Y_n)\|^2 + 3^2\|U(Y_n)\|^2$$

but

$$\sum \varrho_n^2 w_n^{-1}\|Z_{n+1}-\mathsf{E}_n Z_{n+1}\|^2 < +\infty \text{ a.s.}$$

follows from C) and from Lemma 2., further

$$\sum \varrho_n^2 w_n^{-1}\|\mathsf{E}_n Z_{n+1}-U(Y_n)\|^2 < \infty \text{ a.s.}$$

in view of B), as $(1+u)^2 \leqq 2(1+u^2)$ applies for $u = \|Y_n-y_0\|$, while

$$\sum \varrho_n^2 w_n^{-1}\|U(Y_n)\|^2 < +\infty \text{ a.s.}$$

is a consequence of D) and A); thus

$$\sum \varrho_n^2 w_n^{-1}\|Z_{n+1}\|^2 < +\infty \text{ a.s.} \tag{6}$$

On the other hand, the Cauchy inequality, and the elementary estimation

$$\frac{u}{1+u^2} < \frac{2}{1+u} \quad \text{if} \quad u \geqq 0$$

imply that

$$w_n^{-1}\left|\left(Y_n - y_0, \mathsf{E}_n Z_{n+1} - U(Y_n)\right)\right| \leqq w_n^{-1}\|Y_n - y_0\|\,\|\mathsf{E}_n Z_{n+1} - U(Y_n)\| \leqq$$
$$\leqq 2(1+\|Y_n - y_0\|)^{-1}\|\mathsf{E}_n Z_{n+1} - U(Y_n)\|,$$

therefore, as Y_n, ϱ_n and w_n are measurable on $\mathscr{A}_n$, we have from assumption B) that

$$\sum \left|\mathsf{E}_n\left(\varrho_n w_n^{-1}(Y_n - y_0, Z_{n+1} - U(Y_n))\right)\right| < +\infty \text{ a.s.}$$

Further, as

$$\frac{u^2}{(1+u^2)^2} \leqq (1+u^2)^{-1},$$

in the same way as above, we obtain from C) that

$$\sum \mathsf{E}_n\left((\varrho_n w_n^{-1}(Y_n - y_0, Z_{n+1} - U(Y_n)))^2\right) < +\infty \text{ a.s.}$$

that is Lemma 1. applies and yields that

$$\sum_{n=1}^{m} \varrho_n w_n^{-1}\left(Y_n - y_0, Z_{n+1} - U(Y_n)\right) = S_m \tag{7}$$

converges with probability one to a finite limit. Therefore, we know that $s_n = \sum_{i=1}^{n} b_i$ converges with probability one to a finite limit.

Now, let us turn to equality (5). In view of E),

$$w_{n+1} \leqq (1+b_n)w_n \quad \text{if} \quad \|Y_n - y_0\| \geqq r_0.$$

However, in the opposite case

$$w_{n+1} \leqq 1 + (r_0 + \varrho_n\|Z_{n+1}\|)^2,$$

since

$$\|Y_{n+1} - y_0\| \leqq \|Y_n - y_0\| + \varrho_n\|Z_{n+1}\|,$$

but from (6) we know that

$$\lim_n \varrho_n (w_n)^{-\frac{1}{2}}\|Z_{n+1}\| = 0 \text{ a.s.}$$

whence immediately follows that for almost every realization of process (1) we have such a finite number n_0 that

$$\varrho_n\|Z_{n+1}\| \leqq 1 \quad \text{if} \quad \|Y_n - y_0\| \leqq r_0 \quad \text{and} \quad n \geqq n_0.$$

Thus, setting $a = 1 + (r_0+1)^2$,

$$w_{n+1} \leqq \max\{a, (1+b_n)w_n\}$$

holds for $n \geqq n_0$, which implies the statement by Lemma 3. ◇

We are now in a position to prove the main result of this section. As regards the basic ideas of its proof, see [15] and [14].

THEOREM 1. *If the gradient vector $U(y)$ is continuous on R, and assumptions* A)—F) *are satisfied, then* $\lim_n \|U(Y_n)\| = 0$ *a.s. Moreover, if the stationary points of f are isolated from each other, then process* (1) *converges with probability one to one on them.*

PROOF. First we prove that

$$\sum \varrho_n \|U(Y_n)\|^2 < +\infty \text{ a.s.}; \tag{8}$$

since $\sum \varrho_n = +\infty$ a.s., it will be possible to deduce from here the first asserion. Of course, we must exploit the special structure of process (1), since, in general, (8) implies only

$$\underline{\lim}\, \|U(Y_n)\| = 0 \text{ a.s.} \tag{9}$$

Putting $x = Y_{n+1}$ and $y = Y_n$ in F), we obtain

$$f(Y_{n+1}) \leqq f(Y_n) + \varrho_n(Z_{n+1}, U(Y_n)) - L(r)\varrho_n^{\mu}\|Z_{n+1}\|^{\mu} - K(r)\varrho_n^2\|Z_{n+1}\|^2,$$

that is

$$\begin{aligned} \varrho_n\|U(Y_n)\|^2 \leqq f(Y_{n+1}) - f(Y_n) + L(r)\varrho_n^{\mu}\|Z_{n+1}\|^{\mu} + \\ + K(r)\varrho_n^2\|Z_{n+1}\|^2 + \varrho_n(Z_{n+1} - U(Y_n), U(Y_n)), \end{aligned} \tag{10}$$

where $r = \sup_n \|Y_n - y_0\| < +\infty$ a.s. in view of Lemma 4.

We have to show that summing over n in (10), the right hand side remains bounded with probability one. Since, in our case, u^{μ} is a convex function of $u \geqq 0$ and $u^{\mu} \leqq 1 + u^2$, we obtain

$$\|Z_{n+1}\|^{\mu} \leqq (\|Z_{n+1} - \mathsf{E}_n Z_{n+1}\| + \|\mathsf{E}_n Z_{n+1} - U(Y_n)\| + \|U(Y_n)\|)^{\mu} \leqq$$

$$\leqq 3^{\mu-1}\|Z_{n+1} - \mathsf{E}_n Z_{n+1}\|^{\mu} + 3^{\mu-1}\|\mathsf{E}_n Z_{n+1} - U(Y_n)\|^{\mu} + 3^{\mu-1}\|U(Y_n)\|^{\mu} \leqq$$

$$\leqq 3^{\mu-1}(1 + \|Z_{n+1} - \mathsf{E}_n Z_{n+1}\|^2) + 3^{\mu-1}\|\mathsf{E}_n Z_{n+1} - U(Y_n)\|^{\mu} + 3^{\mu-1}\|U(Y_n)\|^{\mu},$$

where

$$\sum \varrho_n^{\mu}\|Z_{n+1} - \mathsf{E}_n Z_{n+1}\|^2 < +\infty \text{ a.s.}$$

in view of C), Lemma 4. and Lemma 2., while

$$\sum \varrho_n^{\mu}\|\mathsf{E}_n Z_{n+1} - U(Y_n)\|^{\mu} < +\infty \text{ a.s.}$$

follows simply from B) and from Lemma 4., and

$$\sum \varrho_n^{\mu}(1 + \|U(Y_n)\|^{\mu}) < +\infty \text{ a.s.}$$

is a consequence of D), Lemma 4. and of the second part of A). Thus we have

$$\sum \varrho_n^{\mu}\|Z_{n+1}\|^{\mu} < +\infty \text{ a.s.} \tag{11}$$

which implies also

$$\sum \varrho_n^2 \|Z_{n+1}\|^2 < +\infty \text{ a.s.} \tag{12}$$

since $\mu \leqq 2$.

As regards the last term in (10), observe that

$$\sum \left|\mathsf{E}_n\left(\varrho_n(Z_{n+1}-U(Y_n), U(Y_n))\right)\right| = \sum \varrho_n \left|\left(\mathsf{E}_n Z_{n+1}-U(Y_n)\right)\right| \leqq$$

$$\leqq \sum \varrho_n \|\mathsf{E}_n Z_{n+1}-U(Y_n)\| \, \|U(Y_n)\| < +\infty \text{ a.s.},$$

as follows from Lemma 4. and from assumptions D), B). Similarly,

$$\sum \mathsf{E}_n\left(\varrho_n(Z_{n+1}-U(Y_n), U(Y_n))\right)^2 \leqq$$

$$\leqq \sum \varrho_n^2 \|U(Y_n)\|^2 \mathsf{E}_n \|Z_{n+1}-U(Y_n)\|^2 < +\infty \text{ a.s.}$$

in view of Lemma 4. and of D) and C). Therefore, Lemma 1. implies that

$$S_m = \sum_{n=1}^{m} \varrho_n\left(Z_{n+1}-U(Y_n), U(Y_n)\right) \text{ converges a.s.} \tag{13}$$

to a finite limit. Thus, taking into account that $f(x)$ is bounded above, the comparison of (10), (11), (12), (13) results in (8).

To deduce $\lim \|U(Y_n)\| = 0$ a.s. from (8), observe first that $U(y)$ is uniformly continuous on the bounded domains of R, that is for any given $\varepsilon > 0$ and $r > 0$ there exists such a $\delta > 0$ that

$$\|U(Y_n)\| \leqq \|U(Y_m)\| + \varepsilon \tag{14}$$

if

$$\|Y_n - Y_m\| = \left\| \sum_{i+m}^{n-1} \varrho_i Z_{i+1} \right\| < \delta, \quad \|Y_n - y_0\| \leqq r$$

and

$$\|Y_m - y_0\| \leqq r.$$

Instead of proving the a.s. convergence of $\sum \varrho_n Z_{n+1}$, we show that

$$S_m = \sum_{n=1}^{m} \varrho_n I_n Z_{n+1} \tag{15}$$

converges a.s. to a finite limit, where

$$I_n = \begin{cases} \varepsilon & \text{if} \quad \|U(Y_n\| \geqq \varepsilon \\ 0 & \text{othervise.} \end{cases}$$

Since I_n is measurable on $\mathscr{A}_n$, in the decomposition

$$\sum_{n=1}^{m} \varrho_n I_n Z_{n+1} = \sum_{n=1}^{m} \varrho_n I_n (Z_{n+1} - \mathsf{E}_n Z_{n+1}) +$$

$$+ \sum_{n=1}^{m} \varrho_n I_n (\mathsf{E}_n Z_{n+1} - U(Y_n)) + \sum_{n=1}^{m} \varrho_n I_n U(Y_n)$$

the a.s. convergence of the first term follows immediately from Lemma 1. by Lemma 4. and by C), while that of the second one is a consequence of B) and of Lemma 4., further,

$$I_n \|U(Y_n)\| \leqq \|U(Y_n)\|^2$$

implies the a.s. convergence of the third term by (8); thus (15) really holds. Moreover, from (6) and from Lemma 4. we have

$$\lim \varrho_n \|Z_{n+1}\| = 0 \text{ a.s.} \tag{16}$$

and simply from Lemma 4.,

$$\sup_n \|Y_n - y_0\| = r < +\infty \text{ a.s.} \tag{17}$$

follows. Now, let us select such a realization of process (1) that the relations in (9), (15), (16) and (17) be true for it; then, given an $\varepsilon > 0$, we have a $\delta > 0$ for which (14) holds, under its first condition. Further, we have such an n_0 that, in view of (15) and (16),

$$\sup_{n, m \geqq n_0} \left\| \varrho_m Z_{m+1} + \sum_{i=m+1}^{n-1} \varrho_i I_i Z_{i+1} \right\| < \delta, \tag{18}$$

and, according to (9), we may suppose that

$$\|U(Y_{n_0})\| \leqq \varepsilon. \tag{19}$$

Therefore, if $n > n_0$, $\|U(Y_n)\| > \varepsilon$, and m denotes the greatest integer between n_0 and n with $\|U(Y_m)\| \leqq \varepsilon$, then $I_i = 1$ if $m < i < n$, consequently, (14) applies, and by (18) it yields that $\|U(Y_n)\| \leqq \|U(Y_m)\| + \varepsilon \leqq 2\varepsilon$, which proves the first assertion of the theorem.

The proof of the second statement is very simple, since the continuity of $U(y)$ means that any limit point of process (1) must be a stationary point of $f(x)$. However, according to Lemma 4., each subsequence has at least one limit point, and they are isolated from each other, thus (16) is possible only if the procedure converges with probability one, which completes the proof. ◇

Remark 1. The above results extend immediately to the case, when R is a separable Hilbert space, the only additional assumption needed for this generalization is the weak continuity of $U(y)$; then $\lim \|U(y)\| = 0$ a.s., as well as the weak convergence of the process follow without any modification of the arguments. To

deduce strong convergence, we need a further condition: Let S denote the set of stationary points of $f(x)$, i.e. $U(z)=0$ if, and only if, $z \in S$, and set

$$S(\varepsilon, r) = \{x; \|x-y_0\| \leqq r, \inf_{y \in S} \|x-y\| \geqq \varepsilon\}$$

for $\varepsilon>0$, $r>0$. Then the assumption

G) $$\inf\{\|U(y)\|; y \in S(\varepsilon, r)\} > 0 \text{ for } \varepsilon > 0 \text{ and } r \geqq r_0$$

implies $\lim d(Y_n, S)=0$ a.s., where $d(Y_n, S)$ denotes the distance of Y_n and S, whence even the strong convergence of process (1) follows, provided that the components of S are single points with positive minimal distance between them.

Remark 2. M. B. NEVELSON and R. Z. HASMINSKII [20] are dealing with a similar problem under an assumption of type G). Of course, as G) does not imply the continuity of $U(y)$, they can not derive that $\lim \|U(Y_n)\|=0$ a.s. but $\lim d(Y_n, S)=0$ a.s. The basic idea of this proof is also to deduce convergence from a relation of type (8), see [15] and [14], too. Unfortunately, the theorems in [20] can not be applied to mode estimation, since some further assumptions as $\lim f(x) = -\infty$ if $\lim \|x\| = +\infty$, and $\mathsf{E}_n Z_{n+1} = U(Y_n)$ — if formulated in our terminology — do not hold for density functions. Moreover, it is desirable to express conditions A)—C) in terms of conditional expectations and $\mathscr{A}_n$-measurable ϱ_n-s, since then several accelerating methods can directly be incorporated into the theory. On the other hand, in [20] the assumption that the regression function $U(y)$ is the gradient of a function $f(x)$ is replaced by a weaker one, therefore, it would be useful to give a common generalization of the results in [20] and of Theorem 1.

Remark 3. Under some regularity conditions it is easy to verify that a local minimum or a saddle point can not be limit point of the process; for a more detailed study of this question see [20], where several kinds of such unstable points are described. However, nothing is known on the stability of stationary points with a singular matrix of second derivatives.

Remark 4. If the stationary points are not isolated, then — apart from some simple situations — the convergence of the procedure can not be proved by known methods. But, if $(y-y_0, U(y)) \geqq 0$, or $U(y)=0$ outside a bounded domain and the further stationary points are isolated, then convergence still holds.

The problems mentioned in the above remarks will be discussed in a forthcoming paper.

3. Mode estimation

In order to apply Theorem 1. to mode estimation, we need a good gradient estimation in the sense of B) and C). Estimation methods for the derivatives of a probability density are treated by K. BHATTACHARYA [3] and E. F. SCHUSTER [4], while I. VINCZE [12] has developed one for the logarithmic derivatives.

To obtain a suitable estimate of the gradient of a probability density, we shall use the technique of K. BHATTACHARYA [3] and of E. F. SCHUSTER [4], based on the analitical theory of singular integrals, see [18] and [1], [2]. From now on, $f(x)$ denotes a differentiable density function on R with gradient vector $U(y)$. The regularity condition F) will be replaced by the following stronger one:

F″) For each $y \in R$ there exists an $L_1(y) \leqq M_1(1+\|y-y_0\|)$ such that

$$\|U(x)-U(y)\| \leqq L_1(y)\|x-y\|^{\mu-1}+K_1\|x-y\|$$

holds for each $x \in R$, where M_1 and K_1 are nonnegative, finite constants, and $1<\mu\leqq 2$.

The very same argument that was used to deduce F) from F′) shows that F) is really a consequence of F″).

Our estimate of $U(y)$ will be defined by means of a vector-valued measurable function $T(x) \in R$ and defined on R; in the following conditions and later, only Lebesgue integrable functions are considered on R. $T(x)$ is described as follows.

(i) $\int T(x)\,dx = 0.$

(ii) $\int_R T(x)(x, U)\,dx = U \quad$ for each $\quad U \in R.$

(iii) $\int_R \|x\|^{\mu}\|T(x)\|\,dx < +\infty$, μ is the same as in F''.

(iv) $\int_R \|x\|^2\|T(x)\|\,dx < +\infty.$

(v) $\int_R \|T(x)\|^2\,dx < +\infty.$

Such functions can be obtained by differentiating and normalizing a symmetric density function on R so that (ii) be satisfied. Lemma 5. and Lemma 6. show the role of these assumptions.

LEMMA 5. *If density function $f(x)$ satisfies* F″), *while* (i)—(iv) *hold for* $T(x)$, *then for any* $\lambda>0$ *we have*

$$\left\| U(y)-\lambda^{k+1}\int_R T\big(\lambda(x-y)\big)f(x)\,dx \right\| \leqq \lambda^{1-\mu}L_1(y)\int_R \|x\|^{\mu}\|T(x)\|\,dx+$$

$$+\lambda^{-1}M_1\int_R \|x\|^2\|T(x)\|\,dx,$$

where k is the dimension of R.

PROOF. As, in view of (3), we have

$$\big|f(x)-f(y)-(x-y, U(y))\big| \leqq L_1(y)\|x-y\|^{\mu}+M_1\|x-y\|^2,$$

and

$$\lambda^{k+1}\int_R T(\lambda(x-y))f(y)\,dx = \lambda f(y)\int_R T(x)\,dx = 0,$$

$$\lambda^{k+1}\int_R T(\lambda(x-y))(x-y, U(y))\,dx = \int_R T(x)(x, U(y))\,dx = U(y),$$

further

$$\lambda^{k+1}\int_R \|x-y\|^{\mu}\big\|T(\lambda(x-y))\big\|\,dx = \lambda^{1-\mu}\int_R \|x\|^{\mu}\|T(x)\|\,dx,$$

$$\lambda^{k+1}\int_R \|x-y\|^{2}\big\|T(\lambda(x-y))\big\|\,dx = \lambda^{-1}\int_R \|x\|^{2}\|T(x)\|\,dx,$$

the statement follows directly from

$$\Big\|U(y)-\lambda^{k+1}\int_R T(\lambda(x-y))f(x)\,dx\Big\| \leqq$$

$$\leqq \Big\|U(y)-\lambda^{k+1}\int_R T(\lambda(x-y))\big(f(y)+(x-y, U(y))\big)\,dx\Big\| +$$

$$+\Big\|\lambda^{k+1}\int_R T(\lambda(x-y))\big(f(y)+(x-y, U(y))-f(x)\big)\,dx\Big\|$$

if we apply the elementary inequality

$$\Big\|\int_R V(x)\,dx\Big\| \leqq \int_R \|V(x)\|dx,$$

which is well known for vector-valued integrals. <>

LEMMA 6. *For any* $\lambda>0$ *and* $f(x)\geqq 0$ *we have*

$$\lambda^{2k+2}\int_R \big\|T(\lambda(x-y))\big\|^2 f(x)\,dx \leqq \lambda^{k+2}M_0\int_R \|T(x)\|^2\,dx,$$

where $M_0=\sup f(x)$.

PROOF. The assertion obtains trivially by a change of variables, as $f(x)$ is bounded above. <>

Suppose now that we are given a sequence $X_1, X_2, \ldots, X_n, \ldots$ of independent and identically distributed random vectors with common density function $f(x)$, and let $\mathscr{A}_n$ denote the minimal σ-algebra on which $X_1, X_2, \ldots, X_n$ are measurable, and let $\lambda_n>0$ be measurable on $\mathscr{A}_n$. According to Lemma 5.,

$$Z_{n+1} = \lambda_n^{k+1}T\big(\lambda_n(X_{n+1}-Y_n)\big) \tag{20}$$

defines a sequence of asymptotically unbiased estimates of $U(Y_n)$ if $\lim \lambda_n = +\infty$ and Y_n remains bounded, since

$$\mathsf{E}_n Z_{n+1} = \lambda_n^{k+1}\int_R T\big(\lambda_n(x-Y_n)\big)f(x)dx.$$

Despite the unfortunate situation that the conditional variance of Z_{n+1} goes to infinity if its conditional expectation approaches $U(Y_n)$, it is possible to choose the sequences ϱ_n and λ_n in such a way that Theorem 1. could be applied.

THEOREM 2. *Let us consider process* (1) *with specifications* (20) *and* (i)—(v) *for* Z_{n+1}, *and* I.—IV. *for* ϱ_n *and* λ_n, *as follows:*

I. $\sum \varrho_n = +\infty$ *a.s.*

II. $\lim \lambda_n = +\infty$ *a.s.*

III. $\sum \varrho_n \lambda_n^{1-\mu} < +\infty$ *a.s.*

IV. $\sum \varrho_n^{\mu} \lambda_n^{k+2} < +\infty$ *a.s.*,

where $1 < \mu \leqq 2$ *is fixed. Further, suppose that the underlying probability density* $f(x)$ *satisfies* D), E), F″), *and its gradient* $U(y)$ *is continuous on* R; *then* $\lim \|U(Y_n)\| = 0$ *a.s. Moreover, if the stationary points of* $f(x)$ *are isolated, then the procedure converges to one of them with probability one.*

PROOF. Since we have already pointed out that F″) implies F), it suffices to show that conditions A)—C) are fulfilled. However, the first part of A) is supposed, while the second one follows from II. and IV. Further, B) is a consequence of F″), III., and of Lemma 5., while C) is implied by Lemma 6. and by IV., since

$$\mathsf{E}_n \|Z_{n+1} - \mathsf{E}_n Z_{n+1}\|^2 \leqq \mathsf{E}_n \|Z_{n+1}\|^2 = \int \left|\left|\lambda_n^{k+1} T\big(\lambda_n (x - Y_n)\big)\right|\right|^2 f(x)\, dx = O(\lambda_n^{k+2})$$

as $f(x)$ is bounded above in view of E). Therefore, Theorem 2. is proved as a consequence of Theorem 1. ◇

To see a simple example, let $\mu = 2$, $\varrho_n = \dfrac{1}{n}$, $\lambda_n = (\log n)^2$, and

$$T(x) = x(2\pi)^{-\frac{k}{2}} \exp\left(-\frac{1}{2}\|x\|^2\right).$$

It is easy to verify that in this case the corresponding assumptions of Theorem 2. are satisfied. Let us remark that Theorem 1. and 2. include several acceleration methods, since ϱ_n and λ_n may depend on the past history of the process. The problems of acceleration are to be discussed later.

It is important to observe that Z_{n+1} will be very small if λ_n is large and X_{n+1} is not close to Y_n. Therefore, we may hope for a good rate of convergence only if the initial value Y_1 is appropriately choosen. Of course, in practical situations, where the number of steps is limited, it is not desirable to let λ_n be too large, since it would result in a rather low rate of convergence, too. On the other hand, if $\lambda_n = \lambda = \text{const.}$ then the procedure does not converge to the roots of $U(y)$ but to that of

$$U_\lambda(y) = \lambda^{k+1} \int_R T\big(\lambda(x - y)\big) f(x)\, dx.$$

Let us remark that in the above example $U_\lambda(y)$ is just the gradient of the density function

$$f_\lambda(y) = \int_R \lambda^k (2\pi)^{-\frac{k}{2}} \exp\left(-\frac{1}{2}\|\lambda(x-y)\|^2\right) f(x)\,dx,$$

which is the expectation of the corresponding empirical density function. This means thet the algorithm locates the modes of f_λ if $\lambda_n = \lambda$, as follows easily from Theorem 1., and a similar statement is true for any T derived from a symmetric density function by differentiation.

REFERENCES

[1] ROSENBLATT, M.: Remarks on some non-parametric estimates of a density function, *Ann. Math. Statist.* **27** (1956) 832—837.

[2] PARZEN, E.: On estimation of a probability density function and mode, *Ann. Math. Statist.* **33** (1962) 1065—1076.

[3] BHATTACHARYA, K.: Estimation of a probability density function and its derivatives, *Sankhya* **29** (1967) 373—382.

[4] SCHUSTER, E. F.: Estimation of a probability density function and its derivatives, *Ann. Math. Statist.* **40** (1969) 1187—1195.

[5] RÉVÉSZ, P.: Testing of density functions, *Periodica Math. Hungar.* **1** (1971) 35—44.

[6] RÉVÉSZ, P.: On empirical density function, *Periodica Math. Hungar* (1972) 85—110.

[7] CHERNOFF, H.: Estimation of the mode, *Ann. Inst. Stat. Math.* **16** (1964) 31—41.

[8] GRENANDER, U.: Some direct estimates of the mode *Ann. Math. Statist.* **36** (1965) 131—138.

[9] VENTER, J. H.: On estimation of the mode, *Ann. Math. Statist.* **38** (1967) 1446—1455.

[10] FU, K. S.: *Statistical pattern recognition, in Adaptive, Learning and Pattern Recognition Systems,* Mendel-Fu Ed. Academic Press, New York and London, 1970.

[11] WISHART, D.: CLUSTAN 1A Program Library Services No. 8. (1969), St Andrews University Ed.

[12] VINCZE, I.: Oral communication (1972).

[13] FÁBIAN, V.: A new one dimensional stochastic approximation method for finding a local minimum of a function, *Trans. third Prague Conference,* (1964), 85—105.

[14] FRITZ, J.: On a pattern classification algorithm of R. L. Kashyap, submitted to *Problems of Control and Information Theory.*

[15] AIZERMAN, M. A.—BRAVERMAN, E. M.—ROSONOER, L. I.: *The Method of Potential Functions in the Theory of Machine Learning,* Chap. IV. Nauka, Moscow, (1970), (in Russian).

[16] WASAN, M.: Stochastic Approximation, Mir, Moscow (1972), (in Russian, with the Appen dix of Vaisboard, E. M. and Judin, D. B.)

[17] SCHMETTERER, L.: *L'Approximation Stochastique,* Univ. de Clermont-Ferrand Ed. (1972).

[18] BOCHNER, S.: *Harmonic Analysis and the Theory of Probability,* Univ. of California Press, (1955).

[19] NEVEU, J.: *The Mathematical Foundations of Probability Theory,* Masson, Paris, (1964).

[20] NEVELSON, M. B.—HASMINSKII, R. Z.: *Stochastic Approximation and Iterative Estimation Methods,* Nauka, Moscow, (1972) (in Russian).

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received February 25, 1973)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 323—328.

ДВЕ ТЕОРЕМЫ ОБ ОТКЛОНЕНИИ СУММ ФУРЬЕ

Г. П. НЕВАИ

Обозначим через $\tilde{C}$ Банахово пространство непрерывных, 2π-периодических функций $f(x)$, снабженное с обычной равномерной нормой

$$\|f\| = \max_{|x| \leqq \pi} |f(x)|.$$

Пусть далее[1]

$$\Delta_\delta^R f(x) = 2^{-R} \sum_{k=0}^{R} (-1)^k \binom{R}{k} f(x+k\delta) \quad (R = 1, 2, \ldots)$$

означает R-ую (модифицированную) разность функции $f(x)$ с шагом δ. Обозначим через

$$\omega_R(f; \delta) = \max_{\substack{|x| \leqq \pi \\ |h| \leqq \delta}} |\Delta_h^R f(x)| \qquad (\delta \geqq 0)$$

модуль гладкости порядка R функции $f(x)$ из $\tilde{C}$, а через

$$\omega_R(f; \delta)_{L_1} = \max_{|h| \leqq \delta} \frac{1}{2\pi} \int_{-\pi}^{\pi} |\Delta_h^R f(x)| \, dx \qquad (\delta \geqq 0)$$

R-ый модуль гладкости функции $f(x)$ из $\tilde{C}$ в метрике $\tilde{L}_1$.

Пусть

$$S_n(x, f) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x+t) D_n(t) \, dt \qquad (n = 1, 2, \ldots)$$

где

$$D_n(t) = \frac{\sin\left(n + \frac{1}{2}\right) t}{2 \sin \frac{t}{2}}$$

ядро Дирихле n-го порядка, означает n-ую частную сумму тригонометрического ряда Фурье функции $f(x)$ из $\tilde{C}$.

Пусть $\varphi(n)$ $(n = 1, 2, \ldots)$ такая последовательность чисел, для которой выполняются соотношения

$$\varphi(n) \geqq 1, \quad \varphi(n) = O(n), \quad \varphi(n)^{-1} = o(1). \tag{1}$$

[1] Целесообразность введения нормирующего множителя $(-2)^{-R}$ была показана в нашей работе [6].

Обозначим через $H(\varphi)$ класс тех функций $f(x)$, принадлежащих к $\tilde{C}$, для которых равномерно по всем a и b $(-\pi \leqq a < b \leqq \pi)$ имеет место соотношение

$$\varphi(n)\left|\int_a^b f(\xi)e^{in\xi}\,d\xi\right| = O(1) \quad (n = 1, 2, \ldots) \tag{2}$$

Отметим, что в случае $\varphi(n) \equiv n$, $H(\varphi)$ совпадает с классом функций, имеющих ограниченное отклонение, который был введен Адамаром [3] (см. еще Алексич [1]).

Неш [5] в 1953-ом году доказал следующее утверждение:

Если $f \in H(\varphi)$ и $\theta(n)$ $(n=1, 2, \ldots)$ монотонно неубывающая последовательность чисел, обладающая свойствами

$$1 \leqq \theta(n) \leqq \varphi(n), \quad 1 \leqq \frac{\theta(n+1)}{\theta(n)} \leqq \frac{\varphi(n+1)}{\varphi(n)},$$

то тогда

$$\|f - S_n(f)\| = O\left\{\left[\omega\left(f; \frac{1}{n}\right)\log\theta(n) + \frac{n}{\varphi(n)}\right] + \frac{1}{\theta(n)}\right\} \qquad (n = 1, 2, \ldots).$$

Через год М. Сато [7] — применяя в основном метод Нэша — усилил это утверждение в следующей форме: [2]

Если $f \in H(\varphi)$ и последовательность $\theta(n)$ $(n=1, 2, \ldots)$ монотонно неубывает, далее $1 \leqq \theta(n) \leqq \varphi(n)$, то тогда

$$\|f - S_n(f)\| = O\left[\omega\left(f; \frac{1}{n}\right)\log\frac{n\theta(n)}{\varphi(n)} + \frac{1}{\theta(n)}\right] \qquad (n = 1, 2, \ldots)$$

С тех пор вопрос о том, что можно ли в теореме Сато модуль непрерывности функции $f(x)$ заменить на модули гладкости высших порядков, оставался открытым. Ответ на эту проблему дает следующая

Теорема 1. *Пусть $f \in H(\varphi)$ и R произвольное, но фиксированное, натуральное число. Если последовательность $\theta(n)$ $(n=1, 2, \ldots)$ такова, что* [3]

$$\frac{c_1}{n} \leqq \frac{\theta(n)}{\varphi(n)} \leqq c_2 \qquad (n > n_0) \quad (0 < c_1, c_2)$$

то тогда

$$\|f - S_n(f)\| = O\left[\omega_R\left(f; \frac{1}{n}\right)\log\frac{n\theta(n)}{\varphi(n)} + \frac{1}{\theta(n)}\right] \qquad (n = 1, 2, \ldots). \tag{3}$$

---

[2] См. еще книгу Н.К. Бари [2].

[3] Подчеркнем, что последовательность $\theta(n)$ может и не быть монотонной.

В 1955-ом году М. Сато [8] заметил, что из его результатов [7] вытекает также следующая теорема:

Если $f\in\tilde{C}$, то тогда

$$\|f-S_n(f)\| = O\left\{\left\{1+\log\left[1+\frac{n\omega\left(f;\frac{1}{n}\right)_{L_1}}{\omega\left(f;\frac{1}{n}\right)}\right]\right\}\omega\left(f;\frac{1}{n}\right)\right\} \qquad (n=1,2,\ldots).$$

В связи с этим мы покажем, что справедлива и более общая

Теорема 2. *Если $f\in\tilde{C}$ и R произвольное, но фиксированное, натуральное число, то тогда*

$$\|f-S_n(f)\| = \left\{\left\{1+\log\left[1+\frac{n\cdot\omega_R\left(f;\frac{1}{n}\right)_{L_1}}{\omega_R\left(f;\frac{1}{n}\right)}\right]\right\}\omega_R\left(f;\frac{1}{n}\right)\right\} \qquad (n=1,2,\cdots). \tag{4}$$

Прежде, чем приступить к доказательству этих двух теорем, мы отметим, что они включают в себе несколько любопытных признаков для равномерной сходимости рядов Фурье.

Следствие 1 *(обобщенный признак Дини-Липшица). Если $f\in\tilde{C}$ и при некотором натуральном R*

$$\omega_R(f;\delta) = o\left(\frac{1}{|\log\delta|}\right) \qquad (\delta\to+0)$$

то ряд фурье функции $f(x)$ сходится равномерно.

Для этого достаточно заметить, что в силу Теоремы 31 из книги Харди-Рогозинского [4] [4] для всякой функции $f(x)$ из $\tilde{C}$ найдется последовательность $\varphi(n)\equiv\varphi_0(n)$ со свойствами (1) такова, что $f\in H(\varphi_0)$. Далее, полагая в (3) $\varphi(n)=\varphi_0(n)$ и $\theta(n)=\varphi_0(n)$ мы и получим требуемый результат.

Следствие 2 *(признак Жордана-Дирихле). Если $f\in\tilde{C}$ и $\operatorname{var}_{-\pi}^{\pi} f<\infty$, то ряд фурье функции $f(x)$ сходится равномерно.*

Действительно, известно, что в этом случае $f\in H(\varphi)$ с $\varphi(n)\equiv n$ (см. Адамар [3]), значит достаточно положить в (3) $\varphi(n)\equiv n$, а $\theta(n)$ устремлять достаточно медленно к бесконечности.

Следствие 2 вытекает из Теоремы 2 также, ибо $\omega_R(f;\delta)_{L_1}=O(\delta)$, если $\operatorname{var}_{-\pi}^{\pi} f<\infty$. [5]

---

[4] Эта теорема есть — собственно говоря — не что иное, чем специальная формулировка теоремы Римана-Лебега.

[5] Мы имеем даже $\omega(f;\delta)_{L_1}=O(\delta)$ при условии $\operatorname{var}_{-\pi}^{\pi} f<\infty$.

Следствие 3. *Если* $f\in\tilde{C}$ *и при некотором натуральном* R

$$\omega_R(f;\delta)=o\left(\frac{1}{\underbrace{\log\log\ldots\log}_{k\text{ раз}}\frac{1}{\delta}}\right)\qquad(\delta\to+0)$$

далее $f\in H(\varphi)$ *с* $\varphi(n)=\dfrac{n}{\underbrace{\log\log\ldots\log}_{(k-1)\text{ раз}}n}$, *то ряд Фурье функции* $f(x)$ *сходится равномерно.*

Для доказательства Следствия 3 достаточно положить в (3) $\theta(n)=\underbrace{\log\log\ldots\log}_{(k-1)\text{ раз}}n$.

Следствие 4. *Если* $f\in\tilde{C}$ *и при некотором натуральном* R *и фиксированном* $c(>0)$

$$\omega_R(f;\delta)=o\left(\frac{1}{\log\log\frac{1}{\delta}}\right)\qquad(\delta\to+0)$$

и

$$\omega_R(f;\delta)_{L_1}=O\left[\omega_R(f;\delta)\left(\log\frac{1}{\delta}\right)^c\delta\right]\qquad(\delta\to+0),$$

то ряд фурье функции $f(x)$ *сходится равномерно.*

Доказательство Теоремы 1. Без ограничения всеобщности мы в праве предполагать, что $c_1=e$ и $c_2=\pi$. Соласно Теореме 4.2 из нашей работы [6] [6] мы имеем для всех функций $f(x)$ из $\tilde{C}$

$$S_n(x,f)-f(x)=S_n(x,\Delta^R_{\pi/n}f)+O_R\left[\omega_R\left(f;\frac{1}{n}\right)\right]\qquad(n=1,2,\ldots),$$

значит

$$(5)\qquad S_n(x,f)-f(x)=\frac{1}{\pi}\int_{-\pi}^{0}+\frac{1}{\pi}\int_{0}^{\pi}\Delta^R_{\pi/n}f(x+t)\,D_n(t)\,dt+O_R\left[\omega_R\left(f;\frac{1}{n}\right)\right].$$

Для простоты вычислений мы будем оценивать лишь второй интеграл, стоя-

[6] Возможность применяемости этой теоремы с шагом $\dfrac{\pi}{n}$ следует из результатов С. Н. Бернштейна и В. В. Рогозинского (см. например книгу Н. К. Бари [2]).

щий в правой части (5). После простых преобразований мы получим, что

$$\frac{1}{\pi}\int_0^{\pi} \Delta^R_{\pi/n} f(x+t) D_n(t)\,dt = \frac{1}{\pi}\int_{1/n}^{\pi} \Delta^R_{\pi/n} f(x+t)\frac{\sin nt}{2\operatorname{tg}\frac{t}{2}}\,dt + O\left[\omega_R\left(f;\frac{1}{n}\right)\right] =$$

$$= \frac{1}{\pi}\int_{1/n}^{\theta(n)/\varphi(n)} + \frac{1}{\pi}\int_{\theta(n)/\varphi(n)}^{\pi} \Delta^R_{\pi/n} f(x+t)\frac{\sin nt}{2\operatorname{tg}\frac{t}{2}}\,dt + O\left[\omega_R\left(f;\frac{1}{n}\right)\right] =$$

$$= \frac{1}{\pi}\int_{\theta(n)/\varphi(n)}^{\pi} \Delta^R_{\pi/n} f(x+t)\frac{\sin nt}{2\operatorname{tg}\frac{t}{2}}\,dt + O\left[\omega_R\left(f;\frac{1}{n}\right)\log\frac{n\theta(n)}{\varphi(n)}\right],$$

следовательно, в силу второй теоремы о среднем значении, мы имеем

$$\frac{1}{\pi}\int_0^{\pi} \Delta^R_{\pi/n} f(x+t) D_n(t)\,dt = \tag{6}$$

$$= \frac{1}{2\pi\operatorname{tg}\frac{\theta(n)}{2\varphi(n)}}\int_{\alpha}^{\beta} \Delta^R_{\pi/n} f(x+t)\sin nt\,dt + O\left[\omega_R\left(f;\frac{1}{n}\right)\log\frac{n\theta(n)}{\varphi(n)}\right],$$

где $0 \leqq \alpha < \beta \leqq \pi$. Утверждение Теоремы 1 немедленно вытекает из (5) и (6), ибо при всех α и β $(-\pi \leqq \alpha < \beta \leqq \pi)$ справедлива оценка

$$\int_{\alpha}^{\beta} \Delta^R_{\pi/n} f(x+t)\sin t\,dt = \int_{\alpha^*}^{\beta^*} f(x+t)\sin nt\,dt + O\left(\frac{1}{n}\right) =$$

$$= \int_{\alpha^{**}}^{\beta^{**}} f(t)\cos nt\,dt\sin nx + \int_{\alpha^{**}}^{\beta^{**}} f(t)\sin nt\,dt\cos nx + O\left(\frac{1}{n}\right) =$$

$$= O\left[\left|\int_{\alpha^{**}}^{\beta^{**}} f(\xi)e^{in\xi}\,d\xi\right| + \frac{1}{n}\right],$$

где $-\pi \leqq \alpha^* < \beta^* \leqq \pi$ и $-\pi \leqq \alpha^{**} < \beta^{**} \leqq \pi$ некоторые числа, а условие (2) для класса $H(\varphi)$ выполняется равномерно по всем a и b $(-\pi \leqq \alpha < b \leqq \pi)$.

Доказательство Теоремы 2 весьма просто. Рассмотрим (5), точнее — для простоты — второй интеграл, стоящий в правой части (5). Обозначим

$$m = \frac{\omega_R\left(f;\frac{1}{n}\right)_{L_1}}{\omega_R\left(f;\frac{1}{n}\right)} < 1. \tag{7}$$

Ясно, что

$$(8)\qquad \left|\int_0^{\pi} \Delta^R_{\pi/n} f(x+t)\, D_n(t)\, dt\right| = \left|\int_0^{m} + \int_m^{\pi} \Delta^R_{\pi/n} f(x+t)\, D_n(t)\, dt\right| =$$

$$= O\left[\log(1+mn)\,\omega_R\left(f;\frac{1}{n}\right)+\frac{1}{m}\,\omega_R\left(f;\frac{1}{n}\right)_{L_1}\right] \qquad (n = 1, 2, \ldots),$$

ибо абсолютная величина ядра Дирихле имеет мажоранту, вида

$$\frac{C_n}{1+n|t|} \qquad (-\pi \leqq t \leqq \pi),$$

где константа $C(>0)$ выбрана подходящим образом. Теорема 2 следует теперь из соотношений (5), (7) и (8).

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] Alexits, G.: Über eine Hadamardsche Fragestellung, *Math. Zeitschrift* **30** (1929), 43—46.
[2] Бари, Н. К.: *Тригонометрические ряды,* ГИФМЛ, Москва, 1961.
[3] Hadamard, M. J.: Essai sur l'étude des fonctions données par leur développement de Taylor, *Journ. de Math.* (4) **8** (1892), 101—186.
[4] Hardy, G. H.—Rogosinski, W. W.: *Fourier series,* Cambridge, University Press, 1956.
[5] Nash, J. P.: Uniform convergence of Fourier series, *The Rice Inst. Pamphlet,* Spec. Issue, Nov. 1953, 31—57.
[6] Неваи, Г. П.: Замечания о тригонометрическом интерполировании и суммах Фурье, *Studia Sci. Math. Hung.*
[7] Sato, M.: Uniform convergence of Fourier series, *Proc. of the Japan Acad.* **30** (1954), 528—531.
[8] Sato, M.: Uniform convergence of Fourier series, IV, *Proc. of the Japan Acad.* **31** (1955), 261—263.

Magyar Tudományos Akadémia Matematikai Kutató Intézete
1053 *Budapest, Reáltanoda u. 13—15.*

(*Поступила* 28 *марта* 1973)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 329—340.

A MODIFICATION OF THE ROBBINS—MONRO PROCESS

by

J. KOMLÓS and P. RÉVÉSZ

Introduction

Let $M(x)$ $(-\infty<x<+\infty)$ be a real valued function for which

$$M(0) = 0 \tag{1}$$

$$\mathcal{M}(\delta) = \inf_{|x|\geq\delta} |M(x)| > 0 \quad \text{for} \quad \delta > 0. \tag{2}$$

If $Y_1, Y_2, \ldots$ is a sequence of random variables satisfying for all $n=1, 2, \ldots$ the relations

$$\mathsf{E}(Y_1) = \mathsf{E}(Y_{n+1}|Y_1, Y_2, \ldots, Y_n) = 0, \tag{3}$$

$$\mathsf{E}(Y_1^2) \leqq K, \quad \mathsf{E}(Y_{n+1}^2|Y_1, Y_2, \ldots, Y_n) \leqq K \text{ a.e.} \tag{4}$$

and X_1 is an arbitrary real number, then the sequence defined by the recursion

$$X_{n+1} = X_n - c_n(Y_n + M(X_n)) \tag{5}$$

is called the Robbins—Monro process (R.M.P.). We restrict ourselves to the most investigated case, when the sequence of real numbers $c_n=1/n$.

BLUM [1] proved the following basic

THEOREM A. *If the function $M(x)$ has a "linear upper bound"*

$$|M(x)| \leqq A|x| \qquad (A > 0) \tag{6}$$

then $\mathsf{P}(X_n \to 0)=1$.

CHUNG [2] investigated the moments of an R.M.P. One of his results is the following

THEOREM B. *If the function $M(x)$ has both "linear upper and lower bounds"*

$$B|x| \leqq |M(x)| \leqq A|x|$$

with $B > \frac{1}{2}$, *then*

$$\mathsf{E}X_n^2 = O\left(\frac{1}{n}\right).$$

More general results are formulated on the moments in [3], [4] and [5]:

THEOREM C. *Suppose that*

$$B|x| \leqq |M(x)| \leqq A|x| \qquad (0 < B < A < +\infty;\ -\infty < x < +\infty). \tag{7}$$

Then we have

$$\mathsf{E}X_n^2 = \begin{cases} O\left(\frac{1}{n}\right) & \text{if} \quad B > \frac{1}{2} \\ O\left(\frac{\log n}{n}\right) & \text{if} \quad B = \frac{1}{2} \\ O\left(\frac{1}{n^{2B}}\right) & \text{if} \quad B < \frac{1}{2} \end{cases}. \tag{8}$$

It is well-known that in Theorem A the upper bound (6) cannot be dropped, e.g. for the function $M(x)=x^3$ the R.M.P. diverges even if there is no error, i.e. $Y_n \equiv 0$. (Consequently, it cannot be dropped in Theorems B and C, either.) As to the lower bound in Theorems B and C, it consists of two essentially different parts: a linear lower bound "near 0", and a linear lower bound "near infinity". The first one is known to be essential, e.g. if the derivative $M'(0)$ exists, the behaviour of the process is essentially different in three cases $M'(0) > \frac{1}{2}$, $M'(0) = \frac{1}{2}$, $M'(0) < \frac{1}{2}$. We show in § 3 that, for an R.M.P., the second part of the condition (i.e. the existence of a linear lower bound "near infinity") cannot be omitted either.

However, both the "upper and lower bounds" at the infinity seem unessential (though inevitable), since the process' visit "at the infinity" should be more the exceptional situation than the rule. This suggests the idea of modifying the process in such a way that it be less sensitive to the large values of $M(x)$.

§ 1. The modified process

In [6] we have introduced the following modified Robbins—Monro process (M.R.M.P.) (we repeat the definition in case $c_n = \frac{1}{n}$):

X_n is defined by the following recursion (X_1 is a real number) $X_{n+1} = X_n - Z_n$ where

$$Z_n = \begin{cases} \frac{1}{n}(Y_n + M(X_n)) & \text{if} \quad \left|\frac{1}{n}(Y_n + M(X_n))\right| \leqq 1 \\ \operatorname{sign}\frac{1}{n}(Y_n + M(X_n)) & \text{if} \quad \left|\frac{1}{n}(Y_n + M(X_n))\right| > 1 \end{cases} \tag{9}$$

(the conditions (1), (2), (3) and (4) still apply).

We show that this definition (which seems quite natural to us) makes it possible to eliminate the conditions imposed on $M(x)$ which are unessential, yet necessary for the original process. This is not only of theoretical values, since the function $M(x)$ is unknown. (This, of course, does not apply for the behaviour of $M(x)$ near 0 (say the derivative $M'(0)$), since this behaviour is essential for the whole process; perhaps some sequential methods are needed. See e.g. [7].)

In §§ 1 and 2 X_n always denotes an M.R.M.P.

The basic statements of this paper are the following three theorems:

THEOREM 1. *For any* $\varepsilon>0$ *we have*

$$\mathsf{P}(|X_n| > \varepsilon) = O\left(\frac{1}{n}\right). \tag{10}$$

THEOREM 2.

$$\mathsf{P}(X_n \to 0) = 1. \tag{11}$$

That is to say, for an M.R.M.P. condition (6) in Theorem A can be omitted.

THEOREM 3. *Suppose that for some* $\delta>0$ $M(x)$ *satisfies*

$$|M(x)| > B|x| \quad \textit{for} \quad |x| < \delta. \tag{12}$$

Then

$$\mathsf{E}X_n^2 = \begin{cases} O\left(\frac{1}{n}\right) & \textit{if} \quad B > \frac{1}{2} \\ O\left(\frac{\log n}{n}\right) & \textit{if} \quad B = \frac{1}{2} \\ O\left(\frac{1}{n^{2B}}\right) & \textit{if} \quad B < \frac{1}{2} \end{cases}. \tag{13}$$

That is to say, for an M.R.M.P. the upper part of condition (7) in Theorem C can be omitted, and the lower part is required only in the vicinity of zero.

§ 2. The proofs

First we establish the following

LEMMA 1. *For every* $L\geqq 1$ *there exists an* ε_L *such that*

$$\mathsf{P}(|X_n| > \varepsilon_L) = O\left(\frac{1}{n^L}\right). \tag{14}$$

We will make use of the following inequalities:

If ξ is a random variable of finite variance and $\mathscr{F}$ is a σ-algebra (a sub-σ-algebra

of the underlying σ-algebra) and $\chi(\xi<\alpha)$ stands for the indicator function of the event $\{\xi<\alpha\}$, then we have for $\alpha>0$ with probability 1:

$$\mathsf{P}(\xi > \alpha|\mathscr{F}) = \mathsf{E}(\chi(\xi > \alpha)|\mathscr{F}) \leqq \frac{\mathsf{E}(\xi^2|\mathscr{F})}{\alpha^2} \tag{15}$$

and

$$\mathsf{E}(\xi\chi(\xi > \alpha)|\mathscr{F}) \leqq \frac{\mathsf{E}(\xi^2|\mathscr{F})}{\alpha}. \tag{16}$$

(15) and (16) are easy consequences of the definition of conditional expectation.

$\mathscr{F}_i$ will always stand for the σ-algebra generated by $Y_1, Y_2, \ldots, Y_{i-1}$ or, which is the same, by $X_1, X_2, \ldots, X_i$.

PROOF OF LEMMA 1.

Given $\varepsilon>\delta>0$, $\delta>X_1$ define the events

$$A_k = \{X_k \leqq \delta, X_{k+1} > \delta, \ldots, X_n > \delta\}.$$

Then we have

$$\mathsf{P}(X_n > \varepsilon) = \sum_{k=1}^{n-1} \mathsf{P}(X_n > \varepsilon, A_k).$$

We want to show that choosing ε and δ large enough, we get

$$\mathsf{P}(X_n > \varepsilon, A_k) \leqq \frac{C}{n^{L+1}} \quad \text{for} \quad k = 1, 2, \ldots, n-1.$$

(The probability $\mathsf{P}(X_n<-\varepsilon)$ can be estimated similarly.)

On the set A_k the relation $X_n = X_k - \sum_{i=k}^{n-1} Z_i > \varepsilon$ implies $\sum_{i=k}^{n-1} Z_i < -(\varepsilon-\delta)$. Define the variables

$$U_i = \begin{cases} \dfrac{\mathscr{M}(\delta)+Y_i}{i} & \text{if} \quad \left|\dfrac{\mathscr{M}(\delta)+Y_i}{i}\right| \leqq 1 \\ \operatorname{sign}\dfrac{\mathscr{M}(\delta)+Y_i}{i} & \text{otherwise.} \end{cases}$$

Then we have on the set A_k for $k<i\leqq n$

$$Z_i \geqq U_i$$

for $1>Z_i$ implies

$$1 \geqq Z_i \geqq \frac{M(X_i)+Y_i}{i} \geqq \frac{\mathscr{M}(\delta)+Y_i}{i}.$$

Thus it is enough to estimate the probability

$$\mathsf{P}\left(\sum_{i=k}^{n-1} U_i < -(\varepsilon-\delta)\right)$$

which is more convenient since the U_i's involve only the variables Y_i and not the X_i.

For $T>1$ define

$$U_i' = \begin{cases} U_i & \text{if } |U_i| \leqq \dfrac{1}{T} \\ 0 & \text{otherwise,} \end{cases}$$

$$U_i'' = U_i - U_i'.$$

Then we actually have

$$U_i' = \begin{cases} \dfrac{\mathcal{M}(\delta)+Y_i}{i} & \text{if } \left|\dfrac{\mathcal{M}(\delta)+Y_i}{i}\right| \leqq \dfrac{1}{T} \\ 0 & \text{otherwise.} \end{cases}$$

We can write

$$\mathsf{P}\left(\sum_{i=k}^{n-1} U_i < -(\varepsilon-\delta)\right) = \mathsf{P}\left(\sum_{i=k}^{n-1}(-U_i') > \beta\right) + \mathsf{P}\left(\sum_{i=k}^{n-1}(-U_i'') > \varepsilon-\delta-\beta\right) =$$

$$= \mathsf{P}\left(\sum_{i=k}^{n-1}(-U_i'T) > \beta T\right) + \mathsf{P}\left(\sum_{i=k}^{n-1}(-U_i'') > \varepsilon-\delta-\beta\right) = P_1 + P_2$$

where the number β (as well as T, δ and ε) will be chosen later.

First estimate P_1. Since $|U_i'T| \leqq 1$, we have

$$P_1 = \mathsf{P}\left(\exp\left(\sum_{i=k}^{n-1}(-U_i'T)\right) > e^{\beta T}\right) \leqq e^{-\beta T}\mathsf{E}\exp\left(\sum_{i=k}^{n-1}(-U_i'T)\right) \leqq$$

$$\leqq e^{-\beta T}\mathsf{E}\prod_{i=k}^{n-1}(1-U_i'T+U_i'^2T^2) =$$

$$\leqq e^{-\beta T}\prod_{i=k}^{n-1}\operatorname{ess\,sup}|\mathsf{E}(1-U_i'T+U_i'^2T^2|\mathcal{F}_i)|.$$

Applying (16) we get

$$\mathsf{E}(U_i'|\mathcal{F}_i) = \mathsf{E}\left(\frac{\mathcal{M}(\delta)+Y_i}{i}\chi\left(\left|\frac{\mathcal{M}(\delta)+Y_i}{i}\right| \leqq \frac{1}{T}\right)\middle|\mathcal{F}_i\right) =$$

$$= \frac{\mathcal{M}(\delta)}{i} - \mathsf{E}\left(\frac{\mathcal{M}(\delta)+Y_i}{i}\chi\left(\left|\frac{\mathcal{M}(\delta)+Y_i}{i}\right| > \frac{1}{T}\right)\middle|\mathcal{F}_i\right) \geqq$$

$$\geqq \frac{\mathcal{M}(\delta)}{i} - T\mathsf{E}\left(\left(\frac{\mathcal{M}(\delta)+Y_i}{i}\right)^2\middle|\mathcal{F}_i\right) \geqq \frac{\mathcal{M}(\delta)}{i} - \frac{T}{i^2}(K+\mathcal{M}^2(\delta))$$

and

$$\mathsf{E}(U_i'^2|\mathcal{F}_i) \leqq \mathsf{E}\left(\left(\frac{\mathcal{M}(\delta)+Y_i}{i}\right)^2\middle|\mathcal{F}_i\right) \leqq \frac{K+\mathcal{M}^2(\delta)}{i^2}$$

(both inequalities are meant with probability 1). Since

$$1-U_i'T+U_i'^2T^2 \geqq 0$$

we have

$$P_1 \leqq e^{-\beta T} \prod_{i=k}^{n-1} \left(1 - \mathcal{M}(\delta)\frac{T}{i} + 2\left(K + \mathcal{M}^2(\delta)\right)\frac{T^2}{i^2}\right) \leqq \tag{17}$$

$$\leqq e^{-\beta T} \prod_{i=k}^{n-1} \left(1 - \frac{\mathcal{M}(\delta)}{2}\frac{T}{i}\right) \leqq \exp\left(-\beta T - \frac{\mathcal{M}(\delta)}{2}\sum_{i=k}^{n-1}\frac{T}{i}\right),$$

if only

$$\frac{\mathcal{M}(\delta)T}{i} \leqq 1, \quad 2\left(K + \mathcal{M}^2(\delta)\right)\frac{T^2}{i^2} \leqq \frac{\mathcal{M}(\delta)T}{2i}$$

i.e.

$$T \leqq \frac{k}{4}\frac{\mathcal{M}(\delta)}{K + \mathcal{M}^2(\delta)} \tag{18}$$

holds.

For estimating P_2 we remark that $|U_i''| \leqq 1$ implies $P_2 \leqq \mathsf{P}$(at least $[\varepsilon - \delta - \beta] = \alpha$ of the variables $U_k'', \ldots, U_{n-1}''$ is different from zero). Applying (15) we get

$$\mathsf{P}\left(U_{i_\mu}'' = 0 \ (\mu = 1, 2, \ldots, \alpha)\right) =$$

$$= \mathsf{P}\left(\left|\frac{\mathcal{M}(\delta) + Y_{i_\mu}}{i_\mu}\right| > \frac{1}{T} (\mu = 1, 2, \ldots, \alpha)\right) \leqq \prod_{\mu=k}^{\alpha} \left(K + \mathcal{M}^2(\delta)\right)\frac{T^2}{i_\mu^2},$$

whence

$$P_2 \leqq \frac{\left(\left(K + \mathcal{M}^2(\delta)\right)\left(\sum_{i=k+1}^{n} \frac{T^2}{i^2}\right)\right)^{\alpha}}{\alpha!}. \tag{19}$$

Now we choose the values δ, ε, T and β in such a way that both (17) and (19) would give an estimation $O\left(\frac{1}{n^{L+1}}\right)$.

If $k > \sqrt{n}$, put $\beta = 1$, $T = (L+1)\log n$ $\varepsilon - \delta \geqq 2L+5$ whence $\alpha \geqq 2L+3$. (In this case (18) holds automatically.) By (17) we have

$$P_1 \leqq \exp\left(-(L+1)\log n\right) = \frac{1}{n^{L+1}}$$

and by (19)

$$P_2 \leqq \left(\left(K + \mathcal{M}^2(\delta)\right)\frac{(L+1)^2 \log^2 n}{\sqrt{n}}\right)^{2L+3} \leqq \frac{1}{n^{L+1}}$$

if n is large enough (not depending on k). If $k \leqq \sqrt{n}$, put $\beta = -\frac{M(\delta)}{8}\log n$, $T = \frac{8(L+1)}{M(\delta)}$. By (17) we have

$$P_1 \leqq \exp\left(\frac{\mathcal{M}(\delta)}{8} T \log n - \frac{\mathcal{M}(\delta)}{2} T \frac{\log n}{2}\right) =$$

$$\leqq \exp\left(-\frac{\mathcal{M}(\delta)}{8} T \log n\right) = \frac{1}{n^{L+1}},$$

(provided that (18) holds) and by (19)

$$P_2 \leqq \left(\frac{(K+\mathcal{M}^2(\delta))2T^2}{\frac{\alpha}{3}}\right)^{\alpha} \leqq \left(\frac{6(K+\mathcal{M}^2(\delta))T^2}{-\beta-1}\right)^{-\beta-1} \leqq \frac{1}{n^{L+1}}$$

if n is large enough.

Now we define the numbers $\varepsilon > \delta > 0$. Put

$$\delta = X_1 + 32(L+1)\frac{K+\mathcal{M}^2(\delta)}{\mathcal{M}^2(\delta)} \quad \text{and} \quad \varepsilon = \varepsilon_L = 2L+5+\delta.$$

Since $X_{k+1} \leqq X_1 + k$, if the index k is so small that (18) does not hold then $X_{k+1} \leqq \delta$, whence $\mathsf{P}(A_k) = 0$.

PROOF OF THEOREM 1. We prove only the statement $\mathsf{P}(X_n > \varepsilon) = O\left(\frac{1}{n}\right)$ the probability $\mathsf{P}(X_n < -\varepsilon)$ can be estimated similarly. We have

$$P = \mathsf{P}(X_n > \varepsilon) \leqq \mathsf{P}(X_{cn} > \varepsilon_2) + \mathsf{P}\left(X_n > \varepsilon,\ X_{cn} \leqq \varepsilon_2,\ \bigcap_{cn \leqq i < n}\left\{X_i > \frac{\varepsilon}{2}\right\}\right) +$$

$$+ \mathsf{P}\left(\bigcup_{cn \leqq k < n}\left\{X_n > \varepsilon,\ X_k \leqq \frac{\varepsilon}{2},\ X_{k+1} > \frac{\varepsilon}{2}, \ldots, X_{n-1} > \frac{\varepsilon}{2}\right\}\right)$$

where ε_2 has been introduced in Lemma 1 and the index cn is meant as $[cn]$ $(0 < c < 1)$. The first term is $O\left(\frac{1}{n^2}\right)$, by the definition of ε_2, and we have

$$P \leqq O\left(\frac{1}{n^2}\right) + \mathsf{P}\left(\sum_{cn \leqq i < n} Z_i < \varepsilon_2 - \varepsilon,\ \bigcap_{cn \leqq i < n}\left\{X_i > \frac{\varepsilon}{2}\right\}\right) +$$

$$+ \mathsf{P}\left(\sup_{cn \leqq k < n} \sum_{i=k}^{n-1} Z_i < -\frac{\varepsilon}{2},\ X_{k+1} > \frac{\varepsilon}{2},\ X_{k+2} > \frac{\varepsilon}{2}, \ldots, X_n > \frac{\varepsilon}{2}\right).$$

Define the variables U_i analogously as in the proof of Lemma 1, as follows

$$U_i = \begin{cases} \dfrac{\mathcal{M}\left(\frac{\varepsilon}{2}\right) + Y_i}{i} & \text{if} \quad \left|\dfrac{\mathcal{M}\left(\frac{\varepsilon}{2}\right) + Y_i}{i}\right| \leqq 1 \\ \operatorname{sign}\left(\dfrac{\mathcal{M}\left(\frac{\varepsilon}{2}\right) + Y_i}{i}\right) & \text{otherwise} \end{cases}$$

On the set $\left\{X_i \geqq \frac{\varepsilon}{2}\right\}$ we have — just as before — the inequality

$$Z_i \geqq U_i$$

Thus

$$P \leqq O\left(\frac{1}{n^2}\right)+\mathsf{P}\Big(\sum_{cn\leqq i<n} U_i < \varepsilon_2-\varepsilon\Big)+\mathsf{P}\left(\sup_{cn\leqq k<n}\left(Z_k+\sum_{i=k+1}^{n} U_i\right) < -\frac{\varepsilon}{2}\right) \leqq$$

$$\leqq O\left(\frac{1}{n^2}\right)+\mathsf{P}\left(\sum_{cn\leqq i<n}\frac{\mathscr{M}\left(\frac{\varepsilon}{2}\right)+Y_i}{i} < \varepsilon_2-\varepsilon\right)+$$

$$+\mathsf{P}\left(\sup_{cn\leqq k<n}\sum_{i=k+1}^{n-1}\frac{\mathscr{M}\left(\frac{\varepsilon}{2}\right)+Y_i}{i} < -\frac{\varepsilon}{4}\right)+\sum_{cn\leqq k<n}\mathsf{P}\left(Z_k \leqq -\frac{\varepsilon}{4}\right)+$$

$$+2\mathsf{P}\left(\bigcup_{cn\leqq i<n}\left(U_i \neq \frac{\mathscr{M}\left(\frac{\varepsilon}{2}\right)+Y_i}{i}\right)\right) = O\left(\frac{1}{n}\right)+$$

$$+\mathsf{P}\left(\sum_{cn\leqq i<n}\frac{Y_i}{i} \leqq \varepsilon_2-\varepsilon-\mathscr{M}\left(\frac{\varepsilon}{2}\right)\log\frac{1}{c}\right)+\mathsf{P}\left(\sup_{cn\leqq k<n}\sum_{i=k+1}^{n-1}\frac{Y_i}{i} < -\frac{\varepsilon}{4}\right)+$$

$$+2\mathsf{P}\left(\bigcup_{cn\leqq i<n}\left\{\left|\frac{\mathscr{M}\left(\frac{\varepsilon}{2}\right)+Y_i}{i}\right| > 1\right\}\right).$$

Choosing c small enough so that $\varepsilon_2-\varepsilon-M(\varepsilon/2)\log 1/c<0$ and applying the Kolmogorov inequality (Y_i forms a martingale-difference) we get

$$\mathsf{P}(X_n > \varepsilon) = O\left(\frac{1}{n}\right).$$

PROOF OF THEOREM 2. By Theorem 1 we have

$$\sum_{n=1}^{\infty}\mathsf{P}(|X_{n^2}| > \varepsilon) < \infty.$$

Now

$$\mathsf{P}\Big(\big|\sup_{n^2\leqq i<(n+1)^2} X_i\big| > \varepsilon\Big)$$

has actually been estimated in the proof of Theorem 1, and we have obtained $O\left(\frac{1}{(n+1)^2}\right)$, that proves our theorem.

LEMMA 2 (see [3]). *Let $\{b_n\}$ be a sequence of nonnegative numbers such that for all n (large enough)*

$$b_{n+1} \leqq \left(1-\frac{c_n}{n}\right)b_n+\frac{d}{n^2}$$

where $d>0$ *and* $c_n \to c$ *as* $n \to \infty$. *Then*

$$b_n = \begin{cases} O\left(\frac{1}{n}\right) & \text{if} \quad c > 1 \\ O\left(\frac{\log n}{n}\right) & \text{if} \quad c = 1. \\ O\left(\frac{1}{n^c}\right) & \text{if} \quad c < 1 \end{cases}$$

PROOF OF THEOREM 3. Introduce the following notations:

$$U(x) = M(x) - Bx,$$

$$nZ_n - M(X_n) = \overline{Y}_n$$

(thus $\overline{Y}_n = Y_n$, if $|M(x_n)+Y_n| \leqq n$),

$$F_n = \{|X_n| < \delta\}.$$

Let $a>0$ be a constant for which $\mathsf{P}(|X_n|>a) = O\left(\frac{1}{n^4}\right)$. Set $G_n=\{|X_n|<a\}$. Then we clearly have (for a sufficiently large n).

$$\int_{G_n} |\mathsf{E}(\overline{Y}_n|\mathscr{F}_n)| = \int_{G_n} |\mathsf{E}(\overline{Y}_n - Y_n|\mathscr{F}_n)| \leqq \int_{G_n} \mathsf{E}(|\overline{Y}_n - Y_n||\mathscr{F}_n) = \int_{G_n} |\overline{Y}_n - Y_n| =$$

$$= \int_{G_n \cap \{|M(X_n)+Y_n|>n\}} |\overline{Y}_n - Y_n| \leqq \int_{\left\{|Y_n|>\frac{n}{2}\right\}} (2n+|Y_n|) \leqq \frac{8K}{n} + \sqrt{K}\sqrt{\frac{4K}{n^2}} = O\left(\frac{1}{n}\right)$$

and in a similar way we get

$$\int_{G_n} \mathsf{E}(\overline{Y}_n^2|\mathscr{F}_n) = \int_{G_n} \overline{Y}_n^2 = O(1).$$

Since

$$X_{n+1} = X_n\left(1-\frac{B}{n}\right) - \frac{1}{n}U(X_n) - \frac{1}{n}\overline{Y}_n$$

we get

$$\int_{G_n} X_{n+1}^2 = \left(1-\frac{B}{n}\right)^2 \int_{G_n} X_n^2 + \frac{1}{n^2}\int_{G_n} U^2(X_n) + \frac{1}{n^2}\int_{G_n} \overline{Y}_n^2 -$$

$$-\frac{2}{n}\left(1-\frac{B}{n}\right)\int_{G_n} X_n U(X_n) - \frac{2}{n}\left(1-\frac{B}{n}\right)\int_{G_n} X_n \overline{Y}_n + \frac{2}{n^2}\int_{G_n} U(X_n)\overline{Y}_n.$$

The integrals of the right-hand side can be estimated as follows:

$$\int_{G_n} U^2(X_n) = O(1), \quad \int_{G_n} \overline{Y}_n^2 = O(1),$$

$$-\int_{G_n} X_n U(X_n) = -\int_{G_n \cap F_n} X_n U(X_n) - \int_{G_n \cap \bar{F}_n} X_n U(X_n) \leqq$$

$$\leqq -\int_{G_n \cup \bar{F}_n} X_n U(X_n) = O(1)\mathsf{P}(\bar{F}_n) = O\left(\frac{1}{n}\right),$$

$$\int_{Gn} X_n \overline{Y}_n = \int_{G_n} X_n \mathsf{E}(\overline{Y}_n | \mathscr{F}_n) = O\left(\frac{1}{n}\right),$$

$$\int_{G_n} U(X_n)\, \overline{Y}_n = O(1).$$

Hence we have

$$\int_{G_n} X_{n+1}^2 \leqq \left(1 - \frac{B}{n}\right)^2 \int_{\Omega} X_n^2 + \frac{d'}{n^2}$$

for a suitable d'.

Further, since $|X_{n+1}| \leqq |X_1| + n$, we have

$$\int_{\Omega} X_{n+1}^2 = \int_{G_n} X_{n+1}^2 + \int_{\bar{G}_n} X_{n+1}^2 \leqq \int_{G_n} X_{n+1}^2 + O\left(\frac{1}{n^2}\right)$$

i.e.

$$\int_{\Omega} X_{n+1}^2 \leqq \left(1 - \frac{B}{n}\right)^2 \int_{\Omega} X_n^2 + \frac{d}{n^2}$$

for some $d > 0$, and by Lemma 2 we have our Theorem 3.

§ 3. A counterexample

In this § X_n denotes the original R.M.P. again. If we are given $\varepsilon > 0$, $\alpha > 0$ and a positive integer T then we can construct a R.M.P. such that

$$\lim_{|x| \to \infty} \frac{M(x)}{x^{1-\varepsilon}} = 1,$$

$$M'(0) = \alpha,$$

$$\mathsf{E}(Y_{n+1}^{2T} | Y_1, Y_2, \ldots, Y_n) \leqq K$$

(with probability 1) and

$$\mathsf{E}(X_n^2) \geqq \frac{C}{8(\log n)^{4T/\varepsilon}}$$

where

$$C = \left(2 \sum_{k=1}^{\infty} \frac{1}{k^{2T+2}}\right)^{-1}.$$

PROOF. Let

$$M(x) = \begin{cases} f(x) & \text{if } |x| \leqq 1 \\ (\operatorname{sign} x)|x|^{1-\varepsilon} & \text{if } |x| > 1 \end{cases}$$

where $f(x)$ is a monotonically increasing function defined in $[-1, +1]$ for which

$$f(0) = 0 \quad f'(0) = \alpha$$

$$-1 \leqq f(-1) < f(1) \leqq +1.$$

Further let $X_1 = 0$ and $Y_1, Y_2, \ldots$ be a sequence of i.i.r.v.'s for which

$$\mathsf{P}(Y_n = \pm k) = \frac{C}{k^{2T+2}} \qquad (k = 1, 2, \ldots)$$

Let the events A_n, B_n and C_n defined by

$$A_n = \{Y_1 = -[(\log n)^{2/\varepsilon}]\} = \{X_2 = [(\log n)^{2/\varepsilon}]\};$$

$$B_n = \left\{ \sup_{1 \leqq k \leqq n} \left| \sum_{j=2}^{k} \frac{Y_j}{j} \right| \geqq \frac{\pi}{\sqrt{3}} \right\},$$

$$C_n = A_n \cap \bar{B}_n.$$

Then clearly

$$\mathsf{P}(A_n) = \frac{C}{[(\log n)^{2/\varepsilon}]^{2T+2}}, \quad \mathsf{P}(B_n) \leqq \frac{1}{2},$$

$$\mathsf{P}(C_n) \geqq \frac{C}{2[(\log n)^{2/\varepsilon}]^{2T+2}}$$

and on the set C_n the inequality

$$X_{n+1} \geqq [(\log n)^{2/\varepsilon}] - \sum_{k=2}^{n} \frac{\left[(\log n)^{2/\varepsilon} + \frac{\pi}{\sqrt{3}}\right]^{1-\varepsilon}}{k} - \frac{\pi}{\sqrt{3}} \geqq \frac{(\log n)^{2/\varepsilon}}{2}$$

holds (if n is big enough).

Now we have our statement as follows:

$$\mathsf{E}(X_{n+1}^2) = \int_{\Omega} X_{n+1}^2 \geqq \int_{C_n} X_{n+1}^2 \geqq \frac{(\log n)^{4/\varepsilon}}{4} \cdot \frac{C}{2[(\log n)^{2/\varepsilon}]^{2T+2}} = \frac{8(\log n)^{4T/\varepsilon}}{C}.$$

REFERENCES

[1] BLUM, J. R.: A stochastic approximation method, *Ann. Math. Stat.* **22** (1951) 400—407.
[2] CHUNG, K. L.: On a stochastic approximation method. *Ann. Math. Stat.* **25** (1954) 463—483.
[3] VENTER, J. H.: On Dvoretzky Stochastic Approximation Theorems. *Ann. Math. Statist.* **37** (1966) 1534—1544.
[4] Невельсон, М. Б.—Хасьминский, Р.З.: *Стохастическая аппроксимация и рекуррентное оценивание.* Наука 1972 Москва.
[5] HODGES, J. L.—LEHMAN, E. L.: Two approximations to the Robbins-Monro process. *Proc. Third Berkeley Symp.* (1956) **1** 95—104.
[6] KOMLÓS, J.—RÉVÉSZ, P.: On the rate of convergence of the Robbins-Monro method. *Zeitschrift für Wahrscheinlichkeitstheorie* **25** (1972) 39—47.
[7] VENTER, J. H.: An extension of the Robbins—Monro procedure. *Ann. Math. Statist.* **38** (1967) 181—190.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received March 20, 1973)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 341—345.

BOUNDS ON ERROR PROBABILITY

by

T. NEMETZ

Summary

In this short note we give upper and lower bounds on the error-probability of the decision problem for r simple hypotheses in terms of the error-probabilities of test problems with two simple hypotheses. The bounds are functions of the variation distances of pairs of possible probability measures. An example is given showing that the upper bounds are the best possible in an asymptotic sense.

Introduction and notations

Let ϑ be a finite random variable, called *parameter,* taking on the values $1, 2, \ldots, r$, and let

$$q_k = \text{Prob}\,(\vartheta = k), \qquad k = 1, 2, \ldots, r.$$

Let ξ be another random variable, called *observation,* taking values on a measurable space $\{X, \mathcal{X}\}$, the distribution of which is given by the probability measure P_k under the hypothesis $\mathcal{H}_k = \{\vartheta = k\}$, $k = 1, 2, \ldots, r$.

On the basis of the observation ξ we want to infer the actual value of the parameter. For a given decision rule d, the *error-probability* of the decision $d = d(\xi)$ is defined as $e(d) = \text{Prob}\,(d(\xi) \neq \vartheta)$.

It is well known that there is a decision Δ yielding minimal error probability. Following RÉNYI we call such decisions standard decisions. A variant of the standard decision is the following: Having observed $\xi = x$ let $d(\xi)$ be the smallest value k for which $\vartheta = k$ has greatest posterior probability (see RÉNYI [1]). For this decision it is easy to see that

$$e(\Delta) = \int_X \left[\sum_{i=1}^{r} q_i p_i(x) - \max_{1 \leq i \leq r} q_i p_i(x) \right] \mu(dx) \tag{1}$$

where $p_i(x)$ denotes the Radon—Nikodym derivative of P_i with respect to a dominating (σ-finite) measure μ.

We will use two more notations:

$e(Q, \mathcal{P})$ stands for the error probability $e(\Delta)$ where $Q = \{q_1, \ldots, q_r\}$ is the (prior) distribution of the parameter and $\mathcal{P} = \{P_1, \ldots, P_r\}$ is the set of (conditional) distributions of the observation.

If $Q = \{q_1, q_2 = 1 - q_1\}$, $\mathcal{P} = \{P_1, P_2\}$, we will use the notation $e(q_1, P_1, P_2)$ instead of $e(Q, \mathcal{P})$.

In this paper we give upper and lower bounds on $e(Q, \mathcal{P})$ by

$$\max_{i \neq k} e\left(\frac{1}{2}, P_i, P_k\right).$$

The results

To begin with we state two simple lemmas:

LEMMA 1.

$$2 \cdot \min(q_1, q_2) e\left(\frac{1}{2}, P_1, P_2\right) \leqq e(q_1, P_1, P_2) \leqq \tag{2}$$

$$\leqq 2 \max(q_1, q_2) e\left(\frac{1}{2}, P_1, P_2\right)$$

PROOF. If $r=2$, then (1) is equivalent to the following:

$$e(q_1, P_1, P_2) = \int_X \min[q_1 p_1(x), q_2 p_2(x)] \mu(dx),$$

and here

$$2 \min(q_1, q_2) \min\left(\frac{1}{2} p_1(x), \frac{1}{2} p_2(x)\right) \leqq \min\left(q_1 p_1(x), q_2 p_2(x)\right) \leqq$$

$$\leqq 2 \max(q_1, q_2) \min\left(\frac{1}{2} p_1(x), \frac{1}{2} p_2(x)\right)$$

which proves the lemma.

LEMMA 2. *Let $a_1, a_2, \ldots, a_r$ be arbitrary real numbers. Then*

$$\sum_{i=1}^{r} a_i - \max_{1 \leqq i \leqq r} a_i = \sum_{i=2}^{r} \min[a_i, \max_{1 \leqq k < i} a_k]$$

PROOF. Observing that

$$a_1 + a_2 - \max(a_1, a_2) = \min(a_1, a_2)$$

a simple induction results in the lemma.

Now we are able to prove our results.

THEOREM. *The following inequalities hold between $e(Q, \mathscr{P})$ and $\max_{i \neq k} e\left(\frac{1}{2}, P_i, P_k\right)$:*

$$e(Q, \mathscr{P}) \leqq (r-1)(1 - \min_{1 \leqq i \leqq r} q_i) \cdot 2 \max_{i \neq k} e\left(\frac{1}{2}, P_i, P_k\right) \tag{3}$$

$$e(Q, \mathscr{P}) \leqq \frac{r(r-1)}{2} \max_{1 \leqq i \leqq r} q_i \cdot 2 \max_{i \neq k} e\left(\frac{1}{2}, P_i, P_k\right) \tag{4}$$

$$e(Q, \mathscr{P}) \geqq [1 - \max_{1 \leqq i \leqq r} q_i] \cdot 2 \min_{i \neq k} e\left(\frac{1}{2}, P_i, P_k\right) \tag{5}$$

$$e(Q, \mathscr{P}) \geqq (\min_{1 \leqq i \leqq r} q_i) \cdot 2 \max_{1 \leqq k \leqq r} \sum_{j \neq k} e\left(\frac{1}{2}, P_k, P_j\right) \tag{6}$$

PROOF. Making use of Lemma 2 it follows from (1) that

$$e(Q, \mathscr{P}) = \sum_{i=2}^{r} \int_X \min [q_i p_i(x), \max_{1 \leq k < i} q_k p_k(x)] u(dx) \tag{7}$$

Therefore

$$e(Q, \mathscr{P}) \leqq \sum_{i=2}^{r} \int_X \sum_{k=1}^{i-1} \min \left(q_i p_i(x), q_k p_k(x)\right) \mu(dx)$$

Using Lemma 1, we have

$$e(Q, \mathscr{P}) \leqq \sum_{i=2}^{r} \sum_{k=1}^{i-1} 2 \max (q_i, q_k) \int_X \min \left(\frac{1}{2} p_i(x), \frac{1}{2} p_k(x) \right) \mu(dx) \leqq$$

$$\leqq 2 \max_{i \neq k} e\left(\frac{1}{2}, P_i, P_k\right) \sum_{i=2}^{r} \sum_{k=1}^{i-1} \max (q_i, q_k)$$

Here, without any loss of generality we may assume that $q_1 \leqq q_2 \leqq \ldots \leqq q_r$, and thus we get

$$e(Q, \mathscr{P}) \leqq 2 \max_{i \neq k} e\left(\frac{1}{2}, P_i, P_k\right) \sum_{i=2}^{r} (i-1) q_i$$

Applying the inequality

$$\sum_{i=2}^{r} (i-1) q_i \leqq (r-1) \sum_{i=2}^{r} q_i = (r-1)(1 - \min_{1 \leq i \leq r} q_i)$$

we get (3), while the inequality

$$\sum_{i=2}^{r} \sum_{k=1}^{i-1} \max (q_i, q_k) \leqq \max_{1 \leq i \leq r} q_i \sum_{i=2}^{r} (i-1) = \frac{r(r-1)}{2} \max_{1 \leq i \leq r} q_i$$

yields (4).

On the other hand, since

$$\int_X \min [q_i p_i(x), \max_{1 \leq k \leqq i} q_k p_k(x)] \mu(dx) \geqq \int \min [q_i p_i(x), q_k p_k(x)] \mu(dx) \geqq \tag{8}$$

$$\geqq 2 \min (q_i, q_k) e\left(\frac{1}{2}, P_i, P_k\right), \qquad k = 1, 2, \ldots, i-1,$$

we have

$$e(Q, \mathscr{P}) \geqq 2 \min_{i \neq k} e\left(\frac{1}{2}, P_i, P_k\right) \sum_{i=2}^{r} \min (q_i, q_{i-1}).$$

Supposing again that $q_1 \leqq q_2 \leqq \ldots \leqq q_r$ we get (5).

Finally, (6) is true, for choosing $k=1$ in (8) we have

$$e(Q, \mathscr{P}) \geqq \sum_{i=2}^{r} 2 \min (q_1, q_i) e\left(\frac{1}{2}, P_1, P_i\right) \geqq$$

$$\geqq (\min_{1 \leqq i \leqq r} q_i) 2 \sum_{i=2}^{r} e\left(\frac{1}{2}, P_1, P_i\right).$$

Here the role of P_1 and P_k, $k \neq 1$ may be changed, and this proves (6).

Remark 1: The bound given in (4) is better than the one given in (3), provided $\min q_i > \frac{1}{2} \max q_i$, which is the case, e.g. if $q_i \equiv 1/r$.

If $\left(\frac{1}{2}, P_i, P_k\right)$ does not depend on (i, k), then (5) is stronger than (6), while, if $q_i = 1/r$, $i=1, 2, \ldots, r$ then, in general, (6) is better than (5).

The following example will illustrate, how good the upper bounds in (3) and (4) are.

Example: Let $r = k+n$, let P_i be the uniform distribution on the interval $[(i-1)n, in]$, $i=1, 2, \ldots, k$ and P_{k+i}, $i=1, 2, \ldots, n$ be the uniform distribution on the set

$$\bigcup_{j=1}^{k} \left[(j-1)n+(i-1), (j-1)n+(i-1)\frac{1}{k}\right],$$

$$\mathscr{P} = \{P_1, \ldots, P_{k+n}\}, \quad Q = \{q_1, \ldots, q_{k+n}\},$$

where

$$q_i = \begin{cases} \dfrac{1}{k+1} & \text{if} \quad 1 \leqq i \leqq k \\ \dfrac{1}{n(k+1)} & \text{if} \quad k < i \leqq n+k \end{cases}$$

It is a routin work to check that

$$e(Q, \mathscr{P}) = \frac{1}{k+1}$$

and

$$2 \max_{i \neq j} e\left(\frac{1}{2}, P_i, P_j\right) = 2e\left(\frac{1}{2}, P_1, P_{k+1}\right) = \frac{1}{kn}$$

thus

$$\frac{e(Q, \mathscr{P})}{2 \max\limits_{i \neq j} e\left(\frac{1}{2}, P_i, P_j\right)} = \frac{k}{k+1} n. \tag{9}$$

COROLLARY. *Suppose $\mathscr{P}_r$ consists of r probability distributions and Q_r is concentrated on r values. Then*

$$\lim_{r\to\infty}\sup_{(Q_r,\mathscr{P}_r)}\frac{e(Q_r,\mathscr{P}_r)}{2r\max\limits_{i\neq j}e\left(\frac{1}{2},P_i,P_j\right)}=1. \tag{10}$$

PROOF. It follows from (3) that

$$\frac{e(Q_r,\mathscr{P}_r)}{2r\max\limits_{i\neq j}e\left(\frac{1}{2},P_i,P_j\right)}<1.$$

On the other hand, choosing $k=\sqrt{r}$ in the previous example we see from (9) that the limit in (10) is at least 1.

Remark 2: Our results can be trivially extended to the problem of minimal Bayesian loss, where the loss-function is bounded.

Remark 3: Considering the variation distance

$$\varrho(P,Q)=\sup_{A\in\mathscr{X}}|P(A)-Q(A)|$$

of two probability measures P and Q on $\{X,\mathscr{X}\}$, it is easy to see that

$$2e\left(\frac{1}{2},P,Q\right)=1-\varrho(P,Q).$$

Therefore $2e\left(\frac{1}{2},P_i,P_k\right)$ in the theorem can be replaced by $1-\varrho(P_i,P_k)$. In this form our theorem provides improvements of a result of I. VAJDA, see [2], Th. 2.

REFERENCE

[1] RÉNYI, A.: Statistics and Information Theory, *Studia Sci. Math. Hung.* **2** (1967) 249—256.
[2] VAJDA, I.: On the Statistical Decision Problems with Finite Parameter Space, *Kybernetika,* **3** (1967) № 5, 451—466.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received April 10, 1973)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 347—361.

ON A NON-PARAMETIC ESTIMATION OF THE REGRESSION FUNCTION

by

P. MAJOR

The regression function is usually estimated by a polynomial. In this paper an other method is investigated, similar to one of the methods of density estimation.

It is the following method: Let (ξ_i, η_i) $i=1, 2, \ldots$ be independent identically distributed two-dimensional random variables, $\mathsf{P}(0 \leqq \eta_i \leqq 1)=1$. Let us divide the interval $[0, 1]$ into a_n equal parts. We estimate the function $R(x)=\mathsf{E}(\xi|\eta=x)$ by the following $R_n(x)$

$$R_n(x) = \frac{\sum_{\substack{\eta_i \in \left[\frac{k}{a_n}, \frac{k+1}{a_n}\right] \\ (i \leqq n)}} \xi_i}{\text{the number of } \eta_i - \text{s } (i \leqq n) \text{ falling into the interval } \left[\frac{k}{a_n}, \frac{k+1}{a_n}\right]} \tag{1}$$

$$x \in \left[\frac{k}{a_n}, \frac{k+1}{a_n}\right] \qquad (k = 1, 2, \ldots, a_n).$$

We investigate the order of magnitude of $R_n(x)-R(x)$ in supremum norm. We introduce some notations. $f(x)$ is the density function of η_i,

$$\sigma^2(x) = \mathsf{E}([\xi_i - \mathsf{E}(\xi_i|\eta_i)]^2|\eta_i = x).$$

Our first statement is the following

THEOREM 1. *Let us suppose that $f(x)$, $\sigma(x)$, $R(x)$ are differentiable, $|f'(x)| \leqq K$, $|\sigma'(x)| \leqq K$, $|R'(x)| \leqq K$ and $f(x) > K'$, $\sigma(x) > K'$ for appropriate $K, K' > 0$ and for every x. Let us suppose further that $\mathsf{E}(e^{t_0|\xi|}|\eta = x) \leqq c$ for appropriate $t_0 > 0$ and c for every $0 \leqq x \leqq 1$. If $n^\alpha < a_n < n^\beta$, $\frac{1}{3} < \alpha < \beta < 1$ then*

$$\mathsf{P}\left(\sup_x \sqrt{\frac{nf(x)}{a_n}} \cdot \frac{|R_n(x)-R(x)|}{\sigma(x)} < \sqrt{2\log a_n - \log\log a_n + y}\right) \to \exp\left(-\frac{e^{-y/2}}{\sqrt{\pi}}\right) \tag{2}$$

$$\mathsf{P}\left(\sup_x \sqrt{\frac{nf(x)}{a_n}} \frac{R_n(x)-R(x)}{\sigma(x)} < \sqrt{2\log a_n - \log\log a_n + y}\right) \to \exp\left(\frac{-e^{-y/2}}{2\sqrt{\pi}}\right) \tag{3}$$

as $n \to \infty$.

If the functions $f(x)$, $R(x)$, $\sigma(x)$ are smooth enough, it is worth substituting the estimation $R_n(x)$ by

$$\bar{R}_n(x) = R_n\left(\frac{k-\frac{1}{2}}{a_n}\right) + \left[R_n\left(\frac{k+\frac{1}{2}}{a_n}\right) - R_n\left(\frac{k-\frac{1}{2}}{a_n}\right)\right]\left(a_n x - k - \frac{1}{2}\right) \tag{4}$$

for

$$k-\frac{1}{2} < a_n x < k+\frac{1}{2}.$$

Then we have our.

THEOREM 1′. *If the conditions of Theorem* 1 *are fulfilled and, in addition* $f(x)$, $R(x)$ *and* $\sigma(x)$ *are differentiable twice with bounded second derivatives, then the relations* (2) *and* (3) *hold for* $n^\alpha < a_n < n^\beta$, $\frac{1}{5} < \alpha < \beta < 1$, *if we substitute* $R_n(x)$ *by* $\bar{R}_n(x)$.

Now if we want to construct a "confidence strip" i.e. we want to construct a region T in the plane such that $\mathsf{P}(R(x) \in T) > 1-\varepsilon$ with a prescribed ε, then we are interested in whether $f(x)$ and $\sigma(x)$ in the formulae (2) and (3) can be substituted by their estimations. The answer is in the affirmative.

THEOREM 2. *Let*

$$f_n(x) = \frac{a_n}{n}\left\{\text{the number of } \eta_i - \text{s } (i \leqq n) \text{ for wich } \eta_i \in \left[\frac{k}{a_n}, \frac{k+1}{a_n}\right]\right\}$$

$$\sigma_n^2(x) = \frac{a_n}{n f_n(x)} \sum_{\substack{\eta_i \in \left[\frac{k}{a_n}, \frac{k+1}{a_n}\right] \\ i \leqq n}} \xi_i^2 - \left[\frac{a_n}{n f_n(x)} \sum_{\substack{\eta_i \in \left[\frac{k}{a_n}, \frac{k+1}{a_n}\right] \\ i \leqq n}} \xi_i\right]^2$$

if the condition of Theorem 1 holds

$$x \in \left[\frac{k}{a_n}, \frac{k+1}{a_n}\right] \qquad k = 0, 1, \ldots, a_n - 1.$$

Then

$$\mathsf{P}\left(\sup_x \sqrt{\frac{n f_n(x)}{a_n}} \cdot \frac{|R_n(x) - R(x)|}{\sigma_n(x)} < \sqrt{2\log a_n - \log\log a_n + y}\right) \to \exp\left(\frac{-e^{-y/2}}{\sqrt{\pi}}\right) \tag{2′}$$

and

$$\mathsf{P}\left(\sup_x \sqrt{\frac{n f_n(x)}{a_n}} \cdot \frac{R_n(x) - R(x)}{\sigma_n(x)} < \sqrt{2\log a_n - \log\log a_n + y}\right) \to \exp\left(\frac{-e^{-y/2}}{2\sqrt{\pi}}\right) \tag{3′}$$

if.

As an immediate consequence of Theorem 1 we get that

$$\sup_x \sqrt{\frac{n f(x)}{2 a_n \log a_n}} \cdot \frac{|R_n(x) - R(x)|}{\sigma(x)} \Rightarrow 1$$

where the symbol $\Rightarrow$ means convergence in probability. The question arises whether this expression is also convergent with probability 1.

THEOREM 3. *Under the conditions of theorem* 1 *we have*

$$\mathsf{P}\left(\lim_{n\to\infty}\sup_x \sqrt{\frac{nf(x)}{2a_n\log a_n}}\,\frac{|R_n(x)-R(x)|}{\sigma(x)}=1\right)=1$$

$$\mathsf{P}\left(\lim_{n\to\infty}\sup_x \sqrt{\frac{nf(x)}{2a_n\log a_n}}\,\frac{R_n(x)-R(x)}{\sigma(x)}=1\right)=1$$

if

$$n^\alpha < a_n < n^\beta, \quad \frac{1}{3} < \alpha < \beta < 1.$$

For proving our theorems we need some lemmas:

LEMMA 1. *Let* $\xi_{n1}, \xi_{n2} \dots \xi_{nn}$, $n=1, 2 \dots$ *be a double array of random variables independent and identically distributed within a row. Let* $\mathsf{E}\xi_{n1}=0$, $\mathsf{D}^2\xi_{n1}=1$, $\mathsf{E}\,e^{t|\xi_{n1}|}<c$ *for appropriate* $t>0$ *and* c. *Then*

$$\mathsf{P}\left(\xi_{n1}+\dots+\xi_{nn} > \sqrt{n}\cdot x_n\right) \sim 1-\Phi(x_n) \quad \textit{for} \quad x_n = o\left(\sqrt[6]{n}\right)$$

where $\Phi(x)$ *is the standard normal distribution function.*

The proof is essentially the same as in [2] page 517, and we omit it.

LEMMA 2. *Let* $(\xi_i^{(n)}, \eta_i^{(n)})$ *be pairs of random variables independent and identically distributed for fixed* n,

$$\mathsf{P}(\eta_i^{(n)} = k) = p_k^{(n)}, \qquad k = 1, 2, \dots, a_n,$$

$$\sum_{k=1}^{a_n} p_k^{(n)} = 1, \quad n^\varepsilon < a_n < n^{1-\varepsilon}, \qquad \varepsilon > 0,$$

$$\mathsf{E}(\xi_i^{(n)}|\eta_i^{(n)} = k) = 0, \quad \mathsf{E}(\xi_i^{(n)2}|\eta_i^{(n)} = k) = \sigma_k^{(n)2}, \quad \frac{c_1}{a_n} < p_k^{(n)} < \frac{c_2}{a_n}$$

$$\mathsf{E}(e^{t|\xi_i^{(n)}|}|\eta_i^{(n)}) \leqq c, \quad \sigma_k^{(n)} > c'$$

for appropriate c_1, c_2, c, $t > 0$, c' *for every* k *and* n. *Then*

$$\mathsf{P}\left(\max_k \frac{\sqrt{np_k^{(n)}}}{\sigma_k^{(n)}} \cdot \frac{\sum_{i=1}^n \xi_i^{(n)} \cdot I_{(\eta_i^{(n)}=k)}}{\sum_{i=1}^n I_{(\eta_i^{(n)}=k)}} < \sqrt{2\log a_n - \log\log a_n + y}\right) \to \exp\left(\frac{-e^{-y/2}}{2\sqrt{\pi}}\right)$$

and

$$\mathsf{P}\left(\max_k \frac{\sqrt{np_k^{(n)}}}{\sigma_k^{(n)}} \cdot \left| \frac{\sum_{i=1}^n \xi_i^{(n)} \cdot I_{(\eta_i^{(n)}=k)}}{\sum_{i=1}^n I_{(\eta_i^{(n)}=k)}} \right| < \sqrt{2\log a_n - \log\log a_n + y}\right) \to \exp\left(\frac{-e^{-y/2}}{\sqrt{\pi}}\right)$$

where $I_{(\eta_i^{(n)}=k)} = 1$ *if* $\eta_i^{(n)} = k$ *and* 0 *otherwise.*

PROOF OF LEMMA 2. In the following we omit the superscript n. Let us introduce the notations $s_j = \sum_{i=1}^n \xi_i I_{(\eta_i=j)}$ and $v_j = \sum_{i=1}^n I_{(\eta_i=j)}$. Then

$$\mathsf{P}(s_1 < x_1, \ldots, s_{a_n} < x_{a_n} | v_1 = t_1 \ldots, v_{a_n} = t_{a_n}) =$$

$$= \mathsf{P}(s_1 < x_1 | v_1 = t_1) \ldots \mathsf{P}(s_{a_n} < x_{a_n} | v_{a_n} = t_{a_n})$$

Let $\tilde{\xi}_{ij}$, $i=1, 2 \ldots j=1, 2 \ldots a_n$ be — for fixed j — independent identically distributed random variables with distribution function $\mathsf{P}(\tilde{\xi}_{ij}<x)=\mathsf{P}(\xi_1=x_1|\eta_1=j)$. Then

$$\mathsf{P}(s_j < x_j | v_j = t_j) = \mathsf{P}(\tilde{\xi}_{1,j} + \ldots + \tilde{\xi}_{tj,j} < x_j)$$

Now we prove that

$$\mathsf{P}\left(|v - np_j| > \sqrt{np_j}\, n^{\frac{1-\varepsilon}{3}}\right) \leqq 2\exp\left(-\frac{n^{\frac{2}{3}(1-\varepsilon)}}{3}\right) \tag{5}$$

Really

$$\mathsf{P}\left(v_j - np_j > \sqrt{np_j}\, n^{\frac{1-\varepsilon}{3}}\right) = \mathsf{P}\left(\sum_{i=1}^n [I_{(\eta_i=j)} - p_j] > \sqrt{np_j}\, n^{\frac{1-\varepsilon}{3}}\right) =$$

$$= \mathsf{P}\left(\exp\left(n^{\frac{1-\varepsilon}{3}} \sum_{i=1}^n \frac{I_{(\eta_i=j)} - p_j}{\sqrt{np_j}}\right) > \exp\left(n^{\frac{2}{3}(1-\varepsilon)}\right)\right) \leqq$$

$$\leqq \frac{\left(\mathsf{E}\left[\exp\left(n^{\frac{1-\varepsilon}{3}} \frac{I_{(\eta_i=j)} - p_j}{\sqrt{np_j}}\right)\right]\right)^n}{\exp\left(n^{\frac{2}{3}(1-\varepsilon)}\right)} \leqq \frac{\left[1 + \frac{2}{3} n^{\frac{2}{3}(1-\varepsilon)} \mathsf{E}\left(\frac{I_{(\eta_1=j)} - p_j}{\sqrt{np_j}}\right)^2\right]^n}{\exp\left(n^{\frac{2}{3}(1-\varepsilon)}\right)} \leqq$$

$$\leqq \frac{\exp\left[\frac{2}{3}(1-p_j) n^{\frac{2}{3}(1-\varepsilon)}\right]}{\exp\left(n^{\frac{2}{3}(1-\varepsilon)}\right)} \leqq \exp\left(-\frac{n^{\frac{2}{3}(1-\varepsilon)}}{3}\right)$$

and similarly

$$\mathsf{P}\left(v_j - np_j < -\sqrt{np_j}\, n^{\frac{1-\varepsilon}{3}}\right) = \exp\left(-\frac{n^{\frac{2}{3}(1-\varepsilon)}}{3}\right).$$

Thus (5) is valid.

Now, by lemma 1 and by the relation $\frac{1}{\sqrt{2\pi}\,x} e^{-x^2/2} \sim 1-\Phi(x)$ for $x \to \infty$ we have

$$\mathsf{P}\left(\frac{\sqrt{np_j}}{\sigma_j}\cdot\frac{s_j}{v_j} > \sqrt{2\log a_n - \log\log a_n + y}\,\middle|\, v_j = t\right) =$$

$$= \mathsf{P}\left(\frac{\tilde{\xi}_{1j}+\ldots+\tilde{\xi}_{tj}}{\sigma_j\sqrt{t}} > \sqrt{\frac{t}{np_j}}\sqrt{2\log a_n - \log\log a_n + y}\right) \sim$$

$$\sim \sqrt{\frac{np_j}{2\pi t}}\cdot\frac{1}{\sqrt{2\log a_n - \log\log a_n + y}}\cdot e^{-\frac{np_j}{2t}(2\log a_n - \log\log a_n + y)} \sim \frac{e^{-y/2}}{2\sqrt{\pi}\,a_n}$$

if

$$|t-np_j| < \sqrt{np_j}\, n^{\frac{1-\varepsilon}{3}}$$

On the other hand from (5) we have

$$\mathsf{P}\left(|v_k - np_k| > \sqrt{np_k}\, n^{\frac{1-\varepsilon}{3}} \text{ for some } k\right) \leqq 2a_n \exp\left(-\frac{n^{\frac{2}{3}(1-\varepsilon)}}{3}\right) = o(1)$$

Thus

$$\mathsf{P}\left(\max_k \frac{\sqrt{np_k}}{\sigma_k}\cdot\frac{s_k}{v_k} < \sqrt{2\log a_n - \log\log a_n + y}\right) =$$

$$= \left[1-\frac{e^{-y/2}}{2\sqrt{\pi}\,a_n}(1+o(1))\right]^{a_n} \mathsf{P}\left(|v_k - np_k| < \sqrt{np_k}\, n^{\frac{1-\varepsilon}{3}} \text{ for every } k\right) + o(1) =$$

$$= \exp\left(-\frac{e^{-y/2}}{2\sqrt{\pi}}\right) + o(1)$$

as we have stated.

The second relation of lemma 2 can be proved in the same way.

LEMMA 3. *Under the conditions of lemma* 2

$$\mathsf{P}\left(\max_k \frac{\sqrt{np_k^{(n)}}}{\sigma_k^{(n)}}\cdot\frac{\sum_{i=1}^{n}\xi_i^{(n)} I_{(\eta_i^{(n)}=k)}}{\sum_{i=1}^{n} I_{(\eta_i^{(n)}=k)}} < \sqrt{2\log a_n - 3\log\log a_n - 2\log 2\sqrt{\pi}\,c}\right) = \frac{1}{a_n^{c+o(1)}}$$

$$\mathsf{P}\left(\max_k \frac{\sqrt{np_k^{(n)}}}{\sigma_k^{(n)}}\cdot\left|\frac{\sum_{i=1}^{n}\xi_i^{(n)} I_{(\eta_i^{(n)}=k)}}{\sum_{i=1}^{n} I_{(\eta_i^{(n)}=k)}}\right| < \sqrt{2\log a_n - 3\log\log a_n - 2\log\sqrt{\pi}\,c}\right) = \frac{1}{a_n^{c+o(1)}}$$

The proof is similar to that of Lemma 2.

LEMMA 4. *Under the conditions of Lemma 2 we have*

$$\mathsf{P}\left(\max_k \frac{\sqrt{np_k^{(n)}}}{\sigma_k^{(n)}}\left(1+o\left(\frac{1}{\log a_n}\right)\right)\cdot \frac{\sum_{i=k}^{n} I_{(\eta_i^{(n)}=k)}\left[\xi_i^{(n)}+o\left(\sqrt{\frac{a_n}{n\log a_n}}\right)\right]}{\sum_{i=1}^{n} I_{(\eta_i^{(n)}=k)}} < \right.$$

$$\left. < \sqrt{2\log a_n - \log\log a_n + y}\right) \to \exp\left(-\frac{e^{-y/2}}{2\sqrt{\pi}}\right)$$

$$\mathsf{P}\left(\max_k \frac{\sqrt{np_k^{(n)}}}{\sigma_k^{(n)}}\left(1+o\left(\frac{1}{\log a_n}\right)\right)\cdot \left|\frac{\sum_{i=1}^{n} I_{(\eta_i^{(n)}=k)}\left[\xi_i^{(n)}+o\left(\sqrt{\frac{a_n}{n\log a_n}}\right)\right]}{\sum_{i=1}^{n} I_{(\eta_i^{(n)}=k)}}\right| < \right.$$

$$\left. < \sqrt{2\log a_n - \log\log a_n + y}\right) \to \exp\left(-\frac{e^{-y/2}}{\sqrt{\pi}}\right).$$

Lemma 4 is an easy consequence of Lemma 2.

PROOF OF THEOREM 1. Let us divide the interval [0,1] into a_n equal parts. Obviously

$$m_k^{(n)} = \mathsf{E}\left(\xi_i \,\middle|\, \frac{k-1}{a_n} \leqq \eta_i < \frac{k}{a_n}\right) = \frac{\int_{\frac{k-1}{a_n}}^{\frac{k}{a_n}} R(y)f(y)\,dy}{\int_{\frac{k-1}{a_n}}^{\frac{k}{a_n}} f(y)\,dy}$$

and

$$D_k^{2(n)} = D^2\left(\xi_i \,\middle|\, \frac{k-1}{a_n} \leqq \eta_i < \frac{k}{a_n}\right) = \frac{\int_{\frac{k-1}{a_n}}^{\frac{k}{a_n}} [\sigma^2(y)+R^2(y)]f(y)\,dy}{\int_{\frac{k-1}{a_n}}^{\frac{k}{a_n}} f(y)\,dy} - [m_k^{(n)}]^2.$$

Let us define the random variables $(\xi_i^{(n)}, \eta_i^{(n)})$ $i=1, 2 \ldots n$ in the following way:

$$\eta_i^{(n)} = k \quad \text{and} \quad \xi_i^{(n)} = \frac{\xi_i - m_k^{(n)}}{D_k^{(n)}} \cdot \sigma\left(\frac{k-\frac{1}{2}}{a_n}\right)$$

if

$$\frac{k-1}{a_n} \leqq \eta_i < \frac{k}{a_n} \qquad i = 1, 2, \ldots, n \quad k = 1, 2, \ldots, a_n.$$

Then the conditions of Lemma 2 are satisfied for the sequence $(\xi_i^{(n)}, \eta_i^{(n)})$. Indeed, $p_k^{(n)} = \int_{\frac{k-1}{a_n}}^{\frac{k}{a_n}} f(y)\,dy$, thus $\frac{c_1}{a_n} < p_k^{(n)} < \frac{c_2}{a_n}$ since $f(x)$ is bounded both from above and below. $\mathsf{E}(\xi_i^{(n)} | \eta_i^{(n)} = k) = 0$, $(\sigma_k^{(n)})^2 = \mathsf{D}^2(\xi_i^{(n)} | \eta_i^{(n)} = k) = \sigma^2\left(\frac{k-\frac{1}{2}}{a_n}\right)$ is bounded from below. $\frac{D_k^{(n)}}{\sigma\left(\frac{k-\frac{1}{2}}{a_n}\right)}$ is bounded from below and

$$\mathsf{E}\left(\exp\left(t_0 \frac{D_k^{(n)}}{\sigma\left(\frac{k-\frac{1}{2}}{a_n}\right)} |\xi_i^{(n)}|\right) \middle| \eta_i^{(n)} = k\right) \leqq c \exp\left(t_0 m_k^{(n)}\right)$$

is also bounded as required. By the Taylor formula

$$f(y) = f(x) + (y-x) \cdot f'\big(x + \theta(y-x)\big)$$

$$R(y) = R(x) + (y-x) \cdot R'\big(x + \theta'(y-x)\big) \qquad 0 < \theta, \theta', \theta'' < 1$$

$$\sigma(y) = \sigma(x) + (y-x) \cdot \sigma'\big(x + \theta''(y-x)\big)$$

The first derivatives of $f(x)$, $R(x)$, $\sigma(x)$ are bounded and so it can be seen easily that $m_k^{(n)} = R(x) + O\left(\frac{1}{a_n}\right)$, $p_k^{(n)} = \frac{f(x)}{a_n} + O\left(\frac{1}{a_n^2}\right)$, $\mathsf{D}^2(\xi_i^{(n)} | \eta_i^{(n)} = k) = \sigma^2(x) + O\left(\frac{1}{a_n}\right)$ if $x \in \left[\frac{k-1}{a_n}, \frac{k}{a_n}\right]$. Since $n^\alpha < a_n$ and $\alpha > \frac{1}{3}$, $\frac{1}{a_n} = o\left(\sqrt{\frac{a_n}{n \log a_n}}\right)$ and

$$\sqrt{\frac{nf(x)}{a_n}} \frac{1}{\sigma(x)} = \frac{\sqrt{np_k^{(n)}}}{\sigma_k^{(n)}} \left(1 + o\left(\frac{1}{\log a_n}\right)\right) \quad \text{for} \quad x \in \left[\frac{k-1}{a_n}, \frac{k}{a_n}\right].$$

Hence Lemma 4 can be applied, and we get

$$\mathsf{P}\left(\sup_k \sup_{k \in \left[\frac{k-1}{a_n}, \frac{k}{a_n}\right]} \sqrt{\frac{nf(x)}{a_n}} \frac{\sum_{i=1}^n I_{(\eta_i^{(n)}=k)} [\xi_i - R(x)]}{\sigma(x) \sum_{i=1}^n I_{(\eta_i^{(n)}=k)}} < \sqrt{2 \log a_n - \log \log a_n + y}\right) \to$$

$$\to \exp\left(-\frac{e^{-y/2}}{2\sqrt{\pi}}\right)$$

which is identical to formula (3). The second statement of lemma 4 gives formula (4).

The PROOF of Theorem 1′ is almost the same. By the Taylor expansion up to 2 terms

$$R(x) = R\left(\frac{k-\frac{1}{2}}{a_n}\right) + R'\left(\frac{k-\frac{1}{2}}{a_n}\right)\left(x - \frac{k-\frac{1}{2}}{a_n}\right) +$$

$$+ \frac{1}{2} R''\left(\frac{k-\frac{1}{2}}{a_n} + \theta\left(x - \frac{k-\frac{1}{2}}{a_n}\right)\right)\left(x - \frac{k-\frac{1}{2}}{a_n}\right)^2$$

$$f(x) = f\left(\frac{k-\frac{1}{2}}{a_n}\right) + f'\left(\frac{k-\frac{1}{2}}{a_n}\right)\left(x - \frac{k-\frac{1}{2}}{a_n}\right) +$$

$$+ \frac{1}{2} f''\left(\frac{k-\frac{1}{2}}{a_n} + \theta'\left(x - \frac{k-\frac{1}{2}}{a_n}\right)\right)\left(x - \frac{k-\frac{1}{2}}{a_n}\right)^2$$

$$\sigma(x) = \sigma\left(\frac{k-\frac{1}{2}}{a_n}\right) + \sigma'\left(\frac{k-\frac{1}{2}}{a_n}\right)\left(x - \frac{k-\frac{1}{2}}{a_n}\right) +$$

$$+ \frac{1}{2} \sigma''\left(\frac{k-\frac{1}{2}}{a_n} + \theta''\left(x - \frac{k-\frac{1}{2}}{a_n}\right)\right)\left(x - \frac{k-\frac{1}{2}}{a_n}\right)^2$$

where $0 < \theta, \theta', \theta'' < 1$. It implies that

$$m_k^{(n)} = R\left(\frac{k-\frac{1}{2}}{a_n}\right) + O\left(\frac{1}{a_n^2}\right), \quad p_k^{(n)} = \frac{f\left(\frac{k-\frac{1}{2}}{a_n}\right)}{a_n} + O\left(\frac{1}{a_n^3}\right)$$

and

$$D^2(\xi_i^{(n)} | \eta_i^{(n)} = k) = \sigma^2\left(\frac{k-\frac{1}{2}}{a_n}\right) + O\left(\frac{1}{a_n^2}\right),$$

and these imply, just as in Theorem 1, the limit relation

$$\mathsf{P}\left(\sup_k \sqrt{\frac{nf\left(\frac{k-\frac{1}{2}}{a_n}\right)}{a_n}} \frac{R_n\left(\frac{k-\frac{1}{2}}{a_n}\right) - R\left(\frac{k-\frac{1}{2}}{a_n}\right)}{\sigma\left(\frac{k-\frac{1}{2}}{a_n}\right)} < \sqrt{2\log a_n - \log\log a_n + y}\right) \to$$

$$\to \exp\left(\frac{-e^{-y/2}}{2\sqrt{\pi}}\right).$$

Now by

$$R(x) = R\left(\frac{k-\frac{1}{2}}{a_n}\right) + \left[R\left(\frac{k+\frac{1}{2}}{a_n}\right) - R\left(\frac{k-\frac{1}{2}}{a_n}\right)\right]\left(a_n x - k - \frac{1}{2}\right) + O\left(\frac{1}{a_n^2}\right)$$

for $x \in \left[\frac{k-\frac{1}{2}}{a_n}, \frac{k+\frac{1}{2}}{a_n}\right]$ and (4) the statement of Theorem 1′ is valid.

PROOF OF THEOREM 2. Since we have already proved Theorem 1, it is sufficient to prove that

$$\mathsf{P}\left(\sup_x \frac{\sqrt{f(x)}}{\sigma(x)} \frac{\sigma_n(x)}{\sqrt{f_n(x)}} > 1 + \frac{\delta_n}{\sqrt{\log a_n}}\right) \to 0$$

and

$$\mathsf{P}\left(\inf_x \frac{\sqrt{f(x)}}{\sigma(x)} \frac{\sigma_n(x)}{\sqrt{f_n(x)}} > 1 - \frac{\delta_n}{\sqrt{\log a_n}}\right) \to 0$$

for an appropriate sequence $\delta_n \to 0$, i.e. $\frac{\sqrt{f(x)}}{\sigma(x)}$ and $\frac{\sqrt{f_n(x)}}{\sigma_n(x)}$ are near to each other. We shall estimate $f_n(x) - f(x)$ and $\sigma_n(x) - \sigma(x)$. This will be similar to the proof of Lemma 2.

$$f(x) - a_n p_k^{(n)} = O\left(\frac{1}{a_n}\right) \quad x \in \left[\frac{k+1}{a_n}, \frac{k}{a_n}\right]$$

and

$$f_n(x) - a_n p_k^{(n)} = \frac{a_n}{n} \sum_{i=1}^{n} [I_{(\eta_i^{(n)}=k)} - \mathsf{E} I_{(\eta_i^{(n)}=k)}].$$

Thus, exactly the same way as in proving estimation (5) we get that

$$\mathsf{P}\left(|f_n(x) - f(x)| > n^{\frac{1-\alpha}{3}} \frac{\sqrt{n}}{a_n \sqrt{p_k^{(n)}}}\right) = O\left(\exp\left(-n^{\frac{1-\alpha}{3}}\right)\right).$$

$n^{\frac{1-\alpha}{3}} \frac{\sqrt{n}}{a_n \sqrt{p_k^{(n)}}} < K \cdot n^{-\frac{1-\alpha}{6}}$ for an appropriate constant K. Thus

$$\mathsf{P}\left(\sup_x |f_n(x) - f(x)| > K \cdot n^{-\frac{1-\alpha}{6}}\right) \leqq a_n O\left(\exp\left(-n^{\frac{1-\alpha}{3}}\right) \to 0\right.$$

We have

$$\sigma^2(x) - D_k^{(n)})^2 = O\left(\frac{1}{a_n}\right) \quad \text{and} \quad \sigma_n^2(x) = \frac{\sum_{i=1}^{n} \xi_i^2 I_{(\eta_i^{(n)}=k)}}{\sum_{i=1}^{n} I_{(\eta_i^{(n)}=k)}} - \left[\frac{\sum_{i=1}^{n} \xi_i I_{(\eta_i^{(n)}=k)}}{\sum_{i=1}^{n} I_{(\eta_i^{(n)}=k)}}\right]^2$$

for

$$x \in \left[\frac{k-1}{a_n}, \frac{k}{a_n}\right].$$

Now we estimate

$$\mathsf{P}\left(\left|\frac{\sum_{i=1}^{n} \xi_i I_{(\eta_i^{(n)}=k)}}{\sum_{i=1}^{n} I_{(\eta_i^{(n)}=k)}} - m_k^{(n)}\right| > \frac{n^{\frac{1-\alpha}{3}}}{\sqrt{n p_k^{(n)}}}\right).$$

Let $\bar{\xi}_1, \dots \bar{\xi}_n \dots n=1, 2, \dots$ be independent identically distributed random variables having the distribution

$$\mathsf{P}(\bar{\xi}_i < x) = \mathsf{P}(\xi_1 - m_k^{(n)} < x | \eta_1^{(n)} = k)$$

Then

$$\mathsf{P}\left(\left| \frac{\sum_{i=1}^{n} \xi_i I_{(\eta_i^{(n)}=k)}}{\sum_{i=1}^{n} I_{(\eta_i^{(n)}=k)}} - m_k^{(n)} \right| > \frac{n^{\frac{1-\alpha}{3}}}{\sqrt{np_k^{(n)}}} \,\middle|\, \sum_{i=1}^{n} I_{(\eta_i^{(n)}=k)} = t \right) =$$

$$= \mathsf{P}\left(\left| \frac{\bar{\xi}_1 + \dots + \bar{\xi}_t}{\sqrt{t}} \right| < n^{\frac{1-\alpha}{3}} \sqrt{\frac{t}{np_k^{(n)}}} \right) = O\left(\exp\left(-n^{\frac{1-\alpha}{3}}\right)\right)$$

for $|t - np_k^{(n)}| < n^{\frac{1-\alpha}{3}} \sqrt{np_k^{(n)}}$. On the other hand

$$\mathsf{P}\left(\left| \sum_{i=1}^{n} I_{(\eta_i^{(n)}=k)} - np_k^{(n)} \right| > n^{\frac{1-\alpha}{3}} \sqrt{np_k^{(n)}} \right) = O\left(\exp\left(-n^{\frac{1-\alpha}{3}}\right)\right)$$

Thus

$$\mathsf{P}\left(\left| \frac{\sum_{i=1}^{n} \xi_i I_{(\eta_i^{(n)}=k)}}{\sum_{i=1}^{n} I_{(\eta_i^{(n)}=k)}} - m_k^{(n)} \right| > \frac{n^{\frac{1-\alpha}{3}}}{\sqrt{np_k^{(n)}}} \text{ for some } k \right) \leqq a_n O\left(\exp\left(-n^{\frac{1-\alpha}{3}}\right)\right) \to 0$$

The relation

$$\mathsf{P}\left(\left| \frac{\sum_{i=1}^{n} \xi_i^2 I_{(\eta_i^{(n)}=k)}}{\sum_{i=1}^{n} I_{(\eta_i^{(n)}=k)}} - \frac{\int_{\frac{k-1}{a_n}}^{\frac{k}{a_n}} [\sigma^2(y) + R^2(y)] f(y)\, dy}{\int_{\frac{k-1}{a_n}}^{\frac{k}{a_n}} f(y)\, dy} \right| > \frac{n^{\frac{1-\alpha}{3}}}{\sqrt{np_k^{(n)}}} \text{ for some } k \right) \to 0$$

can be proved in the same way. From these estimations it follows that $|f_n(x) - f(x)|$ and $|\sigma_n(x) - \sigma(x)|$ are less than n^{-c} for appropriate $c > 0$ with a probability near 1. Thus Theorem 2 is valid.

PROOF OF THEOREM 3. First we prove that

$$\mathsf{P}\left(\underline{\lim}\, \sup_x \sqrt{\frac{nf(x)}{2a_n \log a_n}} \frac{|R_n(x) - R(x)|}{\sigma(x)} > 1 - \varepsilon \right) = 1$$

for arbitrary $\varepsilon > 0$.

It is easy to check the validity of the conditions of Lemma 2 for the sequence $(\xi_i^{(n)}, \eta_i^{(n)})$, therefore by Lemma 3 we can state that

$$\mathsf{P}\left(\max_k \sqrt{np_k^{(n)}} \left| \frac{\sum_{i=1}^n \xi_i^{(n)} I_{(\eta_i^{(n)}=k)}}{\sigma\left(\frac{k-\frac{1}{2}}{a_n}\right)\sum_{i=1}^n I_{(\eta_i^{(n)}=k)}} \right| > \sqrt{2\log a_n - 3\log\log a_n - 2\log\sqrt{\pi}\, c}\right) =$$

$$= \frac{1}{a_n^{c+o(1)}}$$

Thus, choosing n_0 large enough, by the Borel—Cantelly lemma we obtain

$$\max_k \sqrt{np_k^{(n)}} \left| \frac{\sum_{i=1}^n \xi_i^{(n)} I_{(\eta_i^{(n)}=k)}}{\sigma\left(\frac{k-\frac{1}{2}}{a_n}\right)\sum_{i=1}^n I_{(\eta_i^{(n)}=k)}} \right| <$$

$$< \sqrt{2\log a_n - 3\log\log a_n - 2\log\sqrt{\pi}\, c} \quad \text{for} \quad n > n_0(\omega)$$

with probability 1. Since

$$\sqrt{np_k^{(n)}} = \sqrt{\frac{nf(x)}{a_n}}\left(1+O\left(\frac{1}{a_n}\right)\right), \quad \sigma\left(\frac{k-\frac{1}{2}}{a_n}\right) = \sigma(x)+O\left(\frac{1}{a_n}\right)$$

and

$$\xi_i^{(n)} I_{(\eta_i^{(n)}=k)} = \left[\xi_i - R(x) + O\left(\frac{1}{a_n}\right)\right] I_{(\eta_i^{(n)}=k)}\left[1+O\left(\frac{1}{a_n}\right)\right]$$

if $x \in \left[\frac{k-1}{a_n}, \frac{k}{a_n}\right]$, our statement follows.

The proof of the relation

$$\mathsf{P}\left(\overline{\lim}\sup_x \sqrt{\frac{nf(x)}{2a_n\log a_n}} \frac{|R_n(x)-R(x)|}{\sigma(x)} < 1-\varepsilon\right) = 1$$

is more involved.

Set

$$\bar{\xi}_i = \frac{\xi_i - \mathsf{E}(\xi_i|\eta_i)}{\mathsf{D}(\xi_i|\eta_i)}.$$

Let us define the random variable

$$T_n(x) = \sum_{\substack{i \leqq n \\ \eta_i \leqq x}} \bar{\xi}_i \qquad 0 \leqq x \leqq 1.$$

We want to prove that

(6)
$$\mathsf{P}(A_n) = \mathsf{P}\left(\sup_{\substack{y-x \leqq \frac{1}{a_n} \\ 0 \leqq x < y \leqq 1}} \frac{1}{\sqrt{f(x)}} |T_n(x) - T_n(y)| > \left(1+\varepsilon+\frac{\sqrt{\varepsilon}}{2}\right) \sqrt{\frac{2n \log a_n}{a_n}} \right) \leqq \frac{12}{\varepsilon \cdot a_n^{\varepsilon}}.$$

First we estimate, for fixed x, the probability

$$\mathsf{P}\left(\sup_{x \leqq y \leqq x+\frac{3+\varepsilon}{3a_n}} \frac{1}{\sqrt{f(x)}} |T_n(x) - T_n(y)| > (1+\varepsilon) \sqrt{\frac{2n \log a_n}{a_n}} \right).$$

Let us order the η_i-s falling into the interval $\left[x, x+\frac{3+\varepsilon}{3a_n}\right]$ in increasing order $x \leqq \eta_1^* < \eta_2^* \ldots < \eta_t^* < x+\frac{3+\varepsilon}{3a_n}$ and denote by $\bar{\xi}_i^*$ the $\bar{\xi}_i$ which is the pair of η_j^*. Then

$$\sup_{x<y<x+\frac{3+\varepsilon}{3a_n}} |T_n(y) - T_n(x)| = \sup_{k \leqq t} \left| \sum_{j=1}^{k} \bar{\xi}_j^* \right|$$

We can state that

$$(7) \quad \mathsf{P}\left(\sup_{k \leqq t} \sum_{j=1}^{k} \bar{\xi}_j^* > (1+\varepsilon) \sqrt{\frac{2nf(x) \log a_n}{a_n}} \,\middle|\, \begin{array}{l} \text{the number of } \eta_i\text{-s } i \leqq n \text{ falling} \\ \text{into the interval } \left[x, x+\frac{3+\varepsilon}{3a_n}\right] \\ \text{is } t,\ \eta_1^* = x_1, \ldots \eta_t^* = x_t \end{array} \right) =$$

$$= \mathsf{P}\left(\sup_{k \leqq t} \sum_{j=1}^{k} \zeta_j > (1+\varepsilon) \sqrt{\frac{2nf(x) \log a_n}{a_n}} \right)$$

where $\zeta_1, \zeta_2, \ldots, \zeta_t$ are independent and $\mathsf{P}(\zeta_j < y) = \mathsf{P}(\xi_1 < y \mid \eta_1 = x_j)$. Now the sequence $\sum_{j=1}^{k} \zeta_j$, $k = 1, 2 \ldots, t$ forms a martingale, so $\exp\left(t \sum_{j=1}^{k} \zeta_j\right)$ is a semimartingale for arbitrary t, and by a well-known martingale inequality see [1] p. 314.,

$$(8) \quad \mathsf{P}\left(\sup_{k \leqq t} \sum_{j=1}^{k} \zeta_j > (1+\varepsilon) \sqrt{\frac{2nf(x) \log a_n}{a_n}} \right) =$$

$$= \mathsf{P}\left(\sup_{k \leqq t} \exp\left(\sqrt{\frac{2a_n \log a_n}{nf(x)}} \sum_{j=1}^{k} \zeta_j \right) > \exp\left(2(1+\varepsilon) \log a_n\right) \right) \leqq$$

$$\leqq \frac{\prod_{i=1}^{t} \mathsf{E} \exp\left(\sqrt{\frac{2a_n \log a_n}{nf(x)}} \zeta_i \right)}{a_n^{2(1+\varepsilon)}}$$

Now

$$\mathsf{E}\exp\left(\sqrt{\frac{2a_n\log a_n}{nf(x)}}\zeta_i\right) \leqq 1+\mathsf{E}\sqrt{\frac{2a_n\log a_n}{nf(x)}}\zeta_i + \tag{9}$$

$$+\mathsf{E}\left[\frac{a_n\log a_n}{nf(x)}\zeta_i^2\exp\left|\sqrt{\frac{2a_n\log a_n}{nf(x)}}\zeta_i\right|\right] \leqq 1+\left(1+\frac{\varepsilon}{2}\right)\frac{a_n\log a_n}{nf(x)} \leqq$$

$$\leqq \exp\left[\left(1+\frac{\varepsilon}{2}\right)\frac{a_n\log a_n}{nf(x)}\right]$$

It can be proved in the same way as relation (5) that

$$\mathsf{P}\left(\left|t-\frac{n\left(1+\frac{\varepsilon}{3}\right)}{a_n}\cdot f(x)\right| > n^{\frac{1-\alpha}{3}}\sqrt{\frac{n}{a_n}f(x)}\right) = O\left(\exp\left(-n^{\frac{1}{3}(1-\alpha)}\right)\right)$$

(t is the number of η_i-s ($i\leqq n$) falling into the interval $\left[x, x+\frac{3+\varepsilon}{3a_n}\right]$), and so by (7), (8) and (9)

$$\mathsf{P}\left(\sup_{x<y<x+\frac{3+\varepsilon}{3a_n}}\left(T_n(y)-T_n(x)\right) > (1+\varepsilon)\sqrt{\frac{2nf(x)\log a_n}{a_n}}\right) < \frac{1}{a_n^{1+\varepsilon}}$$

Using this relation for the sequence $(-\xi_1, \eta_1), \ldots, (-\xi_n, \eta_n)$ we get that

$$\mathsf{P}\left(\sup_{x<y<x+\frac{3+\varepsilon}{3a_n}}|T_n(x)-T_n(y)| > (1+\varepsilon)\sqrt{\frac{2nf(x)\log a_n}{a_n}}\right) < \frac{2}{a_n^{1+\varepsilon}}$$

In the same way (x is fixed) we find

$$\mathsf{P}\left(\sup_{x<y<x+\frac{\varepsilon}{3a_n}}|T_n(y)-T_n(x)| > \frac{\sqrt{\varepsilon}}{2}\sqrt{\frac{2nf(x)\log a_n}{a_n}}\right) < \frac{2}{a_n^{1+\varepsilon}}$$

Thus

$$\mathsf{P}\left(\sup_{\substack{x<y<x+\frac{3+\varepsilon}{3a_n}\\ x=\frac{k\varepsilon}{3a_n}\left(k=0.1,\ldots,\frac{3a_n}{\varepsilon}\right)}}\frac{1}{\sqrt{f(x)}}|T_n(y)-T_n(x)| > (1+\varepsilon)\sqrt{\frac{2n\log a_n}{a_n}}\right) < \frac{6}{\varepsilon a_n^{\varepsilon}}$$

and

$$\mathsf{P}\left(\sup_{\substack{x<y<x+\frac{\varepsilon}{3a_n}\\ x=\frac{k\varepsilon}{3a_n}\left(k=0,1,\ldots,\frac{3a_n}{\varepsilon}\right).}}\frac{1}{\sqrt{f(x)}}|T_n(y)-T_n(x)| > \frac{\sqrt{\varepsilon}}{2}\sqrt{\frac{2n\log a_n}{a_n}}\right) < \frac{6}{\varepsilon a_n^{\varepsilon}}$$

Let us now consider an arbitrary interval $[u, v]$ such that $0<v-u \leqq \frac{1}{a_n}$. Choose the integer k in such a way that for the number $x = \frac{k\varepsilon}{3a_n}$ the inequality $0<u-x<\frac{\varepsilon}{3a_n}$ hold. Then

$$|T_n(v)-T_n(u)| \leqq \sup_{x<y<x+\frac{3+\varepsilon}{3a_n}} |T_n(y)-T_n(x)| + \sup_{x<y<x+\frac{\varepsilon}{3a_n}} |T_n(y)-T_n(x)|,$$

and therefore (6) follows from the last two inequalities. Choosing a $c>\frac{6}{\varepsilon}$ we get $\mathsf{P}(A_{l^c})<\frac{1}{l^2}$, and by the Borel—Cantelli lemma

$$\lim_{l\to\infty} A_{l^c} = \emptyset \tag{10}$$

Given a number n it determines a unique l for which $l^c \leqq n < (l+1)^c$. Then $n-l^c < c(l+1)^{c-1} < 2cn^{c-1}$. By similar calculations as before we get

$$\mathsf{P}\left(\sum_{i=l^c}^{n} \bar{\xi}_i I_{(\eta_i^{(n)}=k)} > \varepsilon\sqrt{\frac{2nf(x)\log a_n}{a_n}}\right) < \exp(-Dn^{1/c})$$

with appropriate $D>0$ for $k=1, 2, \ldots, a_n$, $x\in\left[\frac{k-1}{a_n}, \frac{k}{a_n}\right]$ and

$$\mathsf{P}(B_n) = \mathsf{P}\left(\sup_{\substack{k \\ x\in\left[\frac{k-1}{a_n}, \frac{k}{a_n}\right]}} \left|\frac{1}{\sqrt{f(x)}} \sum_{i=l^c}^{n} \bar{\xi}_i I_{(\eta_i^{(n)}=k)}\right| > \varepsilon\sqrt{\frac{2n\log a_n}{a_n}}\right) < a_n \exp(-Dn^{1/c})$$

This yields the relation

$$\lim B_n = \emptyset. \tag{11}$$

By the definition of A_n and B_n and relations (10) and (11) we have with probability 1

$$\left|\sup \frac{1}{\sqrt{f(x_k)}} \sum_{i=1}^{n} \bar{\xi}_i I_{(\eta_i^{(n)}=k)}\right| < \left(1+2\varepsilon+\frac{\sqrt{\varepsilon}}{2}\right)\sqrt{\frac{2n\log a_n}{a_n}} \tag{12}$$

for

$$x_k\in\left[\frac{k-1}{a_n}, \frac{k}{a_n}\right]$$

for every large n.

An easy computation shows that with probability one

$$\sum_{i=1}^{n} I_{(\eta_i^{(n)}=k)} > (1-\varepsilon)\frac{nf(x_k)}{a_n}, \; x_k\in\left[\frac{k-1}{a_n}, \frac{k}{a_n}\right] \tag{13}$$

for every large n and $k=1, 2, \ldots, a_n$. Since $R(x)$, $\sigma(x)$ and $f(x)$ are smooth enough

$$P\left(\overline{\lim_n} \sup_x \sqrt{\frac{nf(x)}{2a_n \log a_n}} \frac{|R_n(x)-R(x)|}{\sigma(x)} < \frac{1+2\varepsilon+\frac{\sqrt{\varepsilon}}{2}}{1-\varepsilon}\right) = 1$$

from the relations (12) and (13). Since it is true for arbitrary small $\varepsilon>0$, the first statement of theorem 3 is valid. The second statement can be proved in the same way.

LITERATURE

[1] Doob, J. L.: *Stochastic Processes,* Wiley, New York (1953).
[2] Feller, W.: *An Introduction to Probability Theory and Its Applications* 2. Wiley, New York.
[3] Révész, P.: On empirical density function, *Periodica Math. Hungar.* **2** (1972) 85—110.
[4] Надарая, Э. А.: Замечания о непаратмерических оценках плотности вероятности и кривой регрессии. *Теория вероят. и ее примен.* XV. 1 (1970) 139—142.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received April 15, 1973)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 363—367.

EINE FEHLERABSCHÄTZUNG ZUR INTERPOLATION STETIGER FUNKTIONEN

von

H. BRASS und R. GÜNTTNER

§ 1. Das Resultat

Es sei f eine reelle 2π-periodische stetige Funktion. $T[f]$ bedeute das trigonometrische Interpolationspolynom von f zu m äquidistanten Stützstellen, $m = 2n+1$ oder $m=2n$, d.h. denjenigen trigonometrischen Ausdruck der Ordnung n, für den

$$T[f](x_\nu) = f(x_\nu) \qquad \nu = 1, 2, \ldots, m$$

$$x_\nu = \nu \frac{2\pi}{m}$$

gilt, wobei im Falle $m=2n$ (wie üblich) Eindeutigkeit durch die zusätzliche Forderung erzwungen wird, daß kein Glied mit $\sin nx$ vorkommt.

In Verschärfung älterer mit approximationstheoretischen Methoden gewonnener Ergebnisse hat KIS [3] die folgende Abschätzung bewiesen:

$$\|T[f]-f\| \leqq \frac{1}{2}\left(\|T\|+3+\frac{8}{\pi}\right)\omega\left(f; \frac{2\pi}{m}\right)$$

Hier wie im folgenden bedeutet $\omega(f;\cdot)$ den Stetigkeitsmodul von f und $\|\cdot\|$ ist im Sinne der sup-Norm zu verstehen. Wir wollen hier die folgende, von KIS vermutete, weitere Verschärfung beweisen:

SATZ 1. *Es gilt*

$$\|T[f]-f\| \leqq \frac{1}{2}(\|T\|+1)\,\omega\left(f; \frac{2\pi}{m}\right).$$

Daß der Faktor $\frac{1}{2}(\|T\|+1)$ nicht noch weiter verkleinert werden kann, hat KIS schon bewiesen; wir bemerken hier nur, daß diese Tatsache leicht aus der Beziehung

$$\|T-I\| = \|T\|+1$$

(I bedeutet die identische Abbildung) folgt, die für eine recht allgemeine Klasse von Operatoren T gilt, insbesondere für die hier in Rede stehenden Interpolationsoperatoren (vgl. z.B. CHENEY/PRICE [1]).

Der elementare Beweis von Satz 1, den wir in § 2 ausführen, läßt sich nach einigen Modifikationen auf die Interpolation durch algebraische Polynome übertragen. Bezeichnen wir mit g eine auf dem Intervall $I=[x_0, x_{m+1}]$ stetige, reellwertige Funktion und mit $P[g]$ ihr algebraisches Interpolationspolynom zu den Stützstellen $x_1<x_2<\ldots<x_m$ (mit $x_\nu \in [x_0, x_{m+1}]$), so hat man

SATZ 2. *Es gilt*

$$\|P[g]-g\| \leqq \frac{1}{2}(\|P\|+1)\omega(g;\Delta)$$

mit

$$\Delta = \max_{0\leqq i\leqq m} |x_{i+1}-x_i|$$

Wie in Satz 1 ist auch hier der Faktor $\frac{1}{2}(\|P\|+1)$ unverbesserbar. Als Spezialfälle von Satz 2 heben wir hervor:

FOLGERUNG 1. *Es sei g eine auf* $[-1, 1]$ *stetige Funktion und* $P[g]$ *ihr Interpolationspolynom zu den m Stützstellen* $(m>1)$

$$x_\nu = \cos\frac{2\nu-1}{2m}\pi \qquad \nu = 1, 2, \ldots, m.$$

Dann gilt
für m ungerade:

$$\|P[g]-g\| \leqq \frac{1}{2}(\|P\|+1)\,\omega\left(g;\ \sin\frac{\pi}{m}\right),$$

für m gerade:

$$\|P[g]-g\| \leqq \frac{1}{2}(\|P\|+1)\,\omega\left(g;\ 2\sin\frac{\pi}{2m}\right).$$

Dies ist die Verschärfung eines Resultates, das KIS [3] durch die übliche Transformation aus der oben erwähnten Aussage für die trigonometrische Interpolation gewinnt. Auch die übrigen von KIS angegebenen Abschätzungen zur Polynominterpolation ergeben sich (in verschärfter Form) als Spezialfälle von Satz 2.

Die in der nächsten Folgerung genannte Knotenwahl ist von Interesse, weil für sie $\|P\|$ besonders klein wird (dazu und zu den numerischen Werten von $\|P\|$ und $\|T\|$ siehe GÜNTTNER [2]):

FOLGERUNG 2. *Es sei g eine auf* $[-1, 1]$ *stetige Funktion und* $P[g]$ *ihr Interpolationspolynom zu den m Stützstellen* $(m>1)$

$$x_\nu = \frac{\cos\dfrac{2\nu-1}{2m}\pi}{\cos\dfrac{\pi}{2m}} \qquad \nu = 1, 2, \ldots, m.$$

Dann gilt

für m ungerade:

$$\|P[g]-g\| \leqq \frac{1}{2}(\|P\|+1)\,\omega\left(g;\ 2\sin\frac{\pi}{2m}\right),$$

für m gerade:

$$\|P[g]-g\| \leqq \frac{1}{2}(\|P\|+1)\,\omega\left(g;\ 2\tan\frac{\pi}{2m}\right).$$

§ 2. Der Beweis

Wir gehen aus von folgendem

LEMMA. *Sei*

$$\sum_{\mu=0}^{m} a_\mu = 0 \tag{1}$$

und

$$\operatorname{sgn} \sum_{\mu=0}^{\nu-1} a_\mu = -\operatorname{sgn} \sum_{\mu=0}^{\nu} a_\mu \qquad \nu = 1, 2, \ldots, m-1. \tag{2}$$

Dann ist

$$\left| \sum_{\nu=0}^{m} a_\nu b_\nu \right| \leqq \frac{1}{2} \sum_{\nu=0}^{m} |a_\nu| \max_{1 \leqq k \leqq m} |b_{k-1} - b_k|.$$

BEWEIS DES LEMMAS. Partielle Summation ergibt

$$\left| \sum_{\nu=0}^{m} a_\nu b_\nu \right| = \left| \sum_{\nu=1}^{m} (b_{\nu-1} - b_\nu) \sum_{\mu=0}^{\nu-1} a_\mu \right| \leqq \max_{1 \leqq k \leqq m} |b_{k-1} - b_k| \sum_{\nu=1}^{m} \left| \sum_{\mu=0}^{\nu-1} a_\mu \right|.$$

Weiterhin gilt

$$\sum_{\nu=1}^{m} \left| \sum_{\mu=0}^{\nu-1} a_\mu \right| = \left| \sum_{\nu=1}^{m} (-1)^\nu \sum_{\mu=0}^{\nu-1} a_\mu \right| = \frac{1}{2} \left| \sum_{\nu=0}^{m} \left(1 - (-1)^{\nu+m}\right) a_\nu \right| =$$

$$= \frac{1}{2} \left| \sum_{\nu=0}^{m} (-1)^\nu a_\nu \right| \leqq \frac{1}{2} \sum_{\nu=0}^{m} |a_\nu|,$$

womit das Lemma bewiesen ist.

Sowohl $P[f]$ als auch $T[f]$ lassen sich in der Form

$$\sum_{\nu=1}^{m} f(x_\nu) l_\nu(x)$$

mit wohlbestimmten Polynomen bzw. trigonometrischen Ausdrücken l_ν schreiben. Auf die abzuschätzende Differenz

$$\sum_{\nu=1}^{m} f(x_\nu) l_\nu(x) - f(x) \quad \text{mit} \quad x \in]x_{\varkappa-1}, x_\varkappa[$$

kann man das Lemma anwenden, indem man setzt

$$a_\nu = \begin{cases} l_{\nu+1}(x) & \nu = 0, 1, \ldots, \varkappa-2 \\ -1 & \nu = \varkappa-1 \\ l_\nu(x) & \nu = \varkappa, \varkappa+1, \ldots, m \end{cases}$$

$$b_\nu = \begin{cases} f(x_{\nu+1}) & \nu = 0, 1, \ldots, \varkappa-2 \\ f(x) & \nu = \varkappa-1 \\ f(x_\nu) & \nu = \varkappa, \varkappa+1, \ldots, m. \end{cases}$$

Man erkennt, daß das Lemma jetzt unmittelbar Satz 1 und Satz 2 liefert. Es bleibt nachzuweisen, daß die Voraussetzungen (1) und (2) erfüllt sind. Nun ist (1) sicher richtig, denn es bedeutet ja, daß Konstante durch P und T in sich abgebildet werden. Voraussetzung (2) wird erfüllt sein, wenn für $x \in]x_{\varkappa-1}, x_\varkappa[$ gilt

$$\operatorname{sgn} \sum_{\mu=1}^{\nu} l_\mu(x) = (-1)^{\varkappa+\nu+1} \quad \text{für} \quad 1 \leqq \nu \leqq \varkappa-1 \tag{3}$$

$$\operatorname{sgn}\left[\sum_{\mu=1}^{\nu-1} l_\mu(x) - 1\right] = (-1)^{\varkappa+\nu+1} \quad \text{für} \quad \varkappa \leqq \nu \leqq m. \tag{4}$$

(4) ist offenbar äquivalent mit

$$\operatorname{sgn} \sum_{\mu=\nu}^{m} l_\mu(x) = (-1)^{\varkappa+\nu} \quad \text{für} \quad \varkappa \leqq \nu \leqq m. \tag{5}$$

NACHWEIS VON (3) UND (5) FÜR T.
Es reicht aus, den Beweis nur für $\varkappa = n+1$ zu führen, denn ist

$$\max_x |T[f](x) - f(x)| = |T[f](y_0) - f(y_0)|$$

mit $y_0 \in]x_\lambda, x_{\lambda+1}[$, so gilt offenbar

$$\|T[f] - f\| = \max_{x_n < x < x_{n+1}} |T[f_\lambda](x) - f_\lambda(x)|,$$

wo $f_\lambda(x) = f(x + x_\lambda - x_n)$ ist und daher $\omega(f_\lambda; \cdot) = \omega(f; \cdot)$ gilt.
Sei also nun $x \in]x_n, x_{n+1}[$. Aus den bekannten expliziten Ausdrücken

$$l_\nu(x) = \frac{(-1)^\nu}{m} \frac{\sin \frac{m}{2} x}{\sin \frac{1}{2}(x - x_\nu)} \qquad (m = 2n+1)$$

$$l_\nu(x) = \frac{(-1)^\nu}{m} \sin \frac{m}{2} x \cot \frac{1}{2}(x - x_\nu) \qquad (m = 2n)$$

liest man ab

$$|l_1(x)| < |l_2(x)| < \ldots < |l_n(x)|,$$

$$|l_{n+1}(x)| > |l_{n+2}(x)| > \ldots > |l_m(x)|,$$

$$\operatorname{sgn} l_\nu(x) = \begin{cases} (-1)^{\nu+n} & \text{für} \quad 1 \leqq \nu \leqq n \\ (-1)^{\nu+n+1} & \text{für} \quad n+1 \leqq \nu \leqq m, \end{cases}$$

woraus (3) und (5) leicht zu folgern sind.

NACHWEIS VON (3) UND (5) FÜR P.

Die Beweise von (3) und (5) sind ähnlich, wir führen nur den von (3) aus. Für $v=m$ ist die Behauptung sofort als richtig zu erkennen, es sei daher im folgenden $v<m$ vorausgesetzt. Es genügt zu beweisen, daß die Funktion

$$h(x):=\sum_{\mu=1}^{v} l_\mu(x)$$

auf $[x_v, x_{m+1}]$ bei $x_{v+1}, x_{v+2}, \ldots, x_m$ und nur an diesen Stellen das Zeichen wechselt. Wäre es anders, so folgte mit Rolles Satz sofort, daß h' auf $[x_v, x_{m+1}]$ mindestens $m-v$ Nullstellen hat. Da wegen $h(x_\mu)=1$ $(\mu=1, 2, \ldots, v)$ h' auf $]x_1, x_v[$ mindestens $v-1$ Nullstellen hat, hätte dann h' $m-1$ Nullstellen, was einen Widerspruch darstellt, da h' ein nichtverschwindendes Polynom vom Grad $m-2$ ist.

LITERATUR

[1] CHENEY, E. W.—PRICE, K. H.: Minimal Projections, (in A. Talbot (Ed.): *Approximation Theory*, Academic Press London and New York, 1970).

[2] GÜNTTNER, R.: *Abschätzungen für Normen von Interpolationsoperatoren* (Dissertation, Clausthal—Z., 1972).

[3] KIS, O.: Bemerkungen zur Abschätzung des Fehlers beim Interpolieren (Russisch). *Studia Sci. Math. Hungar.* **4** (1969) 279—289.

Math. Institut der Tech. Universität Clausthal

(Eingegangen: 16. April 1973)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 369—377.

ON A RESULT OF SALEM AND ZYGMUND CONCERNING RANDOM POLYNOMIALS

by

G. HALÁSZ

Let ε_k $(k=1, 2, \ldots)$ be independent random variables assuming $+1$ and -1 with probability $1/2$ and consider the random triginometric polynomial

$$f_n(\vartheta) = \sum_{k=1}^{n} \varepsilon_k \cos k\vartheta.$$

It has been found by SALEM and ZYGMUND [1] that with probability 1

$$\frac{1}{2\sqrt{6}} \leqq \liminf_{n\to\infty} \frac{\max\limits_{0\leqq\vartheta\leqq 2\pi} |f_n(\vartheta)|}{\sqrt{n\log n}} \leqq \limsup_{n\to\infty} \frac{\max\limits_{0\leqq\vartheta\leqq 2\pi} |f_n(\vartheta)|}{\sqrt{n\log n}} \leqq 1.$$

They asked whether there is a limit with probability 1. In HAYMAN's "Research Problems..." [2] the same question is raised for power polynomials (Problem 4.17). The aim of this paper is to answer this question in the affirmative with the upper bound of SALEM and ZYGMUND as the limit.

We formulate our result for the trigonometric case but the same holds, with minor difference in the proof that we shall point out at the end of the paper, for power polynomials as well.

THEOREM. *Let ε_k be independent, $\varepsilon_k = +1$ and -1 with probability $1/2$. Then with probability* 1

$$\sqrt{n\log n} - 4\sqrt{\frac{n}{\log n}}\log\log n \leqq \max_{0\leqq\vartheta\leqq 2\pi} |f_n(\vartheta)| = \max_{0\leqq\vartheta\leqq 2\pi} \left|\sum_{k-1}^{n} \varepsilon_k \cos k\vartheta\right| \leqq$$

$$\leqq \sqrt{n\log n} + 3\sqrt{\frac{n}{\log n}}\log\log n$$

for large enough n.

The order of the remainder is probably the right one but no attempt is made to obtain best possible constants. For fixed n one can drop the $\log\log n$ with probability $1-\varepsilon$, 4 and 3 replaced by a $c(\varepsilon)$ and the quantity is likely to have a limit distribution with variance const. $\sqrt{\dfrac{n}{\log n}}$. For the sake of simplicity we restricted ourselves to coefficients ± 1 but one can also get results for the more general cases considered by Salem and Zygmund. More about these generalizations at the end of the paper.

PROOF. We begin by proving the more difficult lower estimation. The upper bound could also be obtained by the method of SALEM and ZYGMUND.

Let $u(x)$, to be defined later, satisfy

$$0 \leqq u(x) \leqq 1 \quad (-\infty < x < +\infty), \qquad u(x) = \begin{cases} 0 & \text{for } |x| \leqq M \\ 1 & \text{for } |x| \geqq M + \Delta. \end{cases}$$

In remainder terms we shall put right away $\frac{3}{4}\sqrt{n \log n} \leqq M \leqq \frac{5}{4}\sqrt{n \log n}$, $\Delta = \sqrt{\frac{n}{\log n}}$ but in main terms we shall keep the parameters in order to motivate their choices. Let further $\vartheta = \frac{2m-1}{2n} 2\pi$ $(m = 1, \ldots, n)$ and consider the random variable

$$\eta = \sum_{m=1}^{n} u(f_n(\vartheta_m)).$$

If $\eta > 0$, then necessarily $\max_{0 \leqq \vartheta \leqq 2\pi} |f_n(\vartheta)| \geqq \max_{m=1,\ldots,n} |f_n(\vartheta_m)| > M$. This we shall show, by applying Tchebyscheff's inequality, to occur with great probability.

We use the characteristic function method, i.e. represent $u(x)$ as a Fourier—Stieltjes transform

$$u(x) = \int_{-\infty}^{\infty} e^{ixt}\, dU(t).$$

In our definition of $u(x)$, $U(t)$ will be of bounded variation on $(-\infty, \infty)$. We have

$$\eta = \sum_{m=1}^{n} \int_{-\infty}^{\infty} e^{it \sum_{k=1}^{n} \varepsilon_k \cos k\vartheta_m}\, dU(t).$$

For evaluating $\mathsf{E}(\eta)$, the expectation of η we have, by independence,

$$\mathsf{E}\left(e^{it \sum_{k=1}^{n} \varepsilon_k \cos k\vartheta_m}\right) = \prod_{k=1}^{n} \mathsf{E}(e^{it\varepsilon_k \cos k\vartheta_m}) = \prod_{k=1}^{n} \cos(t \cos k\vartheta_m) =$$

$$= e^{-\frac{t^2}{2} \sum_{k=1}^{n} \cos^2 k\vartheta_m} + O(nt^4)$$

for $|t| \leqq 1$, by expanding log cos and using the fact that $e^{-a} = e^{-b} + O(|b-a|)$ uniformly for $a, b \geqq 0$. Here

$$\sum_{k=1}^{n} \cos^2 k\vartheta_m = \frac{n}{2} + \frac{1}{2} \sum_{k=1}^{n} \cos 2k\vartheta_m = \frac{n}{2},$$

provided that n is even which we shall assume in the sequel. For $|t| > 1$ we use the trivial bound 1. Multiplying by $dU(t)$, integration and summation give

$$\mathsf{E}(\eta) = n \int_{-1}^{1} e^{-\frac{nt^2}{4}}\, dU(t) + O\left(n^2 \int_{-\infty}^{\infty} t^4 |dU(t)|\right) + O\left(n \int_{|t| \geqq 1} |dU(t)|\right).$$

We extend integration to infinity in the first term on the right, making the same error as in the third term that, in turn, can be included in the second. The estimation of

remainders like the latter we postpone to (4)—(7). By (5) and

$$\Delta = \sqrt{\frac{n}{\log n}},$$

$$\mathsf{E}(\eta) = n \int_{-\infty}^{\infty} e^{-\frac{nt^2}{4}} dU(t) + O(\log^2 n). \tag{1}$$

As to the second moment,

$$\eta^2 = \sum_{m=1}^{n} \sum_{l=1}^{n} u(f_n(\vartheta_m)) u(f_n(\vartheta_l)).$$

The $2n$ terms $m=l$ and $m+l = n+1$ contribute

$$2 \sum_{m=1}^{n} u^2(f_n(\vartheta_m)) \leqq 2 \sum_{m=1}^{n} u(f_n(\vartheta_m)) = 2\eta \tag{2}$$

(note that $0 \leqq u \leqq 1$). For the rest we again represent $u(x)$ by its Fourier—Stieltjes integral, giving rise to terms

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i \sum_{k=1}^{n} \varepsilon_k (t \cos k\vartheta_m + \tau \cos k\vartheta_l)} dU(t)\, dU(\tau). \tag{3}$$

Applying this time,

$$\log \cos x = -\frac{x^2}{2} - \frac{x^4}{12} + O(x^6) \qquad (|x| \leqq 1)$$

and

$$e^{-a} = e^{-b} + (b-a)e^{-b} + O((b-a)^2)$$

uniformly for $a, b \geqq 0$,

$$\mathsf{E}\left(e^{i \sum_{k=1}^{n} \varepsilon_k (t \cos k\vartheta_m + \tau \cos k\vartheta_l)}\right) = \prod_{k=1}^{n} \cos(t \cos k\vartheta_m + \tau \cos k\vartheta_l) =$$

$$= e^{-\frac{1}{2} \sum_{k=1}^{n} (t \cos k\vartheta_m + \tau \cos k\vartheta_l)^2} -$$

$$- \frac{1}{12} \sum_{k=1}^{n} (t \cos k\vartheta_m + \tau \cos k\vartheta_l)^4 e^{-\frac{1}{2} \sum_{k=1}^{n} (t \cos k\vartheta_m + \tau \cos k\vartheta_l)^2} +$$

$$+ O(n(|t| + |\tau|)^6) + O(n^2(|t| + |\tau|)^8)$$

for $|t| \leqq \frac{1}{2}$, $|\tau| \leqq \frac{1}{2}$. Here

$$\sum_{k=1}^{n} (t \cos k\vartheta_m + \tau \cos k\vartheta_l)^2 =$$

$$= t^2 \sum_{k=1}^{n} \cos^2 k\vartheta_m + \tau^2 \sum_{k=1}^{n} \cos^2 k\vartheta_l + t\tau \sum_{k=1}^{n} (\cos k(\vartheta_m - \vartheta_l) + \cos k(\vartheta_m - \vartheta_l)) =$$

$$= \frac{nt^2}{2} + \frac{n\tau^2}{2},$$

for now $m \neq l$, $m+l \neq n+1$. This, together with the trivial bound 1 of the integrand outside the square $|t| \leqq 1/2$, $|\tau| \leqq 1/2$ yields the expectation of (3) as

$$\int_{-1/2}^{1/2}\int_{-1/2}^{1/2} e^{-\frac{nt^2}{4}-\frac{n\tau^2}{4}}\, dU(t)\, dU(\tau) -$$

$$-\frac{1}{12}\sum_{k=1}^{n}\int_{-1/2}^{1/2}\int_{-1/2}^{1/2}(t\cos k\vartheta_m+\tau\cos k\vartheta_l)^4 e^{-\frac{nt^2}{4}-\frac{n\tau^4}{4}}\, dU(t)\, dU(\tau)+$$

$$+O\Big(n\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(|t|+|\tau|)^6\,|dU(t)|\,|dU(\tau)|\Big)+O\Big(n^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(|t|+|\tau|)^8\,|dU(t)|\,|dU(\tau)|\Big)+$$

$$+O\Big(\int_{|t|\geqq 1/2}\int_{-\infty}^{\infty}|dU(t)|\,|dU(\tau)|\Big).$$

We may extend integration to infinity in the first two terms, the errors made are inferior to the third term and so is the last term. The first term then becomes

$$\Big(\int_{-\infty}^{\infty} e^{-\frac{nt^2}{4}}\, dU(t)\Big)^2.$$

The second will consist of products

$$\int_{-\infty}^{\infty} t^p e^{-\frac{nt^2}{4}}\, dU(t)\int_{-\infty}^{\infty} t^r e^{-\frac{nt^2}{4}}\, dU(t)$$

and by (6) (but not directly by (5)), after summation for k, each of these is roughly $O\left(\frac{ne^{-2M^2/n}}{\Delta^4 M^2}\right) = O\left(\frac{1}{n^2}\right)$. The third and fourth terms contain products

$$\int_{-\infty}^{\infty}|t|^p|dU(t)|\int_{-\infty}^{\infty}|t|^r|dU(t)| \qquad (p, r = 0, 1, \ldots,\ p+r = 6 \text{ and } p+r = 8)$$

and by (4) and (5) the contribution is $O\left(n\frac{M}{\Delta^7}+n^2\frac{M}{\Delta^9}\right) = O\left(\frac{\log^5 n}{n^2}\right)$. The expectation of (3) we have just calculated, together with (2) gives

$$\mathsf{E}(\eta^2) \leqq 2\mathsf{E}(\eta)+n^2\Big(\int_{-\infty}^{\infty} e^{-\frac{nt^2}{4}}\, dU(t)\Big)^2+O(\log^5 n).$$

Combining with the result for the first moment, the square of the variance becomes

$$\mathsf{D}^2(\eta) = \mathsf{E}(\eta^2)-\mathsf{E}^2(\eta) \leqq$$

$$\leqq 2\mathsf{E}(\eta)+O(\log^5 n)+O\big(\mathsf{E}(\eta)\log^2 n\big)+O(\log^4 n) = O\big(\mathsf{E}(\eta)\log^2 n+\log^5 n\big).$$

In the event $\eta=0$, η will deviate by $\mathsf{E}(\eta)$ from its expectation and from Tchebycheff's inequality

$$\mathsf{P}(\eta=0)=\frac{O(\mathsf{E}(\eta)\log^2 n+\log^5 n)}{\mathsf{E}^2(\eta)}=O\left(\frac{1}{\log^4 n}\right),$$

provided that $\mathsf{E}(\eta)\gg\log^6 n$. Again by (1) it is enough to show

$$\int_{-\infty}^{\infty} e^{-\frac{nt^2}{4}}\,dU(t)\geqq\frac{\log^6 n}{n}.$$

By Parseval's formula, the Fourier transform of $e^{-\frac{nt^2}{4}}$ being $\sqrt{\frac{4\pi}{n}}e^{-\frac{x^2}{n}}$, the integral can be rewritten as

$$\frac{1}{\sqrt{\pi n}}\int_{-\infty}^{\infty} e^{-\frac{x^2}{n}}u(x)\,dx\geqq\frac{1}{\sqrt{\pi n}}\int_{M+\Delta\leqq|x|\leqq M+2\Delta} e^{-\frac{x^2}{n}}\,dx\geqq\frac{2\Delta}{\sqrt{\pi n}}e^{-\frac{(M+2\Delta)^2}{n}}$$

and the inequality it is to satisfy will indeed hold if $M=\sqrt{n\log n}-3{,}5\sqrt{\frac{n}{\log n}}\log\log n$. Recalling our starting remark, we have thus proved

$$\mathsf{P}\left(\max_{0\leqq\vartheta\leqq 2\pi}|f_n(\vartheta)|\geqq\sqrt{n\log n}-3{,}5\sqrt{\frac{n}{\log n}}\log\log n\right)=1-O\left(\frac{1}{\log^4 n}\right).$$

For the upper estimation we have to take a denser set of points than $\{\vartheta_m\}$ and we consider

$$\eta=\int_0^{2\pi}u(f_n(\vartheta))\,d\vartheta=\int_0^{2\pi}\int_{-\infty}^{\infty}e^{it\sum_{k=1}^{n}\varepsilon_k\cos k\vartheta}\,dU(t)\,d\vartheta$$

with the same $\Delta=\sqrt{\frac{n}{\log n}}$ but a different M satisfying again

$$\frac{3}{4}\sqrt{n\log n}\leqq M\leqq\frac{5}{4}\sqrt{n\log n}.$$

The properties of $u(x)$ imply:

measure on $(0, 2\pi)$ of the set $\{\vartheta:|f_n(\vartheta)|\geqq M+\Delta\}\leqq\eta$.

On the other hand, with the notation $K=\max_{0\leqq\vartheta\leqq 2\pi}|f_n(\vartheta)|=|f_n(\vartheta_0)|$, $|f_n'(\vartheta)|\leqq Kn$ by Bernstein's theorem and so $|f_n(\vartheta)|\geqq M+\Delta$ on a whole interval of length $\frac{2(K-M-\Delta)}{Kn}$ around ϑ_0. Hence

$$\frac{2(K-M-\Delta)}{Kn}\leqq\eta$$

and in case $\eta\leqq\frac{1}{n\log n}$,

$$K=\max_{0\leqq\vartheta\leqq 2\pi}|f_n(\vartheta)|\leqq(M+\Delta)\left(1+\frac{1}{\log n}\right)\leqq M+3\Delta.$$

(In the argument we have assumed $K \geqq M+\Delta$, but the conclusion is trivial otherwise.)

As before,

$$\mathsf{E}\left(e^{it\sum_{k=1}^{n}\varepsilon_k \cos k\vartheta}\right) = \prod_{k=1}^{n} \cos(t\cos k\vartheta) = e^{-\frac{t^2}{2}\sum_{k=1}^{n}\cos^2 k\vartheta} -$$

$$-\frac{t^4}{12}\sum_{k=1}^{n}\cos^4 k\vartheta\, e^{-\frac{t^2}{2}\sum_{k=1}^{n}\cos^2 k\vartheta} + O(nt^6+n^2t^8)$$

for $|t| \leqq 1$ and

$$\mathsf{E}(\eta) = \int_0^{2\pi}\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}\sum_{k=1}^{n}\cos^2 k\vartheta}\, dU(t)\, d\vartheta -$$

$$-\frac{1}{12}\sum_{k=1}^{n}\int_0^{2\pi}\cos^4 k\vartheta \int_{-\infty}^{\infty} t^4 e^{-\frac{t^2}{2}\sum_{k=1}^{n}\cos^2 k\vartheta}\, dU(t)\, d\vartheta +$$

$$+O\left(\int_{-\infty}^{\infty}(nt^6+n^2t^8)\,|dU(t)|\right) =$$

$$= O\left(\frac{\sqrt{n}\, e^{-\frac{M^2}{n}}}{M}\right) + O\left(\frac{n}{\Delta^6}+\frac{n^2}{\Delta^8}\right) = O\left(\frac{e^{-\frac{M^2}{n}}}{\sqrt{\log n}}\right)$$

by (7) ($p=0$ and $p=4$) and (5) ($p=6$ and $p=8$), respectively. By Markoff's inequality

$$\mathsf{P}\left(\eta \leqq \frac{1}{n\log n}\right) \geqq 1 - O\left(ne^{-\frac{M^2}{n}}\sqrt{\log n}\right) = 1 - O\left(\frac{1}{\log^4 n}\right),$$

provided that

$$M = \sqrt{n\log n} + 2{,}5\sqrt{\frac{n}{\log n}}\log\log n$$

and this, as we remarked, implies

$$\mathsf{P}\left(\max_{0\leqq\vartheta\leqq 2\pi}|f_n(\vartheta)| \leqq \sqrt{n\log n} + 2{,}7\sqrt{\frac{n}{\log n}}\log\log n\right) = 1 - O\left(\frac{1}{\log^4 n}\right).$$

Let now n_j be an increasing sequence of even integers. If in both the lower and the upper estimation the sum of the exceptional probabilities for $n=n_j$ is finite, i.e. if

$$\sum_j \frac{1}{\log^4 n_j} < +\infty,$$

then we have proved our theorem for the sequence $n=n_j$. Owing to Lemma (4.4.1) of SALEM and ZYGMUND [1] (that could be proved the same way we have obtained our upper bound) stating

$$\max_{n_j < n \leqq n_{j+1}} \max_{0\leqq\vartheta\leqq 2\pi} |f_n(\vartheta) - f_{n_j}(\vartheta)| \leqq 3\sqrt{(n_{j+1}-n_j)\log n_{j+1}}$$

with probability 1 for large enough j (for any increasing sequence n_j), the validity of our Theorem follows from its validity for the sequence $\{n_j\}$, if $n_{j+1}-n_j = O\left(\frac{n_j}{\log^2 n_j}\right)$, since then $\sqrt{(n_{j+1}-n_j)\log n_{j+1}} = O\left(\sqrt{\frac{n}{\log n}}\right)$ for $n_j<n\leqq n_{j+1}$. (The constants 3,5 and 2,7 are to be enlarged somewhat.) We choose $n_j=2[e^{j^{1/3}}]$. Then

$$\sum_j \frac{1}{\log^4 n_j} \leqq 4 \sum_j \frac{1}{j^{4/3}} < +\infty$$

and

$$\frac{n_{j+1}}{n_j} \leqq \frac{2}{n_j} + e^{(j+1)^{1/3}-j^{1/3}} = \frac{2}{n_j}+1+O\left(\frac{1}{j^{2/3}}\right) = 1+O\left(\frac{1}{\log^2 n_j}\right).$$

Now we construct $u(x)$ for the remaining $M\leqq|x|\leqq M+\Delta$. Let us fix a ten times continuously differentiable function in [0, 1] which is 0 at 0, 1 at 1 and is between 0 and 1 between 0 and 1, with vanishing derivatives up to the 10th order at 0 and 1. Transforming [0, 1] linearly into $[M, M+\Delta]$ and symmetrically into $[-M, -M-\Delta]$ we arrive at a function $u(x)$ that is 0 for $|x|\leqq M$, 1 for $|x|\geqq M+\Delta$ and ten times differentiable on $(-\infty, \infty)$ with $|u^{(r)}(x)| \leqq \frac{\text{const.}}{\Delta_r}$ $(r=0, 1, \ldots, 10)$. This is all we need to deduce the necessary estimations.

Since $u(x)$ is 1 in a neighbourhood of $\pm\infty$, $U(t)$ must have a jump of unity at $t=0$. Apart from this it is even absolutely continuous with the derivative given by Plancherel's formula

$$v(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}(u(x)-1)e^{-itx}\,dx = \frac{1}{2\pi}\int_{-M-\Delta}^{M+\Delta}(u(x)-1)e^{-itx}\,dx.$$

We have

$$|v(t)| \leqq \frac{2(M+\Delta)}{2\pi} = O(M)$$

and by partial integration

$$|t^r v(t)| \leqq 2\Delta \max|u^{(r)}(x)| = O\left(\frac{1}{\Delta^{r-1}}\right) \qquad (r=1, \ldots, 10).$$

Using the first for $|t| \leqq \frac{1}{\Delta}$ and the second with $r=2$ for $|t| \geqq \frac{1}{\Delta}$

$$\int_{-\infty}^{\infty}|dU(t)| = 1+\int_{-\infty}^{\infty}|v(t)|\,dt = O\left(\frac{M}{\Delta}\right) \tag{4}$$

and using the second with $r=p$ for $|t| \leqq \frac{1}{\Delta}$ and with $r=p+2$ for $|t| \geqq \frac{1}{\Delta}$

$$\int_{-\infty}^{\infty}|t|^p\,|dU(t)| = \int_{-\infty}^{\infty}|t|^p\,|v(t)|\,dt = O\left(\frac{1}{\Delta^p}\right) \qquad (p=1, \ldots, 8). \tag{5}$$

Next, by the formula we have already used, the Fourier transform of $t^p dU(t)$ being $i^{-p}u^{(p)}(x)$,

$$(6) \qquad \left| \int_{-\infty}^{\infty} e^{-\frac{nt^2}{4}} t^p \, dU(t) \right| = \frac{1}{\sqrt{\pi n}} \left| \int_{-\infty}^{\infty} e^{-\frac{x^2}{n}} u^{(p)}(x) \, dx \right| =$$

$$= O\left(\frac{1}{\sqrt{n}\, \Delta^p} \int_{|x| \leqq M} e^{-\frac{x^2}{n}} dx \right) = O\left(\frac{\sqrt{n}}{M\Delta^p} e^{-\frac{M^2}{n}} \right) \qquad (p = 0, \ldots, 8).$$

The same way

$$\int_{-\infty}^{\infty} e^{-\frac{t^2}{2} \sum_{k=1}^{n} \cos^2 k\vartheta} t^p \, dU(t) = O\left(\frac{\sqrt{n}}{M\Delta^p} e^{\frac{-M^2}{2 \sum_{k=1}^{n} \cos^2 k\vartheta}} \right).$$

Here

$$\sum_{k=1}^{n} \cos^2 k\vartheta = \frac{n}{2} + \frac{\sin(2n+1)\vartheta}{4 \sin \vartheta} - \frac{1}{4} \leqq \begin{cases} \dfrac{n}{2} + \dfrac{n}{2 \log n} & \text{for } \dfrac{\log n}{n} \leqq |\vartheta| \leqq \dfrac{\pi}{2} \\ n & \text{everywhere.} \end{cases}$$

Hence

$$\int_0^{2\pi} e^{-\frac{M^2}{2 \sum_{k=1}^{n} \cos^2 k\vartheta}} d\vartheta = 4 \int_0^{\pi/2} e^{-\frac{M^2}{2 \sum_{k=1}^{n} \cos^2 k\vartheta}} d\vartheta \leqq \frac{4 \log n}{n} e^{-\frac{M^2}{2n}} + 2\pi e^{\frac{M^2}{n + \frac{n}{\log n}}}.$$

Since

$$\frac{\log n}{n} \leqq e^{-\frac{M^2}{2n}}$$

and

$$e^{-\frac{M^2}{n + \frac{n}{\log n}}} = e^{-\frac{M^2}{n}} e^{\frac{M^2}{n(\log n + 1)}} = O\left(e^{-\frac{M^2}{n}}\right),$$

we get

$$(7) \qquad \int_0^{2\pi} \left| \int_{-\infty}^{\infty} e^{-\frac{t^2}{2} \sum_{k=1}^{n} \cos^2 k\vartheta} t^p \, dU(t) \right| d\vartheta = O\left(\frac{\sqrt{n}}{M\Delta^p} e^{-\frac{M^2}{n}} \right) \qquad (p = 0, \ldots, 8)$$

and the proof is completed.

Remarks on generalizations. The relatively short proof for the lower bound was due to the fact the values $f_n(\vartheta_m)$ for different m are close to independent.*

* We could not find a general theorem in the literature that would enable us to substantially reduce the amount of calculation for getting this independence and we have, therefore, preferred not to use any reference. As a variant of our method of proof, in place of real Fourier technique we could have applied the saddle point method in the complex plane as is usual in theorems on large deviations. However, our deviations being not too large, the advantage would not be significant. Prof. P. ERDŐS informs us that in his attempt to solve the problem of SALEM and ZYGMUND he, too, tried to prove independence (by a different method).

This is no longer true if we are to prove the sharper result for fixed large n mentioned at the beginning where we have to take a denser set just as in the upper estimation. However, the same proof still applies and possibly one can also get a limit distribution along these lines by examining higher moments than just the second, only the calculations become more troublesome. (For the upper bound one has to use Bernstein's theorem for $f_n''(\vartheta)$ and $f_n'(\vartheta_0)=0$.) The exceptionally strong independence ceases to hold also in the case of more general coefficients. Here, in addition, to obtain sharp results, one has to choose $u(x)$ smoother so as to make $U(t)$ as small as possible.

As to the power polynomials $P_n(z) = \sum_{k=1}^{n} \varepsilon_k z^k$ ($|z|=1$), the same lower bound holds, since $f_n(\vartheta)=\operatorname{Re} P_n(e^{i\vartheta})$ and the upper bound can be obtained by taking $\operatorname{Re} e^{i\alpha} P_n(e^{i\vartheta})$ in place of $f_n(\vartheta)$, with a number of fixed α's. For more precise information, especially for the limit distribution, one has to consider $P_n(z)$ as a vector valued variable and apply two dimensional Fourier-technique. The limit distribution is probably similar but shifted by a

$$c\sqrt{\frac{n}{\log n}}\log\log n \qquad (c>0).$$

REFERENCES

[1] SALEM, R. and ZYGMUND, A.: Some properties of trigonometric series whose terms have random signs, *Acta Mathematica* **91** (1954), 245—301.

[2] HAYMAN, W. K.: *Research Problems in Function Theory*, The Athlone Press, University of London, 1967.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received April 20, 1973)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 379—390.

A THREE PART SPERNER THEOREM

by

G. KATONA

Introduction

Let X be a finite set of n elements, and let $\mathcal{A}=\{A_1, \ldots, A_m\}$ be a family of different subsets of X such that $A_i \not\supset A_j$ for $i \neq j$. SPERNER [1] proved that in this case

$$m \leqq \binom{n}{\left[\frac{n}{2}\right]}. \tag{1}$$

KLEITMAN [2] and KATONA [3] independently proved, that if we divide X into two disjoint parts ($X=X_1 \cup X_2$, $X_1 \cap X_2=\emptyset$) and the family $\mathcal{A}$ satisfies the property that for any different i and j

$$A_i \cap X_1 = A_j \cap X_1 \quad \text{implies} \quad A_i \cap X_2 \not\supset A_j \cap X_2$$

and

$$A_i \cap X_2 = A_j \cap X_2 \quad \text{implies} \quad A_i \cap X_1 \not\supset A_j \cap X_1,$$

then (1) holds again. Here, the conditions are weaker, because $A_i \supset A_j$ is excluded only in some particular cases, when A_i and A_j are equal in X_1 or in X_2, respectively. However, the maximal m under this weaker condition is the same as in the Sperner theorem.

The question arises if the conditions of this result can be still weakened. The natural way would be to divide X into 3 disjoint parts and to exclude the inclusion $A_i \supset A_j$ only if A_i and A_j are equal in two of the parts. However, under this condition it is possible to choose $m > \binom{n}{\left[\frac{n}{2}\right]}$ as it is shown by an example in [3]. The main aim of this paper is to give an additional condition which ensures (1).

We treat the problem in a more general context which is described in [4] and [5], but the definitions and the basic idea are briefly restated here.

Using the method of these papers Sperner type theorems can be reduced to problems of determining a set of points maximal with respect to certain conditions in a two or three dimensional lattice-configuration. We think that these problems are of independent interest.

Definitions and the Theorem

We say that the finite set is a *partially ordered set* if a relation $<$ is defined on G with the following properties: a) at most one of the relations $g_1<g_2$, $g_1=g_2$, $g_2<g_1$ holds; (We say that g_1 and g_2 are *comparable* if one of them is true.) b) if $g_1<g_2$ and $g_2<g_3$ then $g_1<g_3$.

g_2 covers g_1 if $g_1<g_2$ and there is no g_3 satisfying $g_1<g_3<g_2$, that is, if g_2 is the immediate successor of g_1. Assume that there is a rank function $r(g)$ which makes correspond a nonnegative integer to every element of G, so that if g_2 covers g_1 then $r(g_2)=r(g_1)+1$ and there is at least one element $g\in G$ for which $r(g)=0$. We say in this case that G is a *partially ordered set with a rank function.*

A *chain of length* l is a sequence $g_1, \ldots, g_l \in G$, where g_l covers g_{l-1}, g_{l-1} covers $g_{l-2}, \ldots, g_2$ covers g_1. A chain is *symmetrical* if $r(g_1)+r(g_2)=n$, where $n=\max_{g\in G} r(g)$. A partially ordered set is a *symmetrical chain set* if it can be divided into disjoint symmetrical chains.

If G and H are partially ordered sets, then the *direct sum* $G+H$ is the set of ordered pairs (g, h), $g\in G$, $h\in H$ with the ordering $(g_1, h_1)\leqq(g_2, h_2)$ iff $g_1\leqq g_2$ and $h_1\leqq h_2$.

If the rank functions of G and H are r and s, respectively then we can define a rank function on $G+H$ as follows:

$$t((g, h)) = r(g)+s(h).$$

For example, the subsets of a finite set X of n elements form a partially ordered set with rank function $r(A)=|A|$, (that is the number of elements) if we order them by inclusion. A covers B $(A, B\subset X)$ iff $A\supset B$ and $|A-B|=1$. By a theorem of DE BRUIJN, KRUSYSWIJK and TENGEBERGEN [6] we know that this is a symmetrical chain set. (More generally this is proved for the set of points of an n-dimensional rectangle with integer coordinates.)

If G is the partially ordered set of the subsets of a set X_1 and H is the same for X_2 $(X_1\cap X_2=\emptyset)$, then $G+H$ is the partially ordered set of the subsets of $X_1\cup X_2$. The same is true for the disjoint sets X_1, X_2, X_3.

Now, we can formulate our theorem.

THEOREM. *Let F, G and H be symmetrical chain sets with respective rank functions p, r and s*

$$\max_{f\in F} p(f)+\max_{g\in G} r(g)+\max_{h\in H} s(h) = n.$$

Assume $(f_1, g_1, h_1), \ldots, (f_m, g_m, h_m)$ are different elements of $F+G+H$, such that there are no two different triples among them for which

$$(f_i, g_i, h_i) < (f_j, g_j, h_j) \tag{2}$$

and two of the corresponding components are equal and there are no four different triples for which

$$f_i = f_j, \quad f_k = f_l$$

$$\begin{aligned} g_i, g_k > g_j, g_l \qquad & g_i, g_k \text{ are comparable} \\ & g_j, g_l \text{ are comparable} \\ h_i, h_l > h_j, h_k \qquad & h_i, h_l \text{ are comparable} \\ & h_j, h_k \text{ are comparable} \end{aligned} \tag{3}$$

holds for any permutation of the letters f, g, h. Then m is maximal if we choose all the elements of $F+G+H$ with rank $\left[\frac{n}{2}\right]$.

Applying the theorem for the partially ordered sets of the subsets of the disjoint sets X_1, X_2, X_3 we obtain a Sperner type theorem, where we exclude some pairs and 4-tuples of subsets and the maximal size of the family under these conditions is $\binom{n}{\left[\frac{n}{2}\right]}$.

We formulate another particular case of the theorem as a separate lemma. The proof of the theorem is based on it.

LEMMA. *Let T be the direct sum of the chains $0<1<\ldots<a_1$, $0<1<\ldots<a_2$ and $0<1<\ldots<a_3$ ($a_3 \geqq a_2 \geqq a_1$ are integers). Assume $(x_1, y_1, z_1), \ldots, (x_m, y_m, z_m)$ are different elements of T ($0 \leqq x_i \leqq a_1, 0 \leqq y_i \leqq a_2$, $0 \leqq z_i \leqq a_3$ integers; $1 \leqq i \leqq m$) such that there are no two different triples with the property,*

(4) *two corresponding components are equal,*

and there are no 4 different triples for which

$$x_i = x_j, \quad x_k = x_l$$

(5)
$$y_i, y_k > y_j, y_l,$$

$$z_i, z_l > z_j, z_k$$

holds.

Then m is maximal if we choose all the elements of rank

$$x+y+z = \left[\frac{a_1+a_2+a_3}{2}\right].$$

Here it is not necessary to assume (5) for the other permutations of x, y and z.

PROOF of the theorem. From the fact that F, G and H can be split into disjoint symmetrical chains it follows that $F+G+H$ can be split into disjoint symmetrical 3-dimensional parallelotopes of type T. The word "symmetrical" means that the sum of the minimal and maximal rank in T is n. Assuming the lemma is true, we obtain the maximal number of elements of T satisfying conditions (2) and (3) ((4) and (5) follow from them) by choosing all the elements of "middle rank". By the symmetry of T they have rank $\left[\frac{n}{2}\right]$. So, under conditions (2) and (3) the maximal set of points consists of all the elements of rank $\left[\frac{n}{2}\right]$. The proof is completed. We have to prove only our lemma.

The proof of the lemma

Before starting the proof let us consider a simple example: $a_1=a_2=a_3=1$. Here $\left[\frac{a_1+a_2+a_3}{2}\right] = 1$, the number of points with $x+y+z = 1$ is 3: (0, 0, 1), (0, 1, 0), (1, 0, 0). If we add the point (1, 1, 1), we have 4 points and there are no two of them satisfying (4). Thus we really need some additional conditions excluding the 4-tuples similar to (0, 0, 1), (0, 1, 0), (1, 0, 0), (1, 1, 1). Similar examples show that we have to exclude somewhat more general 4-tuples. One possibility is (5).

An *x-plane* is the subset of T containing the points with first component equal to x. In (5) $x_i=x_j$ means that the points $P_i=(x_i, y_i, z_i)$ and $P_j=(x_j, y_j, z_j)$ are in the same x-plane. The same is true for $P_k=(x_k, y_k, z_k)$ and $P_l=(x_l, y_l, z_l)$.

$y_i>y_j$ and $z_i>z_j$ mean that these two points form an increasing function (as a function $z(y)$). On the other hand $y_k>y_l$, $z_l>z_k$ mean that these points form a decreasing function. (5) contains the further $(y_i>y_l,\ y_k>y_l,\ z_i>z_k,\ z_l>z_j)$ conditions, which mean that the intervals (y_j, y_i) and (y_l, y_k) are not disjoint, and similarly, (z_j, z_i) and (z_k, z_l) are not disjoint.

1. First let us investigate the points in one x-plane. A *row* is a set of points in an x-plane with a fixed z-component. (The *column* is defined similarly.) By (4) there are no two points in one row or column. So we can consider the points as a function. We shall prove that this *function is essentially monotonically increasing or decreasing.*

Define p by $y_p=\min\{y_j : x_j=x\}$ and q by $y_q=\max\{y_j : x_j=x\}$. If $p=q$, there is only one point in the plane, we have no problem. Otherwise $y_p<y_q$ holds. Assume $z_p<z_q$.

1a. There are no two points (x_i, y_i, z_i) and (x_l, y_l, z_l) satisfying $x_i=x_l=x$, $y_l<y_i$, $z_i, z_l>z_q$, because otherwise (5) holds with $p=j$ and $q=k$. Similarly, there are no two points (x_i, y_i, z_i) and (x_l, y_l, z_l) with $x_i=x_l=x$, $y_l>y_i$, $z_i, z_l>z_q$, changing the role of l and i. Summarizing, there are no two different points (x_i, y_i, z_i) and (x_l, y_l, z_l) with $x_i=x_l=x$ and $z_i, z_l>z_q$. It means, the function $z(y)$ has no two different values $>z_q$. We can see in a similar way that there are no two different values $<z_p$.

1b. On the other hand, assume that we have the different points (x_k, y_k, z_k) and (x_l, y_l, z_l) satisfying the conditions $k\neq p, q$, $l\neq p, q$, $x_k=x_l=x$, $y_k>y_l$, $z_l>z_k$. If $z_k<z_l<z_p$ (or $z_q<z_k<z_l$) this case is settled above. If $z_l>z_p$ (or $z_k<z_q$) then we obtain a contradiction. The points with $i=q$, $j=p$ satisfy (5). That means, there are no two values of $z(y)$ different from z_p and z_q which are in decreasing position.

1c. Finally, if (x_l, y_l, z_l) satisfies $x_l=x$ and $z_l>z_q$, then there is no point (x_k, y_k, z_k) with $x_k=x$ and $y_l<y_k<y_q$, because otherwise these points satisfy (5) choosing $i=p$, $j=q$. (If $z_k>z_l$, it belongs to the case 1a, thus we can assume $z_k<z_l$).

Summarizing our results, if we assume $z_p<z_q$, then the function $z(y)$ almost monotonically increasing; the exceptional pairs have either (x_p, y_p, z_p) or (x_q, y_q, z_q) as a member (by 1b.).

To both (x_p, y_p, z_p) and (x_q, y_q, z_q) we can have at most one exceptional pair (by 1a). If (x_p, y_p, z_p) and (x_i, y_i, z_i) form such an exceptional pair, then y_i is the second value of the domain of $z(y)$ (That is, there is no (x_k, y_k, z_k) with $x_k=x$, $y_p<y_k<y_i$). Similarly, if (x_i, y_i, z_i) and (x_q, y_q, z_q) is an exceptional pair, then

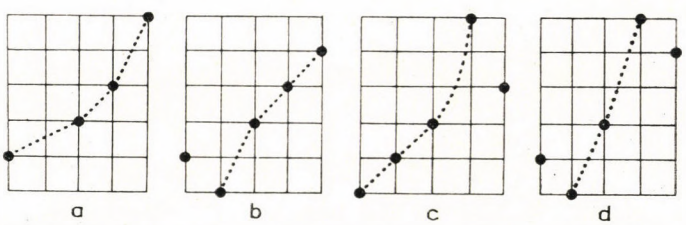


Fig. 1

y_i is the last but one element of the domain of $z(y)$. We call such a function *essentially monotonically increasing function*. Figure 1 gives the typical diagrams of such functions.

Of course, if we assume $z_p > z_q$ we obtain *essentially monotonically decreasing functions*.

2. If (x_i, y_i, z_i) and (x_j, y_j, z_j) are two points with $x_i = x_j$, $y_i < y_j$, $z_i < z_j$, then the rectangle $y_i \leqq y \leqq y_j$, $z_i \leqq z \leqq z_j$ is called a *red rectangle*. On the other hand, if $x_i = x_j$, $y_i < y_j$ and $z_i > z_j$ hold, then the rectangle $y_i \leqq y \leqq y_j$, $z_j \leqq z \leqq z_i$ is called a *blue rectangle. A red rectangle of a plane cannot contain an inner point of a blue rectangle of another plane* (and vice versa), because otherwise the 4 points determining them satisfy (5).

3. We say that an x-plane is *increasing (decreasing)* if $z_p < z_q$ ($z_q < z_p$). We now prove that *if a plane is increasing, then the union of the red rectangles (determined by its $c(x)$ points) covers a $c(x) \times c(x)$ square possibly without its two opposite corners.*

If the points of the plane are of form a, then the rectangle determined by (x_p, y_p, z_p) and (x_q, y_q, z_q) has sizes $\geqq c(x)$, thus the statement is proved in this case.

If the points are of form b, then (x_p, y_p, z_p) and (x_q, y_q, z_q) determine a rectangle with at least $c(x)$ columns and at least $c(x)-1$ rows. In this case, however, there is a point (x_r, y_r, z_r) $(x_r = x)$ satisfying $z_r < z_p$. Thus, (x_r, y_r, z_r) and (x_q, y_q, z_q) determine a rectangle with at least $c(x)-1$ columns and at least $c(x)$ rows. The union of these two rectangles cover a $c(x) \times c(x)$ square possibly except one corner.

The same holds for the case c.

In the case d, by similar arguments, we obtain the $c(x) \times c(x)$ square without two corners. In the case b the lowest point of the first column can be missing from the square, in the case c the topmost point of the last column, and in the case d both of them. We call these figures *incomplete squares.*

The same holds, of course, with blue rectangles for decreasing planes.

4. *If an increasing x-plane contains $c(x) \geqq 0$ points and a decreasing x'-plane contains $c(x') \geqq 0$ points then*

$$c(x) + c(x') \leqq a_3 + 3 \tag{6}$$

with equality only if $a_3 = a_2$.

It is easy to see that any inner point of a blue incomplete square is an inner point of some blue rectangle of the same plane. Using this fact and the statement of section 2 we obtain that the red incomplete square in the plane x cannot contain any inner point of the blue incomplete square in the plane x'. Of course, they can touch each other.

If none of the two squares contains an inner point of the other one, including now the missing corners, then one of the sizes of the rectangle $(a_2+1) \times (a_3+1)$ is $\geqq c(x) + c(x') - 1$. Using $a_3 \geqq a_2$, (6) follows.

On the other hand, if a complete and an incomplete corners of the two squares are touching each other, then both

$$a_3 + 1 \geqq c(x) + c(x') - 2 \tag{7}$$

and

$$a_2 + 1 \geqq c(x) + c(x') - 2 \tag{8}$$

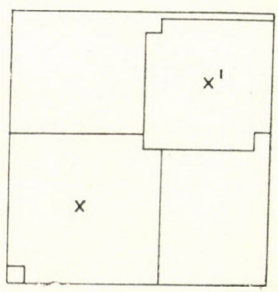


Fig. 2

hold. If (7) is satisfied with equality, then (8) results in $a_2 \geqq a_3$. Using $a_3 \geqq a_2$ the statement of this section is proved.

5. We try now count the number of points contained in the plane x from the optimal system given by the lemma. In other words, what is the number of solutions in y, z $(0 \leqq y \leqq a_2, 0 \leqq z \leqq a_3)$ of the equation

$$x+y+z = \left[\frac{a_1+a_2+a_3}{2}\right]$$

for a fixed x?

Assume

$$0 \leqq x \leqq \left[\frac{a_1+a_2-a_3}{2}\right].$$

If $y < \left[\frac{a_1+a_2-a_3}{2}\right] - x$, then obviously

$$z = \left[\frac{a_1+a_2+a_3}{2}\right] - x - y > \left[\frac{a_1+a_2+a_3}{2}\right] - \left[\frac{a_1+a_2-a_3}{2}\right] = a_3$$

holds, y, z is not a good solution. We may assume

$$\left[\frac{a_1+a_2-a_3}{2}\right] - x \leqq y \leqq a_2.$$

In this case we have always exactly one z satisfying $0 \leqq z \leqq a_3$:

$$0 \leqq a_3 - a_2 = \left[\frac{a_1+a_2+a_3}{2}\right] - \left[\frac{a_1+a_2-a_3}{2}\right] - a_2 \leqq z =$$

$$= \left[\frac{a_1+a_2+a_3}{2}\right] - x - y \leqq \left[\frac{a_1+a_2+a_3}{2}\right] - \left[\frac{a_1+a_2+a_3}{2}\right] = a_3.$$

The number of points is

$$a_2 - \left[\frac{a_1+a_2+a_3}{2}\right] + x + 1 = \left\{\frac{-a_1+a_2+a_3}{2}\right\} + x + 1$$

($\{a\}$ denotes the smallest integer $\geqq a$).

It is easy to determine the number of points in the plane x for the remaining values of x:

$$(9) \qquad c^*(x) = \begin{cases} \left\{\frac{-a_1+a_2+a_3}{2}\right\} + x + 1 & \text{if } 0 \leqq x \leqq \left[\frac{a_1+a_2-a_3}{2}\right] \\ a_2 + 1 & \text{if } \left[\frac{a_1+a_2+a_3}{2}\right] \leqq x \leqq \left[\frac{a_1-a_2+a_3}{2}\right] \\ \left[\frac{a_1+a_2+a_3}{2}\right] - x + 1 & \text{if } \left[\frac{a_1-a_2+a_3}{2}\right] \leqq x \leqq a_1. \end{cases}$$

Obviously

$$\min_{0\leqq x\leqq a_1} c^*(x) = \left[\frac{-a_1+a_2+a_3}{2}\right]+1 \tag{10}$$

holds if $a_3\leqq a_1+a_2$. If $a_3\geqq a_1+a_2$, then $c^*(x) = a_2+1$ $(0\leqq x\leqq a_1)$.

6. Let us prove the lemma for the *case* $a_1+a_2\leqq a_3$. In a fixed x-plane every column contains at most one point, hence $c(x)\leqq a_2+1$. However, in this case $c^*(x) = a_2+1$. The construction given in the lemma is the best possible in every x-plane.

7. Now, let us investigate the *case* $a_2<a_3$ $(a_1+a_2>a_3)$. We fix a_2 and a_3, and prove the statement by induction over a_1. If $a_1 = a_3-a_2$, then the statement holds by section 6. Assume now $a_1>a_3-a_2$, but the statement is true for smaller a_1. We have to prove

$$\sum_{x=0}^{a_1} c(x) \leqq \sum_{x=0}^{a_1} c^*(x). \tag{11}$$

7a. Assume there exist both increasing (say x') and decreasing (say x) x-planes. By section 4 and by $a_3>a_2$ we have

$$c(x)+c(x') \leqq a_3+2.$$

Hence either $c(x)$ or $c(x') \leqq \left[\frac{a_3+2}{2}\right]$ holds, that is, a term (say $c(x')$) in the left hand side of (11) $\leqq \left[\frac{a_3+2}{2}\right]$. Omitting this plane we obtain a new 3-dimensional parallelotope with sizes a_1-1, a_2, a_3. We can use the induction hypothesis:

$$\sum_{\substack{x=0\\ x\neq x'}}^{a_1} c(x) = \sum_{x=0}^{a_1-1} c'(x) \leqq \sum_{x=0}^{a_1-1} c^{*\prime}(x), \tag{12}$$

where $c'(x)$ is the number of points in the plane x after omitting the x'-plane, and $c^{*\prime}(x)$ is the same number for the optimal system in the new parallelotope. What is the connection between the numbers $c^*(x)$ and $c^{*\prime}(x)$? It is easy to see from (9) that $c^*(0), \ldots, c^*(a_1)$ is a sequence of integers starting with $\left\{\frac{-a_1+a_2+a_3}{2}\right\}+1$, increasing 1 by 1 up to a_2+1 (it is really larger, since $a_3<a_1+a_2$), there are a_3-a_2+1 $(\geqq 2)$ numbers a_2+1 and decreasing 1 by 1 up to $\left[\frac{-a_1+a_2+a_3}{2}\right]+1$. Hence, it is clear, that the sequence $c^{*\prime}(0), \ldots, c^{*\prime}(a_1-1)$ consists of the same numbers with the only difference that either the first or the last number is missing, depending on the parity of $-a_1+a_2+a_3$. If $-a_1+a_2+a_3$ is even, then the missing first term is $\frac{-a_1+a_2+a_3}{2}+1$ if it is odd, then the missing last term is $\left[\frac{-a_1+a_2+a_3}{2}\right]+1$. In both cases the missing term is a minimal one.

$$\sum_{x=0}^{a_1} c^*(x) = \sum_{x=0}^{a_1-1} c^{*\prime}(x)+\left[\frac{-a_1+a_2+a_3}{2}\right]+1. \tag{13}$$

However, $a_2 \geqq a_1$ results in

$$\left[\frac{a_3+2}{2}\right] \leqq \left[\frac{-a_1+a_2+a_3}{2}\right]+1, \tag{14}$$

and (11) follows from (12), $c(x') \leqq \left[\frac{a_3+2}{2}\right]$, (13) and (14).

7b. If all the x-planes are increasing or if all of them are decreasing we have to find in a different way an x' for which

$$c(x') \leqq \left[\frac{-a_1+a_2+a_3}{2}\right]+1 \tag{15}$$

holds. For sake of simplicity, let us assume, all the x-planes are decreasing.

7ba. Assume, there is a point (x_i, y_i, z_i) satisfying

$$y_i+z_i \leqq \left[\frac{-a_1+a_2+a_3}{2}\right]-1. \tag{16}$$

There are at most y_i points (x_j, y_j, z_j) such that $x_j=x_i$, $y_j<y_i$, since every column contains at most one point. On the other hand, the number of points (x_j, y_j, z_j) satisfying $x_j=x_i$, $y_j>y_i$ is at most z_i+1, because the plane x_i is decreasing, thus $z_j<z_i$ holds for all but possibly one points. Hence we have

$$c(x_i) \leqq 1+y_i+z_i+1 \leqq \left[\frac{-a_1+a_2+a_3}{2}\right]+1 \tag{17}$$

by (16). (15) is satisfied with $x'=x_i$, the proof of section 7a can be repeated here, too.

7bb. Assume, now there is a point (x_i, y_i, z_i) satisfying

$$y_i+z_i \geqq \left\{\frac{a_1+a_2+a_3}{2}\right\}. \tag{18}$$

The number of points (x_i, y_i, z_i) such that $x_j=x_i$, $y_j>y_i$ is at most a_2-y_i. On the other hand $x_j=x_i$, $y_j<y_i$ results in $z_j>z_i$ with at most one exception $(z_j<z_i)$.

7bba. If there is no exception, then the number of points of the latter type is at most a_3-z_i; the total number of points in the plane is at most $1+a_2-y_i+a_3-z_i$, or

$$c(x_i) \leqq 1+a_2-y_1+a_3-z_i \leqq \left[\frac{-a_1+a_2+a_3}{2}\right]+1 \tag{19}$$

by (18). (15) is satisfied by $x_i=x'$; this case is settled.

7bbb. If there is an exceptional point (x_j, y_j, z_j) with $x_j=x_i$, $y_j<y_i$, $z_j<z_i$, then, by section 1b, either $j=p$ or $i=q$ holds (but not both of them, since $z_p>z_q$).

7bbba. $j=p$. Then by section 1a (x_i, y_i, z_i) is the only point in the plane x_i such that $z_i>z_p$. We distinguish two cases.

7bbbaa. $z_i<a_3$. If there is at least one $z>z_i$ for which we cannot find (x_l, y_l, z_l) such that $x_l=x_i$, $z_l=z$, then the number of points satisfying $y_j<y_i$ is at most, a_3-z_i; (19) holds, again. On the other hand, if for any z there exists such an (x_l, y_l, z_l) then $z_p = a_3-1$, and $z_p<z_i=a_3$. This contradicts our supposition.

7bbbab. $z_l=a_3$. If strict equality holds in (18) then (19) follows despite the existence of the exceptional point. Thus we may suppose $y_i+z_i=\left\{\frac{a_1+a_2+a_3}{2}\right\}$. Our point has the coordinates

$$z_i=a_3,\quad y_i=\left\{\frac{a_1+a_2+a_3}{2}\right\}-a_3=\left\{\frac{a_1+a_2-a_3}{2}\right\}.$$

7bbbaba. There is a point (x_l, y_l, z_l) with $z_l=a_3$, $y_l=\left\{\frac{a_1+a_2-a_3}{2}\right\}-1$.

In this case we cannot find a point (x_k, y_k, z_k) such that $x_k=x_l$, $y_k>y_l$ $(z_k<z_l)$ because otherwise (x_i, y_i, z_i), $(x_p, y_p, z_p)=(x_j, y_j, z_j)$, (x_k, y_k, z_k) and (x_l, y_l, z_l) satisfy (5), a contradiction. On the other hand, if the point (x_k, y_k, z_k) possesses the properties $x_k=x_l$, $y_k<y_l$ and $z_k<z_l$, then using the assumption that the plane is decreasing we obtain either $k=p'$ or $l=q'$ (p' and q' are p and q of this plane). In both cases there are at most two points in the plane $x_l=x_k$, they are in increasing position, which is a contradiction. It follows $c(x_l)=1$, and (15) trivially holds for $x'=x_l$.

7bbbabb. There is no point (x_l, y_l, z_i) with $z_l=a_3$, $y_l=\left\{\frac{a_1+a_2-a_3}{2}\right\}-1$.

We have proved earlier that if there is any point satisfying (18) with strict inequality (7bbbab), or with equality but with $z_i<a_3$ (7bbbaa) then the statement is true.

The same holds, if some point satisfies (16). Thus we may suppose, that all of our points (x_r, y_r, z_r) satisfy

$$\left[\frac{-a_1+a_2+a_3}{2}\right]\leqq y_r+z_r\leqq\left\{\frac{a_1+a_2+a_3}{2}\right\}-1 \tag{20}$$

with the exception of (x_i, y_i, z_i), where $z_i=a_3$, $y_i=\left\{\frac{a_1+a_2+a_3}{2}\right\}-a_3$. If we fix a pair y_r, z_r there is at most one point (x_r, y_r, z_r) with these second and third coordinates, by the supposition of the lemma. The maximal number od points is the number of solutions of (20) in $0\leqq y_r\leqq a_2$, $0\leqq z_r\leqq a_3$, since we have an additional point (x_i, y_i, z_i) but the solution $y_r=\left\{\frac{a_1+a_2-a_3}{2}\right\}-1$, $z_r=a_3$ is omitted by the suppositions of this (7bbbabb) case. However, the number of solutions of (20) is not larger than the expected optimum of the lemma. We will see this in the case 7bc.

7bbbb. $i=q$.

7bbbba. $y_i<a_2$. The number of points in the plane is $c(x_i)=2+a_3-z_i$. Using (18) and $y_i<a_2$ we obtain

$$c(x_i)\leqq 2+a_3-\left\{\frac{a_1+a_2+a_3}{2}\right\}+y_i\leqq$$

$$\leqq 2+a_3-\left\{\frac{a_1+a_2+a_3}{2}\right\}+a_2-1=\left[\frac{-a_1+a_2+a_3}{2}\right]+1,$$

that is, (15) and the statement holds in this case.

7bbbbb. $y_i = a_2$. Like in case 7bbbab we may assume

$$y_i + z_i = \left\{\frac{a_1 + a_2 + a_3}{2}\right\}.$$

Thus,

$$z_i = \left\{\frac{a_1 + a_2 + a_3}{2}\right\} - a_2 = \left\{\frac{a_1 - a_2 + a_3}{2}\right\}.$$

7bbbbba. There is a point (x_k, y_k, z_k) with $z = \left\{\frac{a_1 - a_2 + a_3}{2}\right\} - 1$, $y_k = a_2$.

In this case we cannot find a point (x_l, y_l, z_l) such that case we cannot find a point (x_l, y_l, z_l) such that $x_l = x_k$, $z_l > z_k (y_l < y_k)$, because otherwise $(x_i, y_i, z_i) = (x_q, y_q, z_q)$, (x_j, y_j, z_j), (x_k, y_k, z_k) and (x_l, y_l, z_l) satisfy (5), which contradiction (i and j defined in 7bb and 7bbb). On the other hand, if (x_l, y_l, z_l) possesses the properties $x_l = x_k$, $z_l < z_k$, then similarly to the case 7. bbbaba we have $c(x_k) = 1$, which proves the statement.

7bbbbbb. There is no point (x_k, y_k, z_k) with

$$z_k = \left\{\frac{a_1 - a_2 + a_3}{2}\right\} - 1, \quad y_k = a_2.$$

The proof of this case is the same as in the case 7bbbabb.

7bc. All the points (x_r, y_r, z_r) satisfy (20). y_r and z_r uniquely determine x_r; it is sufficient to count the number of solutions of (20), and to see that this number $\leqq$ the number of points of the construction given in the lemma.

The number of solutions of

$$\left[\frac{-a_1 + a_2 + a_3}{2}\right] \leqq y_r + z_r \leqq \left[\frac{a_1 + a_2 + a_3}{2}\right] \tag{21}$$

is not smaller than that of (20). Define x_r' by

$$x_r' + y_r + z_r = \left[\frac{a_1 + a_2 + a_3}{2}\right]. \tag{22}$$

To every solution of (21) ($0 \leqq y_r \leqq a_2$, $0 \leqq z_r \leqq a_3$) there is a solution of (22) with $0 \leqq x_r' \leqq a_1$, and vice versa. Thus the number of solution of (20) and (21) the number of solution of (22). However, (22) defines the construction of the optimal system. The proof in case 7 is completed.

8. Let us prove the case $a_2 = a_3$. We prove this case by induction on a_1 for fixed $a_2 = a_3$. If $a_1 = 0$, the statement follows from section 6. Assume $a_1 > 0$ and suppose it is proved for smaller a_1. We can repeat the proof of case 7. The only place where we used the assumption $a_2 < a_3$ is the case 7a. There we showed that there is a plane x' satisfying $c(x') \leqq \left[\frac{a_3 + 2}{2}\right]$. If such a plane exists, we are done. We have to investigate, how can it happen that such a plane does not exist.

We have an increasing plane x and a decreasing plane x'. By section 4 we can state only $c(x)+c(x') \leqq a_3+3$ if $a_2=a_3$.

None of $c(x)$ and $c(x') \leqq \left[\frac{a_3+2}{2}\right]$ only if a_3+3 is even and $c(x)=c(x') = \frac{a_3+3}{2}$. It can happen only if in the proof of section 4 the incomplete squares touch each other at the missing point like on the figure. We have used $c(x') \leqq \left[\frac{a_3+2}{2}\right]$ in (14), only; thus if

$$\frac{a_3+3}{2} \leqq \left[\frac{-a_1+a_2+a_3}{2}\right]+1$$

holds, we are done. Using $a_2 \geqq a_1$, it does not hold only if $a_1=a_2$ $(=a_3)$. Hence, it is sufficient to investigate the case, when there are a_3+1 x-planes and every one contains at least $\frac{a_3+3}{2}$ points. They cannot have more points because each of them is increasing or decreasing, thus we can use it instead of x or x'. This shows that the number of points is $\leqq \frac{a_3+3}{2}$.

Hence

$$\sum_{x=0}^{a_3} c(x) = (a_3+1)\frac{a_3+3}{2}. \tag{23}$$

We are going to prove that $\sum_{x=0}^{a_3} c^*(x)$ is not smaller than (23). Using (9) we have

$$\sum_{x=0}^{a_3} c^*(x) = \frac{a_3+3}{2}+\frac{a_3+5}{2}+\ldots+(a_3+1)+a_3+\ldots+\frac{a_3+3}{2}+\frac{a_3+1}{2}. \tag{24}$$

The number of terms in (23) and (24) are the same. There is only one term in (24) smaller than $\frac{a_3+3}{2}$ by (1). Two terms are equal, all the other terms are $> \frac{a_3+3}{2}$. If there are at least 4 terms in (24) that is if $a_3 \geqq 3$, then (24) $\geqq$ (23). However, the statement of the lemma is trivial for $a_1=a_2=a_3=1$. The lemma is proved.

Problems

Probably, both the Theorem and the lemma can be strengthened. We only formulate a possible generalization of the lemma.

Conjecture. The conclusion of the lemma holds under the following condition weaker than (5)

$$x_i = x_j, \quad x_l = x_k$$

$$y_i = y_k, \quad y_j = y_l$$

$$z_i = z_l < z_j, z_k$$

provided (4) is true.

A further problem is to find an n-part Sperner theorem with weak enough conditions.

REFERENCES

[1] SPERNER, E.: Ein Satz über Untermengen einer endlichen Menge, *Math. Z.* **27** (1928), 544—549.
[2] KLEITMAN, D.: On a Lemma of Littlewood and Offord on the distribution of certain Sums, *Math. Z.* **90** (1965), 251—259.
[3] KATONA, G.: On a conjecture of Erdős and a stronger form of Sperner's theorem, *Studia Sci. Math. Hungar.* **1** (1966) 59—63.
[4] KATONA, G. O. H.: A generalization of some generalizations of Sperner's th. *J. Combinatorial Th.* **12** (1972), 72—81.
[5] KATONA, G. O. H.: Families of subsets having no subset containing an other one with small difference, *Nieuw Archief voor Wiskunde* (3) **20** (1972), 54—67.
[6] DE BRUIJN, N. G., VAN EBBENHORST TENGBERGEN, C. A. and KRUYSWIJK, D.: On the set of divisors of a number, *Nieuw Arch. Wiskunde* (2) **23** (1951), 191—193.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received April 20, 1973)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 391—398.

ROBBINS—MONRO PROCEDURE IN A HILBERT SPACE AND ITS APPLICATION IN THE THEORY OF LEARNING PROCESSES I.

by

P. RÉVÉSZ

Introduction

Let $Y_1, Y_2, \ldots$ be a sequence of r.v's taking values from a (separable) Hilbert space H, for which

$$\mathsf{E}(Y_1) = \mathsf{E}(Y_{k+1}|Y_k, \ldots, Y_1) = 0 \qquad (k = 1, 2, \ldots) \tag{1}$$

(with probability 1) and

$$\mathsf{P}(\|Y_k\| \leqq K) = 1 \qquad (k = 1, 2, \ldots). \tag{2}$$

Further let $M_n(x)$ be a sequence of operators (mapping H into H) for which

$$\|M_n(x)\| \leqq A\|x\| + B \qquad (A > 0;\ B \geqq 0;\ n = 1, 2, \ldots) \tag{3}$$

and

$$(x, M_n(x)) \geqq \alpha\|x\|^2 \qquad (n \geqq n_0;\ \alpha > 0) \tag{4}$$

if $\|x\| \geqq \varepsilon/2$, where ε is a given positive number, and n_0 is a positive integer (may depend on ε).

Finally let $\{a_k\}$ be a sequence of positive numbers.

Consider the sequence $\{X_n\}$ of H-valued r.v.'s defined by the recursion:

$$X_{n+1} = X_n - \frac{M_n(X_n)}{n} - \frac{a_n Y_n}{n} \qquad (n = 1, 2, \ldots) \tag{5}$$

where X_1 is an arbitrary element of H.

The sequence $\{X_n\}$ will be called a Robbins—Monro procedure (RMP). Using a very similar definition of RMP VENTER ([1]) proved that $\mathsf{P}(X_n \to 0) = 1$. In the present paper using definition (5) (and restrictions (1)—(4)) we investigate the rate of convergences of X_n to 0 in the case when $a_k = k^\beta (0 \leqq \beta < 1/2)$. In the forthcoming part of this paper the same question will be investigated in the case when condition (4) is replaced by a much weaker one.

THEOREM 1. *Under conditions* (1)—(4)

$$\mathsf{P}(\|X_n\| \geqq \varepsilon) \leqq \begin{cases} \exp\left(-Cn^{\frac{\alpha}{\alpha+\beta}}\right) & \text{if } \alpha+\beta < \frac{1}{2} \\ \exp\left(-C\dfrac{n^{2\alpha}}{\log 2n^\beta}\right) & \text{if } \alpha+\beta = \frac{1}{2} \\ \exp\left(-Cn^{1-2\beta}\right) & \text{if } \alpha+\beta > \frac{1}{2} \end{cases}$$

where $a_k=k^\beta$ $(0\leqq\beta<1/2)$, ε *is the positive number used in condition* (4) *and* C *is a positive constant (in the sense that it does not depend on* n*).*

Remark. Theorem 1 implies

$$\mathsf{P}(\|X_n\| \geqq \varepsilon) \leqq \exp(-Cn)$$

if $\beta=0$ (for any $\alpha>0$).

The proof of this theorem is based on two lemmas what will be given in § 1. The proof itself is prepared in § 2. § 3. contains two applications.

§ 1 Lemmas

LEMMA 1. *Let* $\xi_1, \xi_2, \ldots$ *be a sequence of (real valued) r.v.'s for which*

$$\mathsf{E}(\xi_1) = \mathsf{E}(\xi_{i+1}|\xi_i, \xi_{i-1}, \ldots, \xi_1) = 0 \qquad (i = 1, 2, \ldots)$$

(*with probability* 1) *and*

$$\mathsf{P}(|\xi_i| \leqq K) = 1 \qquad (i = 1, 2, \ldots;\ K > 0).$$

Then

$$\mathsf{P}\left\{\sum_{i=l}^{n} \frac{1}{i^{1-(\alpha+\beta)}}\,\xi_i \geqq \varepsilon n^\alpha\right\} = \begin{cases} \exp\left(-Ce^{\frac{\alpha}{\alpha+\beta}}\right) & \text{if } 0<\alpha+\beta<\dfrac{1}{2} \\ \exp\left(-C\dfrac{n^{2\alpha}}{\log 2n^\beta}\right) & \text{if } \alpha+\beta=\dfrac{1}{2} \\ \exp(-Cn^{1-2\beta}) & \text{if } \alpha+\beta>\dfrac{1}{2} \end{cases}$$

for any $\varepsilon>0$ *and* $l=l_n=1, 2, \ldots, n$ *where* $0<\alpha$, $0\leqq\beta<1/2$ *and* C *is a positive constant.*

PROOF. Introduce the following notations:

$$\gamma = \alpha+\beta, \quad \varkappa = \left(\frac{\varepsilon}{4}\frac{\gamma}{K}\right)^{1/\gamma}, \quad k = k_n = \max(\varkappa n^{\alpha/\gamma}, l)$$

$$\lambda_n = \begin{cases} \mu_1 n^{\alpha/\gamma-\alpha} & \text{if } \gamma<\dfrac{1}{2}, \\ \mu_2\dfrac{n^\alpha}{\log 2n^\beta} & \text{if } \gamma=\dfrac{1}{2}, \\ \mu_3 n^{1-\gamma-\beta} & \text{if } \gamma>\dfrac{1}{2}, \end{cases}$$

where

$$\mu_1 = \min\left\{\frac{\varkappa^{1-\gamma}}{K}, \frac{\varepsilon}{4}\frac{(1-2\gamma)\varkappa^{1-2\gamma}}{K^2}\right\},$$

$$\mu_2 = \begin{cases} \dfrac{\varepsilon}{8}\dfrac{\beta}{K^2(1-2\alpha)} & \text{if } \beta > 0, \\ \min\left(\dfrac{\log 2\sqrt{\varkappa}}{K}, \dfrac{\varepsilon}{4}\dfrac{\log 2}{K^2 \log\dfrac{1}{\varkappa}}\right) & \text{if } \beta = 0, \end{cases}$$

$$\mu_3 = \min\left\{\frac{1}{K}, \frac{\varkappa^{1-\gamma}}{K}, \frac{\varepsilon}{4}\frac{(2\gamma-1)}{K^2}\right\}.$$

From now on we assume $\varkappa < 1$ (we can do it without the loss of generality). Then we have

$$\left|\sum_{i=l}^{k-1}\frac{1}{i^{1-\gamma}}\xi_i\right| \leqq \begin{cases} \dfrac{K}{\gamma}k^\gamma = \dfrac{K}{\gamma}(\varkappa n^{\alpha/\gamma})^\gamma = \dfrac{\varepsilon}{4}n^\alpha & \text{if } l \leqq \varkappa n^{\alpha/\gamma} \\ 0 & \text{if } l > \varkappa n^{\alpha/\gamma}, \end{cases}$$

hence

$$\text{(6)} \qquad \mathsf{P}\left\{\sum_{i=l}^{n}\frac{1}{i^{1-\gamma}}\xi_i \geqq \varepsilon n^\alpha\right\} \leqq \mathsf{P}\left\{\sum_{i=k}^{n}\frac{1}{i^{1-\gamma}}\xi_i \geqq \frac{\varepsilon}{2}n^\alpha\right\} =$$

$$= \mathsf{P}\left\{\exp\left(\lambda_n\sum_{i=k}^{n}\frac{1}{i^{1-\gamma}}\xi_i\right) \geqq \exp\left(\frac{\varepsilon}{2}\lambda_n n^\alpha\right)\right\} \leqq$$

$$\leqq \exp\left(-\frac{\varepsilon}{2}\lambda_n n^\alpha\right)\mathsf{E}\left(\exp\left(\lambda_n\sum_{i=k}^{n}\frac{1}{i^{1-\gamma}}\xi_i\right)\right).$$

Since (if $k \leqq i \leqq n$ and n is sufficiently large)

$$\lambda_n\frac{K}{i^{1-\gamma}} \leqq \begin{cases} \lambda_n\dfrac{K}{k^{1-\gamma}} \leqq \mu_1 n^{\alpha/\gamma-\alpha}\dfrac{K}{(\varkappa n^{\alpha/\gamma})^{1-\gamma}} = \dfrac{\mu_1 K}{\varkappa^{1-\gamma}} \leqq 1 & \text{if } \gamma < \dfrac{1}{2} \\ \lambda_n\dfrac{K}{k^{1-\gamma}} \leqq \mu_2\dfrac{n^\alpha}{\log 2n^\beta}\dfrac{K}{(\varkappa n^{\alpha/\gamma})^{1-\gamma}} = \dfrac{\mu_2 K}{\sqrt{\varkappa}\log 2n^\beta} \leqq 1 & \text{if } \gamma = \dfrac{1}{2} \\ \lambda_n\dfrac{K}{k^{1-\gamma}} \leqq \mu_3 n^{1-\gamma-\beta}\dfrac{K}{(\varkappa n^{\alpha/\gamma})^{1-\gamma}} \leqq 1 & \text{if } \dfrac{1}{2} < \gamma \leqq 1 \\ \lambda_n K n^{\gamma-1} \leqq \mu_3 n^{1-\gamma-\beta}Kn^{\gamma-1} = \mu_3 K n^{-\beta} \leqq 1 & \text{if } \gamma > 1, \end{cases}$$

we have

$$(7)\qquad \mathsf{E}\left(\exp\left(\lambda_n \sum_{i=k}^{n} \frac{1}{i^{1-\gamma}}\,\xi_i\right)\right) \leqq \mathsf{E}\left(\prod_{i=k}^{n}\left(1+\frac{\lambda_n}{i^{1-\gamma}}\,\xi_i+\frac{\lambda_n^2}{i^{2(1-\gamma)}}\,\xi_i^2\right)\right) \leqq$$

$$\leqq \prod_{i=k}^{n}\left(1+\frac{\lambda_n^2 K^2}{i^{2(1-\gamma)}}\right) \leqq \exp\left(\lambda_n^2 K^2 \sum_{i=k}^{n} \frac{1}{i^{2(1-\gamma)}}\right) \leqq$$

$$\leqq \begin{cases} \exp\left(\lambda_n^2 K^2 \dfrac{1}{1-2\gamma}\,\dfrac{1}{k^{1-2\gamma}}\right) \leqq \exp\left(\dfrac{\varepsilon}{4}\,\lambda_n n^{\alpha}\right) & \text{if } \gamma < \dfrac{1}{2} \\ \exp\left(\lambda_n^2 K^2 \log\dfrac{n}{k}\right) \leqq \exp\left(\dfrac{\varepsilon}{4}\,\lambda_n n^{\alpha}\right) & \text{if } \gamma = \dfrac{1}{2} \\ \exp\left(\lambda_n^2 K^2 \dfrac{n^{2\gamma-1}}{2\gamma-1}\right) \leqq \exp\left(\dfrac{\varepsilon}{4}\,\lambda_n n^{\alpha}\right) & \text{if } \gamma > \dfrac{1}{2}. \end{cases}$$

(6) and (7) clearly prove our Lemma 1.

LEMMA 2. *Under the conditions of Theorem* 1 *we have*

$$\|X_{n+1}\| \leqq \|X_n\|\left(1-\frac{\alpha}{n}+\frac{A^2+\dfrac{4AB}{\varepsilon}}{2n^2}\right)-\frac{a_n}{n}\left(\frac{X_n}{\|X_n\|},\ Y_n\right)+O\left(\frac{a_n^2}{n^2}\right)$$

provided that $\|X_n\| \geqq \dfrac{\varepsilon}{2}$ *and* n *is sufficiently large.*

PROOF. Clearly we have

$$\|X_{n+1}\|^2 = \|X_n\|^2+\frac{1}{n^2}\,\|M_n(X_n)\|^2+\frac{a_n^2}{n^2}\,\|Y_n\|^2-\frac{2}{n}\,(X_n,\ M_n(X_n))-$$

$$-\frac{2a_n}{n}\,(X_n,\ Y_n)+\frac{2a_n}{n^2}\,(M_n(X_n),\ Y_n) \leqq \|X_n\|^2+$$

$$+\frac{1}{n^2}\,(A^2\|X_n\|^2+2AB\|X_n\|+B^2)+\frac{a_n^2}{n^2}\,K^2-\frac{2}{n}\,(X_n,\ M_n(X_n))-$$

$$-\frac{2a_n}{n}\,(X_n,\ Y_n)+\frac{2a_n}{n^2}\,(M_n(X_n),\ Y_n) \leqq$$

$$\leqq \|X_n\|^2\left(1-\frac{2\alpha}{n}+\frac{A^2}{n^2}+\frac{4AB}{\varepsilon n^2}\right)+\frac{B^2+a_n^2K^2}{n^2}-\frac{2a_n}{n}\,(X_n,\ Y_n)+\frac{2a_n}{n^2}\,(M_n(X_n),\ Y_n) =$$

$$= \|X_n\|^2\left\{1-\frac{2\alpha}{n}+\frac{A^2+\dfrac{4AB}{\varepsilon}}{n^2}+\frac{B^2+a_n^2K^2}{n^2\|X_n\|^2}-\frac{2a_n}{n}\,\frac{(X_n,\ Y_n)}{\|X_n\|^2}+\frac{2a_n}{n^2}\,\frac{(M_n(X_n),\ Y_n)}{\|X_n\|^2}\right\} \leqq$$

$$\leqq \|X_n\|^2\left\{1-\frac{\alpha}{n}+\frac{A^2+\dfrac{4AB}{\varepsilon}}{2n^2}-\frac{a_n}{n}\,\frac{(X_n,\ Y_n)}{\|X_n\|^2}+\frac{B^2+a_n^2K^2}{\varepsilon n^2\|X_n\|}+\frac{a_n}{n^2}\,\frac{K}{\|X_n\|}\left(A+\frac{2B}{\varepsilon}\right)\right\}^2,$$

hence we have our Lemma 2.

§ 2. Proof of theorem 1.

Introduce the following notations:

$$\eta_n = -\left(\frac{X_n}{\|X_n\|},\ Y_n\right)$$

$$\mathfrak{A}_k^{(n)} = \mathfrak{A}_k = \left\{\|X_k\| \leqq \frac{\varepsilon}{2},\ \|X_{k+1}\| > \frac{\varepsilon}{2},\ \|X_{k+2}\| > \frac{\varepsilon}{2},\ \ldots,\ \|X_{n-1}\| > \frac{\varepsilon}{2},\ \|X_n\| > \varepsilon\right\}$$

$$(k = 1, 2, \ldots, n-1)$$

and

$$\mathfrak{A}_0 = \left\{\|X_1\| > \frac{\varepsilon}{2},\ \|X_2\| > \frac{\varepsilon}{2},\ \ldots,\ \|X_{n-1}\| > \frac{\varepsilon}{2},\ \|X_n\| > \varepsilon\right\}.$$

Then clearly we have

$$\mathsf{P}(\|X_n\| > \varepsilon) = \sum_{k=0}^{n-1} \mathsf{P}(\mathfrak{A}_k) \tag{8}$$

and by Lemma 2

$$\|X_{n+m+1}\| \leqq \|X_n\| \prod_{k=n}^{n+m}\left(1 - \frac{\alpha}{k} + \frac{C}{k^2}\right) + \sum_{k=n}^{n+m} \frac{a_k}{k}\eta_k \prod_{j=k+1}^{n+m}\left(1 - \frac{\alpha}{j} + \frac{C}{j^2}\right) + o_n(1)$$

where

$$C = \frac{1}{2}\left(A^2 + \frac{4AB}{\varepsilon}\right)$$

i.e.

$$(n+m)^\alpha \|X_{n+m+1}\| \leqq [(n+1)^\alpha + C(n+1)^{\alpha-1}]\|X_n\| + \sum_{k=n}^{n+m} \frac{a_k}{k^{1-\alpha}}\eta_k + o\big((n+m)^\alpha\big) \tag{9}$$

provided that

$$\|X_n\| \geqq \frac{\varepsilon}{2},\ \|X_{n+1}\| \geqq \frac{\varepsilon}{2},\ \ldots,\ \|X_{n+m}\| \geqq \frac{\varepsilon}{2}.$$

Hence the probability of the event $\mathfrak{A}_l$ ($l=0, 1, 2, \ldots, n-1$) can be estimated as follows. By (9) and by Lemma 1 we have

$$\mathsf{P}(\mathfrak{A}_l) \leqq \mathsf{P}\left\{\sum_{k=l}^{n} \frac{a_k}{k^{1-\alpha}}\eta_k \geqq \frac{\varepsilon}{4}n^\alpha\right\} \leqq \tag{10}$$

$$\leqq \begin{cases} \exp\left(-Cn^{\frac{\alpha}{\alpha+\beta}}\right) & \text{if } \alpha+\beta < \frac{1}{2} \\ \exp\left(-C\dfrac{n^{2\alpha}}{\log 2n^\beta}\right) & \text{if } \alpha+\beta = \frac{1}{2} \\ \exp\left(-Cn^{1-2\beta}\right) & \text{if } \alpha+\beta > \frac{1}{2} \end{cases}$$

and our Theorem 1 follows from (8) and (10).

§ 3. Applications

a) *Density estimation.*

Let $\xi_1, \xi_2, \ldots$ be a sequence of i.i.d.r.v's with density function $f(x)$. C. T. WOLWERTON and T. I. WAGNER ([2]) gave the following definition of the empirical density function

$$f_n(x) = \frac{1}{n}\sum_{k=1}^{n}\frac{1}{a_k}h\left(\frac{x-\xi_k}{a_k}\right) = \frac{n-1}{n}f_{n-1}(x)+\frac{1}{na_n}h\left(\frac{x-\xi_n}{a_n}\right) \tag{11}$$

where $h(x)$ is an arbitrary density function and $\{a_n\}$ is a monotonically decreasing sequence of real numbers going to 0.

In [3] under some further conditions on the smoothness of f and h and on the behaviour of $\{a_n\}$ we have proved that

$$\mathsf{P}\left\{\int_{-\infty}^{+\infty}(f_n-f)^2\,dx \geqq \varepsilon\right\} \leqq e^{-Cna_n}. \tag{12}$$

At present we prove (12) as a simple consequence of Theorem 1. More precisely we prove:

THEOREM 2. *Suppose that f and h are square integrable density functions and $a_n=n^{-2\beta}$ $(0<\beta<1/2)$. Then* (12) *holds.*

PROOF. Introduce the notations: $X_1\equiv 0$,

$$X_n = X_n(x) = f_n(x)-f(x) \qquad (n = 2, 3, \ldots),$$

$$M_n(X) = \left[X+f-\frac{1}{a_{n+1}}\int h\left(\frac{x-y}{a_{n+1}}\right)f(y)\,dy\right]\frac{n}{n+1},$$

$$Y_n = -\frac{n}{(n+1)}\frac{1}{n^\beta a_{n+1}}\left[h\left(\frac{x-\xi_{n+1}}{a_{n+1}}\right)-\int h\left(\frac{x-y}{a_{n+1}}\right)f(y)\,dy\right].$$

Then by (11) we have

$$X_{n+1} = X_n-\frac{M_n(X_n)}{n}-\frac{n^\beta Y_n}{n} \qquad (n = 1, 2, \ldots).$$

Hence we get our Theorem if we prove that conditions (1)—(4) hold. One can get (1)—(4) (with $\alpha = 1-\delta$, $\delta>1$) using the properties of singular integrals (see e.g. [4] Theorem 1.3.2). The details will be omitted.

b) *Function learning*

Let $f(x)$ be a bounded function defined in [0, 1] (the function what we want to learn), let $g(x)$ be a density function on [0, 1] for which

$$0 < a \leqq g(x) \leqq b < +\infty$$

and let $\zeta_1=(\xi_1, \eta_1), \ldots, \zeta_n=(\xi_n, \eta_n)$ be a sequence of independent vector valued r.v's for which

$$\mathsf{P}(\xi_i < x) = \int_0^x g(t)\,dt,$$

$$\mathsf{E}(\eta_i|\xi_i) = f(\xi_i),$$

$$\mathsf{P}(|\eta_i - f(\xi_i)| \leqq K) = 1.$$

Our aim is to estimate the function $f(x)$ using the learning points $\zeta_1, \zeta_2, \ldots, \zeta_n$. We propose to estimate f by the sequence of functions f_n defined by the recursion

$$f_1(x) \equiv 0,$$

$$f_{n+1}(x) = f_n(x) + h\left(\frac{x-\xi_n}{a_n}\right)(\eta_n - f_n(x))\frac{1}{na_n} \tag{13}$$

where $h(x)$ is a bounded density function and

$$a_n = \frac{1}{n^{2\beta}} \qquad \left(0 < \beta < \frac{1}{2}\right).$$

We prove the following

THEOREM 3.

$$\mathsf{P}\left\{\int_{-\infty}^{+\infty} (f_n - f)^2\,dx \geqq \varepsilon\right\} \leqq e^{-Cn^{1-2\beta}} \qquad (C > 0) \tag{14}$$

provided that $a > \frac{1}{2}$.

PROOF. Introduce the notations

$$X_n = f_n(x) - f(x),$$

$$M_n(X) = \frac{1}{a_n}\left[X\int h\left(\frac{x-y}{a_n}\right)g(y)\,dy - \int h\left(\frac{x-y}{a_n}\right)(f(y)-f(x))g(y)\,dy\right]$$

$$Y_n = \frac{X_n}{\sqrt{a_n}}\left[h\left(\frac{x-\xi_n}{a_n}\right) - \int h\left(\frac{x-y}{a_n}\right)g(y)\,dy\right] -$$

$$-\frac{1}{\sqrt{a_n}}\left[h\left(\frac{x-\xi_n}{a_n}\right)(\eta_n - f(x)) - \int h\left(\frac{x-y}{a_n}\right)(f(y)-f(x))g(y)\,dy\right]$$

Then (13) can be written in the following from

$$X_{n+1} = X_n - \frac{M_n(X_n)}{n} - \frac{n^\beta Y_n}{n}.$$

Using again the properties of the singular integrals and the fact that $|X_n| \leqq K$ one can prove that (1)—(4) hold with $\alpha = a - \delta$ ($\delta > 0$). Hence Theorem 1 implies (14).

REFERENCES

[1] VENTER, I. H.: On Dvoretzky stochastic approximation theorems. *Ann. Math. Statist.* **37** (1966) 1534—1544.
[2] WOLVERTON, C. T.—WAGNER, T. I.: Asymptotically optimal discriminant functions for pattern classification. *IEEE Transactions on Information Theory* TI—**15** (169) 258—266.
[3] REJTŐ, L.—RÉVÉSZ, P.: Density estimation and pattern classification. *Problems of Control and Information theory*, **2** (1973) 67—80.
[4] BOCHNER, S.: *Harmonic analysis and the theory of probability*. Univ. of Calif. Press Berkeley 1955.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received April 26, 1973)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 399—404.

ON THE L_p-NORMS OF ORTHONORMAL HERMITE FUNCTIONS

by

G. FREUD and G. NÉMETH

1. Introduction

On the Conference on Theory of Approximation, held in Poznan, August 22—26 1972 Professor Z. CIESIELSKI raised the following problem (see [1], Problem 12): "Let $\mathbf{H}_n(x)$ denote the Hermite function normalized in $L_2(R)$. Show that

$$\sup_{x \in R} |\mathbf{H}_n(x)| \geqq \sup_{x \in R} |\mathbf{H}_{n+1}(x)| \quad \text{for all} \quad n.$$

Similar question about the monotonicity of the $L_p(R)$ norm of $\mathbf{H}_n$ as a function of n."

In our present paper we give an affirmative answer to the second part of CIESIELSKI's problem for the case $p=4$. Actually we prove the more precise result that the forth power of the corresponding sequence is totally monotone. Our result raises the problem whether the same would be valid for p-th powers of L_p norms for every $2 \leqq p \leqq \infty$. Unfortunately our method does not apply for general p though it might be generalized to finite even integer values of p.

2. Results

Let us fix our notations. Taking $\{H_n(x)\}$ as the sequence of Hermite orthogonal polynomials with the usual normalization, the orthonormal Hermite functions are

$$\mathbf{H}_n(x) \equiv \chi_n(x) = \left(2^n n! \sqrt{\pi}\right)^{-1/2} H_n(x) e^{-x^2/2} \tag{1}$$

Consequently the special case $p=4$ of CIESIELSKI's conjectures states that the sequence

$$\sigma_n = \int_{-\infty}^{\infty} \chi_n^4(x)\, dx \tag{2}$$

is decreasing. We prove the more general

THEOREM 1. *The sequence* $\{\sigma_n\}$ *is totally monotone.*

By a well known theorem of F. HAUSDORFF [3] a sequence is totally monotone iff it is the moment sequence of a nonnegative measure the support of which lies

in [0, 1]. As a matter of fact we are going to prove Theorem 1 through an explicite construction of this nonnegative measure:

THEOREM 2. *We have*

$$\sigma_n = 2^{-1/2}\pi^{-3/2}\int_0^1 t^n(1-t)^{-1/2}\,{}_2F_1\left(\frac{1}{2},\ \frac{1}{2};\ 1;\ t\right)dt \tag{3}$$

where

$${}_2F_1\left(\frac{1}{2},\ \frac{1}{2};\ 1;\ t\right) = \sum_{j=0}^{\infty}\left[\frac{\Gamma\left(j+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)j!}\right]^2 t^j \tag{4}$$

We see from (4) that the measure

$$d\mu(t) = (1-t)^{-1/2}\,{}_2F_1\left(\frac{1}{2},\ \frac{1}{2};\ 1;\ t\right)dt$$

is nonnegative. It is also finite: the simplest way to show this is by applying (3) for $n=0$. Consequently Theorem 1 is a consequence of Theorem 2.

3. Lemmata

LEMMA 1. *The double sequence*

$$\eta_{nm} = \int_{-\infty}^{\infty} \chi_n^2(x)\,\chi_m^2(x)\,dx \tag{5}$$

generates the double power series

$$\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\eta_{mn}u^m v^n = G(u,v) = \frac{1}{\sqrt{2\pi}}[(1-u)(1-v)(1-uv)]^{-1/2}. \tag{6}$$

PROOF. By (1) and Mehler's generator series (see e.g. A. ERDÉLYI [2]) we have

$$\sqrt{\pi}\sum_{n=0}^{\infty}\chi_n^2(x)z^n = e^{-x^2}\sum_{n=0}^{\infty}\frac{H_n^2(x)}{2^n n!}z^n = \tag{7}$$

$$= e^{-x^2}(1-z^2)^{-1/2}\exp\left\{\frac{2z}{1+z}x^2\right\} = (1-z^2)^{-1/2}\exp\left\{-\frac{1-z}{1+z}x^2\right\}$$

and from this follows

$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\eta_{mn}u^m v^n = \tag{8}$$

$$= \frac{1}{\pi}(1-u^2)^{-1/2}(1-v^2)^{-1/2}\int_{-\infty}^{\infty}\exp\left\{-\left(\frac{1-u}{1+u}+\frac{1-v}{1+v}\right)x^2\right\}dx$$

and from (8) we obtain (6) by elementary calculation, q.e.d.

Taking the binomial series

$$(1-z)^{-1/2} = \sum_{\nu=0}^{\infty}\frac{\Gamma\left(\nu+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\nu!}z^{\nu}$$

in consideration, we can expand (6) in

$$G(u,v) = \frac{1}{\sqrt{2\pi}}\sum_{\lambda=0}^{\infty}\sum_{\mu=0}^{\infty}\sum_{\nu=0}^{\infty}\frac{\Gamma\left(\lambda+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\lambda!}\frac{\Gamma\left(\mu+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\mu!}\frac{\Gamma\left(\nu+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\nu!}\cdot(uv)^{\lambda}u^{\mu}v^{\nu}. \tag{9}$$

Comparing the coefficients of (6) and (9) we obtain from (2) and (5)

$$\sigma_n = \eta_{nn} = \frac{1}{\sqrt{2\pi}}\sum_{j=0}^{n}\left[\frac{\Gamma\left(j+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)j!}\right]^2\frac{\Gamma\left(n-j+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)(n-j)!}. \tag{10}$$

LEMMA 2. *We have*

$$\sigma_n = 2^{-1/2}\pi^{-3/2}\frac{(-1)^n}{n!}\int_0^1\omega^n(1-\omega)^{-1/2}\left[\left(\frac{d}{d\omega}\right)^n\omega^{-1/2}(1-\omega)^n\right]d\omega. \tag{11}$$

PROOF. By Leibnitz's formula

$$\left(\frac{d}{d\omega}\right)^n[\omega^{-1/2}(1-\omega)^n] = \tag{12}$$

$$= \sum_{j=0}^{n}\frac{n!}{j!(n-j)!}\left[\left(\frac{d}{d\omega}\right)^j\omega^{-1/2}\right]\left[\left(\frac{d}{d\omega}\right)^{n-j}(1-\omega)^n\right] =$$

$$= \frac{(-1)^n(n!)^2}{\Gamma\left(\frac{1}{2}\right)}\sum_{j=0}^{n}\frac{\Gamma\left(j+\frac{1}{2}\right)}{(j!)^2(n-j)!}\omega^{-(1/2)-j}(1-\omega)^j.$$

By integrating we get from (12)

$$(13)\qquad \frac{(-1)^n}{n!}\int_0^1 \omega^n(1-\omega)^{-1/2}\left(\frac{d}{d\omega}\right)^n[\omega^{-1/2}(1-\omega)^n]\,d\omega =$$

$$= \frac{n!}{\Gamma\left(\frac{1}{2}\right)}\sum_{j=0}^{n}\frac{\Gamma\left(j+\frac{1}{2}\right)}{(j!)^2(n-j)!}\int_0^1 \omega^{n-j-(1/2)}(1-\omega)^{j-(1/2)}\,d\omega =$$

$$= \frac{n!}{\Gamma\left(\frac{1}{2}\right)}\sum_{j=0}^{n}\frac{\Gamma\left(j+\frac{1}{2}\right)}{(j!)^2(n-j)!}\,\frac{\Gamma\left(n-j+\frac{1}{2}\right)\Gamma\left(j+\frac{1}{2}\right)}{\Gamma(n+1)} =$$

$$= \left[\Gamma\left(\frac{1}{2}\right)\right]^2\sum_{j=0}^{n}\left[\frac{\Gamma\left(j+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)j!}\right]^2\frac{\Gamma\left(n-j+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)(n-j)!}$$

From (10) and (15) follows (11). Q.e.d.

4. Proof of Theorem 2.

In the power series

$$(14)\qquad (1-\omega)^{-1/2} = \sum_{\nu=0}^{\infty}\frac{\Gamma\left(\nu+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\nu!}\omega^\nu$$

all terms are positive for $0<\omega<1$. Consequently the partial sums of (14) have the majorant $(1-\omega)^{-1/2}$. Clearly

$$(15)\qquad \left|\omega^n\left\{\left(\frac{d}{d\omega}\right)^n[\omega^{-1/2}(1-\omega)^n]\right\}\right| \leqq c_n\omega^{-1/2} \qquad (0<\omega<1).$$

By (14) and (15) and by Lebesgue's dominated convergence theorem we infer from (11)

$$(16)\quad \sigma_n = 2^{-1/2}\pi^{-3/2}\frac{(-1)^n}{n!}\sum_{j=0}^{\infty}\frac{\Gamma\left(j+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)j!}\int_0^1 \omega^{n+j}\left\{\left(\frac{d}{d\omega}\right)^n[\omega^{-1/2}(1-\omega)^n]\right\}d\omega.$$

By partial integration we get

$$\int_0^1 \omega^{n+j}\left\{\left(\frac{d}{d\omega}\right)^n [\omega^{-1/2}(1-\omega)^n]\right\} d\omega = \tag{17}$$

$$= (-1)^n \frac{(n+j)!}{j!} \int_0^1 \omega^{j-(1/2)}(1-\omega)^n d\omega = (-1)^n \frac{(n+j)!}{j!} \frac{\Gamma\left(j+\frac{1}{2}\right) n!}{\Gamma\left(n+j+\frac{3}{2}\right)}.$$

From (16) and (17)

$$\sigma_n = 2^{-1/2}\pi^{-3/2} \sum_{j=0}^{\infty} \left[\frac{\Gamma\left(j+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right) j!}\right]^2 \frac{(n+j)!\,\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(n+j+\frac{3}{2}\right)} \tag{18}$$

Inserting (4) in (3) and integrating term by term the resulting series with positive terms we obtain (18). This proves Theorem 2, q.e.d.

5. Asymptotic formula for σ_n

By Stirling's asymptotic series we have

$$\frac{\Gamma\left(j+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right) j!} = \frac{1}{\sqrt{\pi j}} + O(j^{-3/2}) \tag{19}$$

so that by (4) we have uniformly in $t \in [0, 1)$

$${}_2F_1\left(\frac{1}{2}, \frac{1}{2}; 1; t\right) = \frac{1}{\pi} \log \frac{1}{1-t} + O(1). \tag{20}$$

By (3)

$$\sigma_n = 2^{-1/2}\pi^{-5/2} \int_0^1 t^n (1-t)^{-1/2} \log \frac{1}{1-t}\, dt + O(1) \int_0^1 t^n (1-t)^{-1/2}\, dt = \tag{21}$$

$$= 2^{-1/2}\pi^{-5/2} \int_0^1 t^n (1-t)^{-1/2} \log \frac{1}{1-t}\, dt + O(n^{-1/2}).$$

To evaluate the last integral we differentiate with respect to α the formula

$$\int_0^1 t^n(1-t)^{\alpha-1}\,dt = \frac{n!\,\Gamma(\alpha)}{\Gamma(n+\alpha+1)}$$

We obtain

$$\int_0^1 t^n(1-t)^{\alpha-1}\log(1-t)\,dt = \frac{n!\,\Gamma(\alpha)}{\Gamma(n+\alpha+1)}[\psi(\alpha)-\psi(n+\alpha)] \tag{22}$$

where $\psi(z)=\Gamma'(z)/\Gamma(z)$. Referring to the fact that $\psi(z)$ satisfies the functional equation

$$\psi(z+1) = z^{-1}+\psi(z)$$

(see e.g. N. N. LEBEDEV [4]) we get

$$\int_0^1 t^n(1-t)^{\alpha-1}\log\frac{1}{1-t} = \frac{n!\,\Gamma(\alpha)}{\Gamma(n+\alpha+1)}\left[\frac{1}{\alpha}+\frac{1}{\alpha+1}+\ldots+\frac{1}{\alpha+(n-1)}\right] = \tag{23}$$

$$= \Gamma(\alpha)\frac{\log n}{n^{\alpha}}+O(n^{-\alpha})$$

Inserting $\alpha=1/2$ in (23) we obtain from (21)

$$\sigma_n = 2^{-1/2}\pi^{-2}n^{-1/2}\log n+O(n^{-1/2}). \tag{24}$$

REFERENCES

[1] CIESIELSKI, Z.: Conference Theory of Approximation, *Problems*. Poznan 22—26 August 1972 p. 5.

[2] ERDÉLYI, A.: Über eine erzeugende Funktion von Produkten Hermitescher Polynome. *Math, Zeitschr.* **44** (1938), 201—211.

[3] HAUSDORFF, F.: Momentenprobleme für ein endliches Intervall. *Math. Zeitschr.* **16** (1923), 220—248.

[4] Лебедев, Н. Н.: *Специальные функции и их приложения*. 2. изд,. Москва, 1963.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received April 28, 1973)

Studia Sciantiarum Mathematicarum Hungarica 8 (1973) 405—409.

AN EXTENSION OF THE CLASS OF PERFECT GRAPHS

by

J. KÖRNER

Summary

A new class of graphs containing all the perfect graphs is introduced. This class is closed under normal product and contains the complement of all its graphs. Many not perfect graphs, e.g. any odd cycle of length at least 9 belong to it.

Introduction

Perfect graph is a concept of CLAÙDE BERGE [1] having some connections with a well-known information-theoretical problem of C. E. SHANNON [4]. In a joint paper with G. LONGO [2] we considered an other information-theoretical problem which has led us to the concept of normal graphs. Some analogies of the two problems suggested to ask whether perfect graphs have the normal property. In this paper we shall answer this question in the affirmative. After proving some simple theorems on normal graphs we shall state an interesting open problem. This problem is connected with a certain kind of graph multiplication, first introduced by SHANNON in [4], and called "normal" in BERGE's monograph [1]. This product is relevant to the connections existing between perfect and normal graphs, because in general the third normal power of a perfect graph is no more perfect, while all the normal powers of a normal graph, and in particular, all the normal powers of a perfect graph still remain normal.

The treatment of the subject makes no use and requires no knowledge of information theory. For those interested in this background of the present paper a motivation is given in [2].

Definitions. Perfect graphs and the class of normal graphs

We shall adopt the terminological conventions of BERGE [1]. Let us recall all the same some standard definitions. By a *graph* we shall always mean indirected finite graphs without loops and multiple edges. A subgraph is called *complete* if any pair of its distinct vertices is joined by an edge. A subgraph is called an *independent set*, if no pair of its distinct vertices is joined by an edge. ("Subgraph" means induced subgraphs only.) Let $\gamma(G)$ denote the chromatic number of the graph G. $\omega(G)$ denotes the maximum number of vertices which constitute a complete subgraph of G. $\omega(G)$ is called the *clique number* of G. Remark that $\omega(G) \leqq \gamma(G)$.

Definition 1. G is a *perfect graph*, if for every subgraph $G' \subseteq G$

$$\gamma(G') = \omega(G').$$

In [2] we introduced the following concept:

Definition 2. G is a *normal graph*, if there exist a covering of its vertex set by complete subgraphs of G and another one by independent sets of G with the property that any complete subgraph of the former has a non-void intersection with all the independent sets of the latter.

One observes immediately that if G is normal, so is $\bar{G}$.

We can formulate now

PROPOSITION 1. *Every perfect graph is normal.*

The PROOF of this fact is based on a theorem of LOVÁSZ (see Theorem 1 of [3]) which served him to prove that BERGE's perfect graph conjecture is true. We shall use an operation on graphs called *substitution* in [3]. This is the following: let G and H denote two vertex-disjoint graphs. The *substitution of a vertex* $x \in G$ *by* H means deleting x and joining every vertex of H with all those vertices of G which have been adjacent with x. Now LOVÁSZ's theorem says that a graph resulting from a perfect graph by the substitution of a vertex by another perfect graph is again perfect.

1. Let K be a graph with the following two properties:

A: $\omega(K) = \gamma(K)$.

B: K contains complete subgraphs of cardinality $\gamma(K)$ the union of which covers the vertex set of K.

First we remark that such a graph K is normal. Actually, consider any colouring of the vertices of K with $\gamma(K)$ different colours. Obviously, the vertices of a same colour form an independent set and the complete subgraphs of cardinality $\omega(K) = \gamma(K)$ intersect every colour class. However both the colour classes and the complete subgraphs of cardinality $\gamma(K)$ cover the vertex set of K.

2. Let G be now any perfect graph. Then we especially have $\gamma(G) = \omega(G)$. We already know that if the greatest complete subgraphs covered the whole vertex set of G, then G would be normal. Therefore we may suppose that G is not of this type. Consider now those complete subgraphs of G which do not have maximum cardinality, and pick out one which has greatest cardinality among *these* subgraphs. (It may contain at most $\gamma(G) - 1$ vertices.) Choose any vertex of this complete subgraph and substitute it by an edge, i.e. by a subgraph H consisting of two adjacent vertices. By the theorem of LOVÁSZ the resulting graph G' is perfect, too.

3. We can continue iterating on G' the procedure of substitution described in 2. until we arrive at a perfect graph K yielding $\gamma(K) = \omega(K) = \omega(G)$ and having property **B**. In fact, every graph resulting from G by successive substitutions is perfect and has the same clique number as G. The iteration stops only arriving at a graph no more vertex of which is left uncovered by the union of its greatest complete subgraphs. It is obvious that such a procedure must always stop in a finite number of steps.

Let us fix any coloration of the vertices of K with $\gamma(G) = \gamma(K)$ colours. For any vertex of K we can decide from which vertex of G it resulted by successive substitutions. We shall construct a covering of G by its — not necessarily disjoint —

independent sets in the following way: to a vertex x of G we assign all the colours which occur in the colouring of any vertex of K resulted from x by successive substitutions. Vertices of a same colour form an independent set of our covering. To a greatest complete subgraph of K we make correspond that complete subgraph of G from which it vertices have resulted by successive substitutions. Since a vertex of G has obtained all the colours of any vertex which resulted from it, the so defined complete subgraphs of G intersect every colour class, and therefore we have obtained the two coverings requested in order to establish the normality of G.

Example: A graph which does not have the normal property: It is easy to see that the odd cycle of length 5 (see Fig. 1) is not normal. In fact, if the covering by complete subgraphs contained the edge (a, b), then the covering by independent sets could not contain the pair $\{c, e\}$. Different edges exclude different pairs of non-adjacent vertices. However at least 3 edges and 3 non-adjacent pairs are necessary to cover the vertex set, and since there are only 5 different pairs of non-adjacent elements in the whole graph, it therefore can not have the normal property.

We remark that—as it is rather easy to check—neither the odd cycle of length 7 is normal. Therefore, the following statement, though its proof is immediate, may surprise:

PROPOSITION 2. *Every odd cycle of length at least* 9 *is normal.*

We remark that every proper subgraph of an odd cycle is perfect, and therefore, odd cycles of length at least 9 are examples of graphs any subgraph of which is normal. In other words Proposition 2 implies that *neither the normality of all the subgraphs involves the perfection of a graph.*

PROOF. We construct first the two coverings requested for normality for the case of the odd cycle of length 9. The construction can be read from Fig. 2. Colour x by red (marked by R in Fig. 2) and blue, y by blue and green, whereas

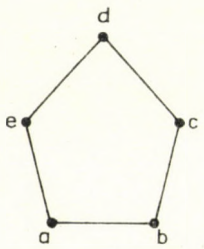


Fig. 1

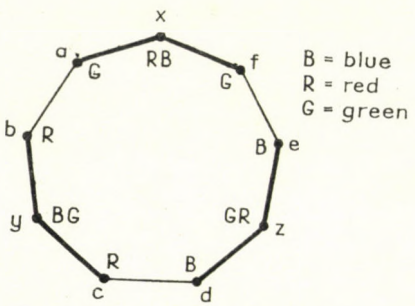


Fig. 2

z gets the colours green and red. The remaining vertices are adjacent with exactly one of these three. Therefore the colour they receive is now uniquely determined, when we use only the former 3 colours. The covering by complete subgraphs consists of all the edges which are adjacent with a doubly-coloured vertex of Fig. 2. (These are the bold edges of Fig. 2).

Let C denote the odd cycle of length $2k+1$ ($k \geqq 5$). Pick out 10 consecutive vertices of C and colour them as shown in Fig. 3. (Consecutive is used here in the sense of any transitive orientation of the edges of the cycle.)

(After making coincide the two endpoints of the chain of Fig. 3 this clouring becomes exactly the same as that of the cycle of length 9 as shown in Fig. 2.) As to the remaining $2k-9$ vertices, colour them successively, every second with green, and the others with both red and blue. The corresponding covering by complete subgraphs consistsaga in of all the edges adjacent with a vertex coloured with two different colours.

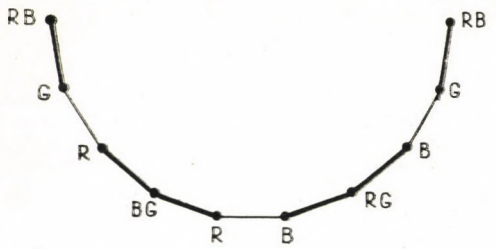


Fig. 3

On the normal product of perfect and normal graphs

As mentioned in the introduction, information-theoretical problems led in a rather natural way to the concept of *normal product*. (See [1], [2] and [4].)

One can think of a graph as of a set of symbols for some device which transmits infromation. If we join by an edge those points of the set which "can be confounded during the transmission", it becomes clear that if we transmit longer sequences of juxtaposed symbols and look at those ones which can be confounded with eachother, then the graph reflecting this situation will be a kind of power of our original graph. This intuitive picture is formulated mathematically in the following

Definition 3: (BERGE [1], p. 303). Let G and H be two graphs. Let $E(G)$ and $E(H)$ denote the respective sets of edges of the two graphs. *The normal product* $G \cdot H$ is a graph the vertex set of which is the cartesian product of the vertex sets of G and H. Two vertices (g, h) and (g', h') of $G \cdot H$ are joined by an edge if either

a) $(g, g') \in E(G)$ and $(h, h') \in E(H)$
b) $(g, g') \in E(G)$ and $h = h'$
c) $g = g'$ and $(h, h') \in E(H)$.

PROPOSITION 3. *The only perfect graphs every normal power of which is again perfect, are vertex-disjoint unions of complete graphs. If G is not of this type, G^3 is no longer perfect.*

The PROOF is easy. It is clear that if H is a vertex-disjoint union of complete graphs, so is any normal power of H, hence all the normal powers of H are perfect. However if G is not of this former type, then it contains a subgraph such as L (see Fig. 4). Now $L^3 \subseteq G^3$ and L^3 contains an odd cycle of length 5 with the following vertices (using the notation of Fig. 4): (100), (211), (122), (021), (010). Finally, a normal product, one factor of which is not perfect, can not be perfect.

One can ask what properties of perfect graphs are preserved by normal multiplication. The answer is given by the following (trivial)

PROPOSITION 4. *The normal product of normal graphs is again normal.*

The PROOF is in fact immediate. We shall prove that the normal product of two normal graphs is normal. Let G be a normal graph. Suppose that $G = \bigcup_i C_i = \bigcup_j F_j$ where the C_i's are complete subgraphs and the F_j's independent sets with $C_i \cap F_j \neq \emptyset$. For H, another normal graph we suppose the existence of two coverings, $H = \bigcup_k D_k = \bigcup_l P_l$ where again the D_k's are complete subgraphs and the P_l's independent sets with $D_k \cap P_l \neq \emptyset$. Then $G \cdot H$ is covered by both the complete $C_i \times D_k$'s and the independent $F_j \times P_l$'s. By our supposition every $C_i \times D_k$ has a non-void intersection with every $F_j \times P_l$, and hence it follows that $G \cdot H$ really has the normal property.

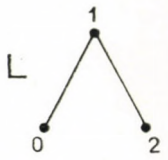


Fig. 4

On the other hand, there exist graphs, no normal power of which is normal. Such a graph is obtained from an odd cycle of length 5 by adding a new vertex adjacent with exactly one vertex of the cycle. In fact, one easily sees that such a graph is not normal. Further, if G is not normal, and in the graph H there is at least one vertex which is contained in exactly one maximal complete subgraph of H, then $G \cdot H$ can not be normal.

Finally, we state an open problem connected with normal product.

Problem: Do there exist two graphs G and H such that their normal product $G \cdot H$ is normal but G is not.

REFERENCES

[1] BERGE, CLAUDE: *Graphes et hypergraphes,* Dunod, Paris, (1970).
[2] KÖRNER, J.—LONGO, G.: Two-step encoding for finite sources, *IEEE Transactions on Information Th.* **IT—19** (1973), 778—782
[3] LOVÁSZ, L.: Normal hypergraphs and the perfect graph conjecture, *Discrete Mathematics.* **2**, no. 3 (1972), 253—267.
[4] SHANNON, C. E.: The zero-error capacity of a noisy channel, *Comp. Information Theory, IRE Trans.* **3** (1956), 3—15.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received April 28, 1973)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 411—423.

FUNCTIONAL LIMIT THEOREMS FOR LACUNARY TRIGONOMETRIC AND WALSH SERIES

by

I. BERKES

1. Introduction

Let $n_1, n_2, \ldots$ be a sequence of positive integers satisfying the condition

$$\frac{n_{k+1}}{n_k} \geqq q > 1 \qquad (k = 1, 2, \ldots). \tag{1.1}$$

It is a well-known fact (see [1], [2]) that the sequence $\sqrt{2}\cos 2\pi n_k x$ $(k=1, 2, \ldots)$ exhibits certain properties of independent random variables, e.g. the central limit theorem and the law of the iterated logarithm hold for it. The same is valid for the sequence $w_{n_k}(x)$ $(k=1, 2, \ldots)$ where $w_n(x)$ denotes the n-th Walsh function (see [3], [4], [5]). The purpose of the present paper is to prove the functional version of these theorems.

Definition. Let $Y_1, Y_2, \ldots$ be a sequence of random variables, $S_N = Y_1 + \ldots + Y_N$ $(S_0=0)$ and let $\chi_N(t)$ denote the random function defined in $[0, 1]$ which is linear in every interval $\left[\frac{k-1}{N}, \frac{k}{N}\right]$ $(1 \leqq k \leqq N)$ and $\chi_N\left(\frac{k}{N}\right) = S_k$ $(0 \leqq k \leqq N)$. Put $\varphi_N(t) = N^{-1/2}\chi_N(t)$, $\psi_N(t) = (2N \log\log N)^{-1/2}\chi_N(t)$. We say that the sequence $Y_1, Y_2, \ldots$ obeys the functional central limit theorem (invariance principle) if the sequence $\varphi_N(t)$ of processes converges weakly to the Wiener-process in the ordinary sense, i.e. the measures μ_N on the space $C[0, 1]$, corresponding to the processes $\varphi_N(t)$, converge weakly to the Wiener-measure. On the other hand, we say that the sequence $Y_1, Y_2, \ldots$ obeys the functional (Strassen-type) law of the iterated logarithm if the sequence $\psi_N(t)$ is equicontinuous with probability one and the set of its norm-limit points (in the norm $C[0, 1]$) coincides with the set

$$K = \left\{x(t) : x(t) \text{ is absolutely continuous in } [0, 1],\ x(0) = 0 \text{ and } \int_0^1 (\dot{x}(t))^2\, dt \leqq 1\right\}.$$

In this paper we shall prove the following theorems:

THEOREM 1. *If $n_1, n_2, \ldots$ are positive integers satisfying* (1.1) *then the sequence $\sqrt{2}\cos 2\pi n_k x$ obeys the functional central limit theorem and the functional law of the iterated logarithm (with respect to the Lebesgue-measure).*

THEOREM 2. *If the positive integers $n_1, n_2, \ldots$ satisfy* (1.1) *then the sequence $w_{n_k}(x)$ obeys the functional central limit theorem and the functional law of the iterated logarithm (with respect to the Lebesgue-measure).*

The phrase "with respect to the Lebesgue-measure" means that the underlying probability space in Theorems 1 and 2 is $(\Omega_0, \mathscr{A}_0, P_0)$ where $\Omega_0=[0, 1]$, $\mathscr{A}_0$ is the class of Borel-subsets of the interval [0, 1] and P_0 is the Lebesgue-measure on $\mathscr{A}_0$.

We formulate also a third theorem which is a common generalization of Theorem 1 and Theorem 2:

THEOREM 3. *Let $X_1, X_2, \ldots$ be a uniformly bounded sequence of random variables satisfying the following three conditions:*

(i) $\mathsf{E}X_n = 0$ $(n = 1, 2, \ldots)$

(ii) *For every $1 \leqq n < m$ there exist constants a_ν $(m-n \leqq \nu \leqq m+n)$ such that*

$$X_m X_n = \sum_{\nu=m-n}^{m+n} a_\nu X_\nu$$

(iii) *For every $n \geqq 1$ there exists a constant c_n such that*

$$X_n^2 = 1 + c_n X_{2n}.$$

Let further $n_1, n_2, \ldots$ be a sequence of positive integers satisfying (1.1). *Then the sequence X_{n_k} $(k=1, 2, \ldots)$ obeys the functional central limit theorem and the functional law of the iterated logarithm.*

As a matter of fact, the author knows no sequences satisfying the conditions of Theorem 3 except the trigonometric and Walsh systems. Our main purpose with the formulation of Theorem 3 was to treat Theorems 1 and 2 in a unified manner.

To deduce Theorem 1 from Theorem 3, it suffices to mention the identities $2 \cos mx \cos nx = \cos (m+n)x + \cos (m-n)x$ and $2 \cos^2 nx = 1 + \cos 2nx$. The situation is equally simple in the case of Theorem 2 since $w_n^2 = 1$, furthermore for every $1 \leqq n < m$ we have $w_m(x) w_n(x) = w_k(x)$ where k is an integer such that $m-n \leqq k \leqq m+n$ (see [5], p. 63).

Remarks. 1. The statement of Theorem 1 is valid also for the sequence $\sqrt{2} \sin 2\pi n_k x$ and more generally for the sequence $\sqrt{2} \cos 2\pi (n_k x + \alpha_k)$ where α_k are arbitrary. This can easily be seen from the proof of Theorem 3.

2. In the case $q \geqq 2$ the random variables w_{n_k} are independent, thus the statement of Theorem 2 is implied by classical results for such values of q. The first statement of Theorem 1 (the functional central limit theorem) is also known in a special case; namely, if $q \geqq 3$ then the sequence $\sqrt{2} \cos 2\pi n_k x$ is a strongly multiplicative system and the functional central limit theorem for such sequences has been recently proved by MCLEISH (see [6]). Except for these special cases, Theorems 1—3 are new.

2. Proof of Theorem 3.

The difficult part of the theorem is the second statement i.e. the functional law of the iterated logarithm; in its proof we shall make use of our previous result ([7]; Theorem 2). The functional central limit theorem will follow automatically from this proof but it could be proved independently by simpler methods. The present derivation, however, yields som additional information on the rate of convergence.

Put $S_k = X_{n_1}+X_{n_2}+\ldots+X_{n_k}$ $(S_0=0)$ and let $\chi_N(t)$, $\varphi_N(t)$, $\psi_N(t)$ be defined by means of the S_k-s similarly as in the Definition in section 1; let us further introduce the piecewise constant random function $\tilde{\varphi}_N(t)=N^{-1/2}S_{[Nt]}$ $(0\leqq t\leqq 1)$. For every fixed $0\leqq t_1<t_1'\leqq\ldots\leqq t_r<t_r'\leqq 1$ $(r=1, 2, \ldots)$ let $f_N^{(t_1,t_1',\ldots,t_r,t_r')}(\lambda_1, \ldots, \lambda_r)$ denote the characteristic function of the random vector

$$\left(\frac{\tilde{\varphi}_N(t_1')-\tilde{\varphi}_N(t_1)}{\sqrt{t_1'-t_1}}, \ldots, \frac{\tilde{\varphi}_N(t_r')-\tilde{\varphi}_N(t_r)}{\sqrt{t_r'-t_r}}\right). \tag{2.1}$$

Let us finally introduce the notation $\|\lambda\|$ for the absolute value of a vector $\lambda=(\lambda_1, \ldots, \lambda_r)$, i.e. put $\|\lambda\| = (\lambda_1^2+\ldots+\lambda_r^2)^{1/2}$.

We shall prove the following two statements:

A. For every $k\geqq 1$, $t\geqq 0$ and every $a_1, a_2, \ldots$ we have

$$\mathsf{P}(|a_1X_{n_1}+\ldots+a_kX_{n_k}|\geqq t\sqrt{a_1^2+\ldots+a_k^2}) \leqq C_1e^{-C_2t^2} \tag{2.2}$$

where C_1, C_2 are positive constants depending only on q and the sequence $X_1, X_2, \ldots$.

B. For every $0\leqq t_1<t_1'\leqq\ldots\leqq t_r<t_r'\leqq 1$, $N\geqq 1$ and for every $\lambda=(\lambda_1, \ldots, \lambda_r)$ satisfying $\|\lambda\|\leqq C_3(Nt_0)^{1/10}$ we have

$$\left|f_N^{(t_1,\ldots,t_r')}(\lambda_1, \ldots, \lambda_r)-\exp\left(-\frac{1}{2}\sum_{k=1}^{r}\lambda_k^2\right)\right| \leqq \frac{C_4}{(Nt_0)^{1/20}} \tag{2.3}$$

where $t_0=\min(t_1'-t_1, \ldots, t_r'-t_r)$, C_3, C_4 are positive constants which depend only on q, r and the sequence $X_1, X_2, \ldots$ (and are independent of $N, t_1, t_1', \ldots, t_r, t_r'$).

By Theorem 2 of [7], A and B together imply that the sequence X_{n_k} obeys the functional law of the iterated logarithm and we can easily see that they imply also the functional central limit theorem for the sequence X_{n_k}. In fact, we obtain from (2.2) for every $1\leqq k<l$:

$$\mathsf{E}(|X_{n_k}+X_{n_{k+1}}+\ldots+X_{n_l}|^4) \leqq$$

$$\leqq \sum_{s=0}^{\infty}[(s+1)\sqrt{l-k+1}]^4\mathsf{P}(|X_{n_k}+X_{n_{k+1}}+\ldots+X_{n_l}|\geqq s\sqrt{l-k+1}) \leqq$$

$$\leqq (l-k+1)^2\sum_{s=0}^{\infty}(s+1)^4C_1e^{-C_2s^2} = C_5(l-k+1)^2$$

where C_5 depends only on C_1, C_2; it is easy now to deduce from this moment inequality (cf. [8] p. 59 and p. 94) that the sequence $\varphi_N(t)$ of processes is tight in the usual measure-theoretic sense. Furthermore, statement B and the trivial estimate $\sup_{0\leqq t\leqq 1}|\varphi_N(t)-\tilde{\varphi}_N(t)| \leqq N^{-1/2}\max_{1\leqq k\leqq N}|X_{n_k}| \leqq N^{-1/2}K_1$ (where K_1 denotes a common upper bound of the r.v.s $|X_N|$) imply that the finite dimensional distributions of the process $\varphi_N(t)$ converge to the finite dimensional distributions of the Wiener-process. Applying PROHOROV's theorem ([8] p. 54), we obtain the desired result.

Before the proof of statements A and B we mention a few simple properties of the sequence $X_1, X_2, \ldots$ which will be used repeatedly in the sequel. Let us agree that d denotes the least positive integer such that $q^d\geqq 3$.

1. If $1 \leqq m_1 < m_2 < \ldots < m_k$ $(k=2, 3, \ldots)$ are integers satisfying $m_k > m_1 + \ldots + m_{k-1}$ then we have

$$X_{m_1} X_{m_2} \ldots X_{m_k} = \sum_{v = m_k - (m_{k-1} + \ldots + m_1)}^{m_k + (m_{k-1} + \ldots + m_1)} c_v X_v \tag{2.4}$$

where c_v are suitable constants.

2a. If $1 \leqq l_1 < l_2 < \ldots < l_k$ $(k=2, 3, \ldots)$ are integers satisfying $\frac{l_{v+1}}{l_v} \geqq 3$ $(v=1, \ldots, k-1)$ then we have

$$\mathsf{E}(X_{l_1} X_{l_2} \ldots X_{l_k}) = 0.$$

2b. If $X_{n_{j_1}}, X_{n_{j_2}}, \ldots, X_{n_{j_s}}$ $(s=2, 3, \ldots)$ is a finite subsystem of the sequence X_{n_k} such that $j_{v+1} - j_v \geqq d$ $(v=1, \ldots, s-1)$ then we have

$$\mathsf{E}(X_{n_{j_1}} X_{n_{j_2}} \ldots X_{n_{j_s}}) = 0.$$

In the case $k=2$ relation (2.4) is identical with condition (ii) of Theorem 3, the case $k>2$ follows by trivial induction from (ii). The integration of (2.4) leads to $\mathsf{E}(X_{m_1} X_{m_2} \ldots X_{m_k}) = 0$ (because of (i)) thus 2a. follows from the estimate $l_{k-1} + l_{k-2} + \ldots + l_1 < l_k \frac{1}{3} + l_k \frac{1}{3^2} + \ldots + l_k \frac{1}{3^{k-1}} < l_k \left(\frac{1}{3} + \frac{1}{3^2} + \ldots + \frac{1}{3^{k-1}} + \ldots \right) = \frac{1}{2} l_k < l_k$. Finally, 2b. is a special case of 2a. since (1.1) and $j_{v+1} - j_v \geqq d$ imply $\frac{n_{j_{v+1}}}{n_{j_v}} \geqq q^d \geqq 3$.

We shall need also the following lemma, due to AZUMA:

LEMMA. *Let $Y_1, Y_2, \ldots, Y_N$ be a sequence of r.v.-s such that $|Y_i| \leqq M$ $(i=1, 2, \ldots, N)$ and $\mathsf{E}(Y_{i_1} Y_{i_2} \ldots Y_{i_l}) = 0$ for every $1 \leqq i_1 < i_2 < \ldots < i_l \leqq N$ $(l=1, 2, \ldots, N)$. Then for every $b_1, \ldots, b_N$, $t \geqq 0$ and $h \geqq 0$ we have*

$$\mathsf{E}(\exp\{t|b_1 Y_1 + \ldots + b_N Y_N|\}) \leqq C_6 \exp\{C_7 (b_1^2 + \ldots + b_N^2) t^2\} \tag{2.5}$$

and

$$\mathsf{P}\left(|b_1 Y_1 + \ldots + b_N Y_N| \geqq h \sqrt{b_1^2 + \ldots + b_N^2}\right) \leqq C_8 \exp(-C_9 h^2) \tag{2.6}$$

where C_6, C_7, C_8, C_9 are positive constants depending only on M.

For the proof of (2.5) see [9]; (2.6) follows from (2.5) by a routine application of the Markov inequality.

Turning to the proof of statement A, let us estimate the left side of (2.2) as follows:

$$\mathsf{P}\left(|a_1 X_{n_1} + \ldots + a_k X_{n_k}| \geqq t \sqrt{a_1^2 + \ldots + a_k^2}\right) \leqq \tag{2.7}$$

$$\leqq \sum_{l=0}^{d-1} \mathsf{P}\left(\Big| \sum_{\substack{1 \leqq v \leqq k \\ v \equiv l \pmod d}} a_v X_{n_v} \Big| \geqq \frac{t}{2d} \sqrt{a_1^2 + \ldots + a_k^2} \right).$$

Using property 2b. and the lemma we obtain for $0 \leqq l \leqq d-1$:

$$\mathsf{P}\left(\left|\sum_{\substack{1 \leqq \nu \leqq k \\ \nu \equiv l \,(\mathrm{mod}\, d)}} a_\nu X_{n_\nu}\right| \geqq \frac{t}{2d}\sqrt{a_1^2+\ldots+a_k^2}\right) \leqq$$

$$\leqq \mathsf{P}\left(\left|\sum_{\substack{1 \leqq \nu \leqq k \\ \nu \equiv l \,(\mathrm{mod}\, d)}} a_\nu X_{n_\nu}\right| \geqq \frac{t}{2d}\sqrt{\sum_{\substack{1 \leqq \nu \leqq k \\ \nu \equiv l \,(\mathrm{mod}\, d)}} a_\nu^2}\right) \leqq C_{10}\exp\left(-C_{11}\frac{t^2}{4d^2}\right)$$

where C_{10}, C_{11} are positive constants depending only on the uniform bound of the sequence $|X_1|, |X_2|, \ldots$. This estimate, together with (2.7), gives

$$\mathsf{P}\left(|a_1 X_{n_1}+\ldots+a_k X_{n_k}| \geqq t\sqrt{a_1^2+\ldots+a_k^2}\right) \leqq C_{10} d \exp\left(-\frac{1}{4d^2} C_{11} t^2\right)$$

and hence statement A is proved.

For the proof of statement B let us first remark that by (1.1) there exists a $0<\varepsilon_0<1$ such that the intervals $[n_k(1-\varepsilon_0), n_k(1+\varepsilon_0)]$ $(k=1, 2, \ldots)$ are pairwise disjoint. Let us choose a large integer Q such that

$$Q > 2, \quad \frac{Q}{Q-1} < 1+\varepsilon_0, \quad \frac{Q-2}{Q-1} > 1-\varepsilon_0$$

finally let p be a positive integer such that $q^p \geqq Q$. Let us divide the members of the sum $\sum_{\nu=[Nt_k]+1}^{[Nt_k']} X_{n_\nu}$ into blocks in such a manner that the length of every block (i.e. the number of the X_{n_ν}-s belonging to it) be a number between p and $2p$ and the number of blocks be an even number. This is obviously possible if $Nt_0 \geqq 2p+2$, since in this case $[Nt_k']-[Nt_k] \geqq N(t_k'-t_k)-2 \geqq 2p$. Let the blocks be denoted by $g_1^{(k)}, g_2^{(k)}, \ldots, g_{z_k}^{(k)}$ (z_k is even), then we have

$$\sum_{\nu=[Nt_k]+1}^{[Nt_k']} X_{n_\nu} = \sum_{j=1}^{z_k} g_j^{(k)}.$$

A block $g_j^{(k)}$ will be called an even resp. an odd one according as j is even or odd.

Making use of the simple formula

$$e^z = (1+z)e^{(z^2/2)+r(z)} \tag{2.8}$$

where $|r(z)| \leqq C_{12}|z|^3$ for $|z| \leqq C_{13}$ (C_{12}, C_{13} are absolute constants) we obtain

$$\exp\left(\frac{i\lambda_k}{\sqrt{N(t_k'-t_k)}} g_j^{(k)}\right) =$$

$$= \left(1+\frac{i\lambda_k}{\sqrt{N(t_k'-t_k)}} g_j^{(k)}\right)\exp\left\{-\frac{\lambda_k^2}{2N(t_k'-t_k)}(g_j^{(k)})^2 + r\left(\frac{i\lambda_k}{\sqrt{N(t_k'-t_k)}} g_j^{(k)}\right)\right\}$$

$$(1 \leqq k \leqq r, \ 1 \leqq j \leqq z_k)$$

and thus

$$f_N^{(t_1,\dots,t_r')}(\lambda_1,\dots,\lambda_r) = \tag{2.9}$$

$$= \mathsf{E}\left(\exp\left\{\frac{i\lambda_1}{\sqrt{t_1'-t_1}}\left(\tilde\varphi_N(t_1')-\tilde\varphi_N(t_1)\right)+\dots+\frac{i\lambda_r}{\sqrt{t_r'-t_r}}\left(\tilde\varphi_N(t_r')-\tilde\varphi_N(t_r)\right)\right\}\right) =$$

$$= \mathsf{E}\left(\exp\left\{\frac{i\lambda_1}{\sqrt{N(t_1'-t_1)}}\sum_{v=[Nt_1]+1}^{[Nt_1']}X_{n_v}+\dots+\frac{i\lambda_r}{\sqrt{N(t_r'-t_r)}}\sum_{v=[Nt_r]+1}^{[Nt_r']}X_{n_v}\right\}\right) =$$

$$= \mathsf{E}\left(\exp\left\{\frac{i\lambda_1}{\sqrt{N(t_1'-t_1)}}\sum_{j=1}^{z_1}g_j^{(1)}+\dots+\frac{i\lambda_r}{\sqrt{N(t_r'-t_r)}}\sum_{j=1}^{z_r}g_j^{(r)}\right\}\right) =$$

$$= \mathsf{E}\left(\prod_{k=1}^{r}\prod_{j=1}^{z_k}\left(1+\frac{i\lambda_\kappa}{\sqrt{N(t_k'-t_k)}}g_j^{(k)}\right)\exp\left\{-\sum_{k=1}^{r}\frac{\lambda_k^2}{2N(t_k'-t_k)}P_k+\sum_{k=1}^{r}Q_k\right\}\right)$$

where

$$P_k = \sum_{j=1}^{z_k}(g_j^{(k)})^2$$

$$Q_k = \sum_{j=1}^{z_k} r\left(\frac{i\lambda_k}{\sqrt{N(t_k'-t_k)}}g_j^{(k)}\right).$$

We shall prove that for $Nt_0 \geqq 2p+2$ we have

$$\mathsf{E}\left(\prod_{k=1}^{r}\prod_{j=1}^{z_k}\left(1+\frac{i\lambda_k}{\sqrt{N(t_k'-t_k)}}g_j^{(k)}\right)\right) = 1 \tag{2.10}$$

furthermore there exist constants C_{14}, C_{15} (depending on q and the sequence $X_1, X_2, \dots$) such that for $Nt_0 \geqq 2p+2$, $1 \leqq k \leqq r$ we have

$$\mathsf{P}\left(|P_k-([Nt_k']-[Nt_k])| \geqq \left(N(t_k'-t_k)\right)^{3/4}\right) \leqq C_{14}\exp\left(-C_{15}\sqrt{Nt_0}\right). \tag{2.11}$$

To prove (2.10), let us write

$$\prod_{k=1}^{r}\prod_{j=1}^{z_k}\left(1+\frac{i\lambda_k}{\sqrt{N(t_k'-t_k)}}g_j^{(k)}\right) = V_1V_2$$

where

$$V_1 = \prod_{k=1}^{r}\prod_{\substack{j=1\\ j \text{ is odd}}}^{z_k}\left(1+\frac{i\lambda_k}{\sqrt{N(t_k'-t_k)}}g_j^{(k)}\right)$$

$$V_2 = \prod_{k=1}^{r}\prod_{\substack{j=1\\ j \text{ is even}}}^{z_k}\left(1+\frac{i\lambda_k}{\sqrt{N(t_k'-t_k)}}g_j^{(k)}\right).$$

Replacing every $g_j^{(k)}$ by the corresponding sum (according to its definition), carrying out the multiplication and making use of relation (2.4) (the applicability of (2.4) will be justified by the estimates below) we obtain that V_1 can be written as follows:

$$V_1 = 1 + \sum_{\nu} c_\nu X_\nu \tag{2.12}$$

where the sum consists of finitely many terms and c_ν are complex constants. Actually we can say something more about the indices ν occuring in this sum: for every such ν we can find numbers $n_{s_1}, n_{s_2}, \ldots, n_{s_t}$ $(s_1 < s_2 < \ldots < s_t)$ belonging to different odd blocks $g_{j_1}^{(k_1)}, \ldots, g_{j_t}^{(k_t)}$ (the phrase "n_i belongs to the block $g_j^{(k)}$" means that the sum $g_j^{(k)}$ contains the variable X_{n_i}) such that

$$n_{s_t} - (n_{s_{t-1}} + \ldots + n_{s_1}) < \nu < n_{s_t} + (n_{s_{t-1}} + \ldots + n_{s_1}). \tag{2.13}$$

It is easy to see that the interval (n_{s_1}, n_{s_2}) contains at least p different n_i-s. In fact, n_{s_1} and n_{s_2} belong to the different odd blocks $g_{j_1}^{(k_1)}$ and $g_{j_2}^{(k_2)}$ resp., thus we can find an even block $g_{j^*}^{(k^*)}$ such that each n_i of this block lies in the interval (n_{s_1}, n_{s_2}). (We used here the fact that the number of blocks $g_j^{(k)}$ is even for every fixed k.) Now (1.1) implies $\frac{n_{s_2}}{n_{s_1}} \geqq \frac{n_{s_1+p}}{n_{s_1}} \geqq q^p \geqq Q$ and since a similar statement holds for the pairs $(n_{s_2}, n_{s_3}), \ldots, (n_{s_{t-1}}, n_{s_t})$, we can estimate the numbers $n_{s_t} - (n_{s_{t-1}} + \ldots + n_{s_1})$ and $n_{s_t} + (n_{s_{t-1}} + \ldots + n_{s_1})$, appearing in (2.13), as follows:

$$n_{s_t} - (n_{s_{t-1}} + \ldots + n_{s_1}) \geqq$$

$$\geqq n_{s_t} - \left(n_{s_t} \frac{1}{Q} + n_{s_t} \frac{1}{Q^2} + \ldots + n_{s_t} \frac{1}{Q^{t-1}} \right) > n_{s_t} \left(1 - \frac{1}{Q} - \frac{1}{Q^2} + \ldots - \frac{1}{Q^{t-1}} - \ldots \right) =$$

$$= n_{s_t} \frac{Q-2}{Q-1} > n_{s_t}(1-\varepsilon_0)$$

(we see here that the left side of (2.13) is positive and this justifies the applicability of (2.4) above)

$$n_{s_t} + (n_{s_{t-1}} + \ldots + n_{s_1}) \leqq$$

$$\leqq n_{s_t} + \left(n_{s_t} \frac{1}{Q} + n_{s_t} \frac{1}{Q^2} + \ldots + n_{s_t} \frac{1}{Q^{t-1}} \right) < n_{s_t} \left(1 + \frac{1}{Q} + \frac{1}{Q^2} + \ldots + \frac{1}{Q^{t-1}} + \ldots \right) =$$

$$= n_{s_t} \frac{Q}{Q-1} < n_{s_t}(1+\varepsilon_0)$$

i.e. $n_{s_t}(1-\varepsilon_0) < \nu < n_{s_t}(1+\varepsilon_0)$. This result means that every ν occuring in the sum (2.12) lies in the interior of an interval $[n_k(1-\varepsilon_0), n_k(1+\varepsilon_0)]$ where n_k belongs to an odd block. Repeating the argument for V_2 we obtain that

$$V_2 = 1 + \sum_{\mu} d_\mu X_\mu \tag{2.14}$$

where the sum consists of finitely many terms, d_μ are complex constants and every μ occuring in the sum lies in the interior of an interval $[n_k(1-\varepsilon_0), n_k(1+\varepsilon_0)]$ where n_k belongs to an even block. Since the intervals $[n_k(1-\varepsilon_0), n_k(1+\varepsilon_0)]$ $(k=1, 2, \ldots)$ are pairwise disjoint, all the X_ν-s appearing in (2.12) differ from all the X_μ-s appearing in (2.14). Therefore (i) and the pairwise orthogonality of the variables X_k (which is a consequence of (i) and (ii)) give

$$\mathsf{E}(V_1 V_2) = \mathsf{E}(1 + \sum_{\nu} c_\nu X_\nu + \sum_{\mu} d_\mu X_\mu + \sum_{\nu,\mu} c_\nu d_\mu X_\nu X_\mu) = 1$$

and hence (2.10) is proved.

For the proof of (2.11) let us first investigate the expression $P_k = \sum_{j=1}^{z_k} (g_j^{(k)})^2$. Every $g_j^{(k)}$ is of the form

$$g_j^{(k)} = X_{n_s} + X_{n_{s+1}} + \ldots + X_{n_t} \tag{2.15}$$

and thus we have by (ii) and (iii)

$$(g_j^{(k)})^2 = \sum_{\nu=s}^{t} X_{n_\nu}^2 + 2 \sum_{s \leqq \nu < \mu \leqq t} X_{n_\nu} X_{n_\mu} = \tag{2.16}$$

$$= (t-s+1) + \sum_{\vartheta=s}^{t} c'_\vartheta X_{2n_\vartheta} + \sum_{\varrho} d'_\varrho X_\varrho = (t-s+1) + h_j^{(k)}.$$

Here c'_ϑ, d'_ϱ are real constants, furthermore for every ϱ appearing in the sum $\sum_{\varrho} d'_\varrho X_\varrho$ we can find numbers n_ν, n_μ such that $s \leqq \nu < \mu \leqq t$ and $n_\mu - n_\nu \leqq \varrho \leqq n_\mu + n_\nu$. For such numbers n_ν, n_μ we obviously have $n_\mu - n_\nu \geqq n_\mu - n_{\mu-1} \geqq n_\mu - \frac{1}{q} n_\mu = \left(1-\frac{1}{q}\right) n_\mu \geqq$ $\geqq \left(1-\frac{1}{q}\right) n_s$, $n_\mu + n_\nu \leqq n_t + n_t = 2n_t$, thus the quantity $h_j^{(k)}$ in (2.16) is a sum of the form $\sum c''_i X_i$ where all the occuring indices i lie in the interval $\left[\left(1-\frac{1}{q}\right) n_s, 2n_t\right]$. Summing up (2.16) for $1 \leqq j \leqq z_k$ and using the fact that the blocks $g_j^{(k)}$ $(1 \leqq j \leqq z_k)$ contain altogether $[Nt'_k] - [Nt_k]$ members, we obtain

$$P_k = \sum_{j=1}^{z_k} (g_j^{(k)})^2 = [Nt'_k] - [Nt_k] + \sum_{j=1}^{z_k} h_j^{(k)}.$$

Hence (2.11) is equivalent to the following relation:

$$\mathsf{P}\left(\left|\sum_{j=1}^{z_k} h_j^{(k)}\right| \geqq (N(t'_k - t_k))^{3/4}\right) \leqq C_{14} \exp\left(-C_{15}\sqrt{Nt_0}\right). \tag{2.17}$$

Let d^* be the least positive integer such that $q^{pd^*} \geqq \frac{6}{1-(1/q)}$. We shall prove that

for every sequence $j_1<j_2<\ldots<j_\gamma$ $(\gamma=2, 3, \ldots)$ of indices satisfying the relation $j_{\nu+1}-j_\nu \geqq d^*$ $(\nu=1, \ldots, \gamma-1)$ we have

$$\mathsf{E}(h_{j_1}^{(k)} h_{j_2}^{(k)} \ldots h_{j_\gamma}^{(k)}) = 0. \tag{2.18}$$

To this aim, let us first consider the quantities $h_{j_1}^{(k)}$ and $h_{j_2}^{(k)}$. If the corresponding $g_j^{(k)}$-s are $g_{j_1}^{(k)}=X_{n_{s_1}}+\ldots+X_{n_{t_1}}$, $g_{j_2}^{(k)}=X_{n_{s_2}}+\ldots+X_{n_{t_2}}$, then $h_{j_1}^{(k)}$ and $h_{j_2}^{(k)}$ can be written as sums of the form

$$h_{j_1}^{(k)} = \sum c_l^{(1)} X_l \tag{2.19}$$

$$h_{j_2}^{(k)} = \sum c_m^{(2)} X_m \tag{2.20}$$

where $c_l^{(1)}$ and $c_m^{(2)}$ are real constants and the indices l, m appearing in the first and second sum lie in the intervals $\left[\left(1-\frac{1}{q}\right)n_{s_1}, 2n_{t_1}\right]$, $\left[\left(1-\frac{1}{q}\right)n_{s_2}, 2n_{t_2}\right]$, resp. By our choice every block $g_j^{(k)}$ contains at least p summands, thus $j_2-j_1\geqq d^*$ implies that the number of indices i such that $n_{t_1}<n_i<n_{s_2}$ exceeds pd^* and thus

$$\frac{\left(1-\frac{1}{q}\right)n_{s_2}}{2n_{t_1}} \geqq \frac{1-\frac{1}{q}}{2} q^{pd^*} \geqq 3.$$

This implies that every pair (l, m) of indices occuring in (2.19) and (2.20) satisfies the relation $m/l\geqq 3$. A similar statement holds for the pairs $(h_{j_2}^{(k)}, h_{j_3}^{(k)})$, ..., $(h_{j_{\gamma-1}}^{(k)}, h_{j_\gamma}^{(k)})$ and thus, replacing every $h_j^{(k)}$ by the corresponding sum in the left side of (2.18), carrying out the multiplication and making use of property 2a., we obtain (2.18).

Let now K_1 denote a uniform upper bound of the sequence $|X_1|, |X_2|, \ldots$. By (2.15) we have $|g_j^{(k)}|\leqq(t-s+1)K_1\leqq 2pK_1$ and thus (2.16) implies

$$|h_j^{(k)}| = |(g_j^{(k)})^2-(t-s+1)| \leqq 4p^2K_1^2+2p. \tag{2.21}$$

Now we have

$$\mathsf{P}\left(\left|\sum_{j=1}^{z_k} h_j^{(k)}\right| \geqq \left(N(t_k'-t_k)\right)^{3/4}\right) \leqq \tag{2.22}$$

$$\leqq \sum_{l=0}^{d^*-1} \mathsf{P}\left(\left|\sum_{\substack{1\leqq j\leqq z_k \\ j\equiv l \,(\mathrm{mod}\, d^*)}} h_j^{(k)}\right| \geqq \frac{1}{2d^*}\left(N(t_k'-t_k)\right)^{3/4}\right).$$

It follows from relations (2.18) and (2.21) that the members of the sum

$$\sum_{\substack{1\leqq j\leqq z_k \\ j\equiv l \,(\mathrm{mod}\, d^*)}} h_j^{(k)}$$

satisfy the conditions of the lemma above and thus we obtain (using the estimate $z_k \leqq [Nt_k'] - [Nt_k] \leqq N(t_k' - t_k) + 2 \leqq 2N(t_k' - t_k)$ valid for $Nt_0 \geqq 2$):

$$\mathsf{P}\left(\Big|\sum_{\substack{1 \leqq j \leqq z_k \\ j \equiv l \,(\mathrm{mod}\, d^*)}} h_j^{(k)}\Big| \geqq \frac{1}{2d^*}\left(N(t_k' - t_k)\right)^{3/4}\right) \leqq$$

$$\leqq \mathsf{P}\left(\Big|\sum_{\substack{1 \leqq j \leqq z_k \\ j \equiv l \,(\mathrm{mod}\, d^*)}} h_j^{(k)}\Big| \geqq \frac{\frac{1}{2d^*}\left(N(t_k' - t_k)\right)^{3/4}}{\left(2N(t_k' - t_k)\right)^{1/2}} \sqrt{\sum_{\substack{1 \leqq j \leqq z_k \\ j \equiv l \,(\mathrm{mod}\, d^*)}} 1}\right) \leqq$$

$$\leqq C_{16} \exp\left(-C_{17} \frac{1}{4(d^*)^2}\left(N(t_k' - t_k)\right)^{1/2}\right) \leqq C_{16} \exp\left(-C_{17} \frac{1}{4(d^*)^2}(Nt_0)^{1/2}\right)$$

where C_{16} and C_{17} depend only on K_1 and p, i.e. only on q and the sequence $X_1, X_2, \ldots$, From (2.22) and (2.23) we obtain (2.17) and thus (2.11) is proved.

Having relations (2.10) and (2.11), the proof of statement B can be completed by following the method of [11]. Making use of (2.9), (2.10) and the inequality $1 + x \leqq e^x$, we obtain for $Nt_0 \geqq 2p + 2$:

$$\left|f_N^{(t_1, \ldots, t_r')}(\lambda_1, \ldots, \lambda_r) - \exp\left(-\frac{1}{2}\sum_{k=1}^{r} \lambda_k^2\right)\right| =$$

$$= \left|\mathsf{E}\left(\prod_{k=1}^{r} \prod_{j=1}^{z_k}\left(1 + \frac{i\lambda_k}{\sqrt{N(t_k' - t_k)}} g_j^{(k)}\right)\left\{\exp\left(-\sum_{k=1}^{r} \frac{\lambda_k^2}{2N(t_k' - t_k)} P_k + \sum_{k=1}^{r} Q_k\right) - \right.\right.$$

$$\left.\left. - \exp\left(-\frac{1}{2}\sum_{k=1}^{r} \lambda_k^2\right)\right\}\right)\right| \leqq$$

$$\leqq \mathsf{E}\left(\prod_{k=1}^{r} \prod_{j=1}^{z_k}\left(1 + \frac{\lambda_k^2}{N(t_k' - t_k)}(g_j^{(k)})^2\right)^{1/2}\left|\exp\left(-\sum_{k=1}^{r} \frac{\lambda_k^2}{2N(t_k' - t_k)} P_k + \sum_{k=1}^{r} Q_k\right) - \right.\right.$$

$$\left.\left. - \exp\left(-\frac{1}{2}\sum_{k=1}^{r} \lambda_k^2\right)\right|\right) \leqq$$

$$\leqq \mathsf{E}\left(\exp \sum_{k=1}^{r} \frac{\lambda_k^2}{2N(t_k' - t_k)} P_k \left|\exp\left(-\sum_{k=1}^{r} \frac{\lambda_k^2}{2N(t_k' - t_k)} P_k + \sum_{k=1}^{r} Q_k\right) - \right.\right.$$

$$\left.\left. - \exp\left(-\frac{1}{2}\sum_{k=1}^{r} \lambda_k^2\right)\right|\right) =$$

$$= \mathsf{E}\left(\left|\exp \sum_{k=1}^{r} Q_k - \exp \sum_{k=1}^{r} \frac{\lambda_k^2}{2N(t_k' - t_k)}\left(P_k - N(t_k' - t_k)\right)\right|\right) \leqq J_1 + J_2$$

where

$$J_1 = \mathsf{E}\left(\left|\exp \sum_{k=1}^{r} Q_k - 1\right|\right)$$

$$J_2 = \mathsf{E}\left(\left|\exp\left\{\sum_{k=1}^{r} \frac{\lambda_k^2}{2N(t_k' - t_k)}\left(P_k - N(t_k' - t_k)\right)\right\} - 1\right|\right)$$

Now we have $|g_j^{(k)}| \leqq 2pK_1$, thus $\frac{\|\lambda\|}{\sqrt{Nt_0}} 2pK_1 \leqq C_{13}$ (i.e. $\|\lambda\| \leqq C_{18}\sqrt{Nt_0}$) implies that the inequality $\left|\frac{i\lambda_k}{\sqrt{N(t_k - t_k)}} g_j^{(k)}\right| \leqq C_{13}$ holds for every k, j. Hence we obtain for $\|\lambda\| \leqq C_{18}\sqrt{Nt_0}$, $Nt_0 \geqq 2$:

$$\left|\sum_{k=1}^{r} Q_k\right| \leqq \sum_{k=1}^{r} \sum_{j=1}^{z_k} \left| r\left(\frac{i\lambda_k}{\sqrt{N(t_k' - t_k)}} g_j^{(k)}\right)\right| \leqq \tag{2.25}$$

$$\leqq \sum_{k=1}^{r} \sum_{j=1}^{z_k} C_{12} \left|\frac{i\lambda_k}{\sqrt{N(t_k' - t_k)}} g_j^{(k)}\right|^3 \leqq$$

$$\leqq \sum_{k=1}^{r} C_{12} \frac{\|\lambda\|^3 (2pK_1)^3}{(N(t_k' - t_k))^{3/2}} 2N(t_k' - t_k) \leqq C_{19} \frac{\|\lambda\|^3}{(Nt_0)^{1/2}}.$$

Here the constants C_{18}, C_{19} depend only on q, K_1, r; all the further constants $C_{20}, C_{21}, \ldots$ appearing in the sequel will have this property. Using now (2.25) and the inequality $|e^z - 1| \leqq 2|z|$, valid for $|z| \leqq C_{20}$, we obtain

$$J_1 \leqq 2\mathsf{E}\left(\left|\sum_{k=1}^{r} Q_k\right|\right) \leqq 2C_{19} \frac{\|\lambda\|^3}{(Nt_0)^{1/2}} \leqq \frac{2C_{19}}{(Nt_0)^{2/10}} \tag{2.26}$$

provided that $\left|\sum_{k=1}^{r} Q_k\right| \leqq C_{20}$, $\|\lambda\| \leqq C_{18}\sqrt{Nt_0}$, $Nt_0 \geqq 2$, $\|\lambda\| \leqq (Nt_0)^{1/10}$. According to (2.25), each of these conditions is satisfied if $\|\lambda\| \leqq C_{21}(Nt_0)^{1/10}$, $Nt_0 \geqq 2$. Let us now introduce the events

$$D_k = \left\{|P_k - N(t_k' - t_k)| < 2\left(N(t_k' - t_k)\right)^{3/4}\right\} \qquad k = 1, 2, \ldots, r$$

$$D = D_1 D_2 \ldots D_r$$

and let us split J_2 into two summands as follows: $J_2 = J_2^{(1)} + J_2^{(2)}$, where

$$J_2^{(1)} = \int_D |e^U - 1| \, d\mathsf{P} \qquad J_2^{(2)} = \int_{\bar{D}} |e^U - 1| \, d\mathsf{P}$$

$$U = \sum_{k=1}^{r} \frac{\lambda_k^2}{2N(t_k' - t_k)} \left(P_k - N(t_k' - t_k)\right).$$

On the set D we have

$$|U| \leqq \sum_{k=1}^{r} \frac{\lambda_k^2}{2N(t_k' - t_k)} 2\left(N(t_k' - t_k)\right)^{3/4} = \sum_{k=1}^{r} \frac{\lambda_k^2}{\left(N(t_k' - t_k)\right)^{1/4}} \leqq \frac{\|\lambda\|^2}{(Nt_0)^{1/4}} \tag{2.27}$$

and hence we get (using the inequality $|e^z - 1| \leqq 2|z|$ again)

$$J_2^{(1)} \leqq \int_D 2|U| \, d\mathsf{P} \leqq \frac{2\|\lambda\|^2}{(Nt_0)^{1/4}} \leqq \frac{2}{(Nt_0)^{1/20}} \tag{2.28}$$

provided that $|U| \leqq C_{20}$, $\|\lambda\| \leqq (Nt_0)^{1/10}$. According to (2.27), both of these conditions are satisfied if $\|\lambda\| \leqq C_{22}(Nt_0)^{1/10}$, $Nt_0 \geqq 1$. On the other hand, on the set $\bar{D}_k$ we have

$$|P_k - ([Nt_k'] - [Nt_k])| \geqq |P_k - N(t_k' - t_k)| - 2 \geqq 2(N(t_k - t_k))^{3/4} - 2 \geqq (N(t_k' - t_k))^{3/4}$$

(provided that $Nt_0 \geqq 3$), thus (2.11) gives $\mathsf{P}(\bar{D}) \leqq \sum_{k=1}^{r} \mathsf{P}(\bar{D}_k) \leqq C_{23} \exp\left(-C_{15}\sqrt{Nt_0}\right)$. Using this fact, we obtain for $\|\lambda\| \leqq (Nt_0)^{1/10}$, $Nt_0 \geqq C_{24}$:

$$J_2^{(2)} \leqq \int_{\bar{D}} (e^U + 1)\, d\mathsf{P} \leqq \int_{\bar{D}} \left[\exp\left(\sum_{k=1}^{r} \frac{\lambda_k^2}{2N(t_k' - t_k)} P_k\right) + 1\right] d\mathsf{P} \leqq \tag{2.29}$$

$$\leqq C_{23} \exp\left(-C_{15}\sqrt{Nt_0}\right)\left[\exp\left\{\sum_{k=1}^{r} \frac{\lambda_k^2}{2N(t_k' - t_k)} (2pK_1)^2 2N(t_k' - t_k)\right\} + 1\right] \leqq$$

$$\leqq C_{23} \exp\left(-C_{15}\sqrt{Nt_0}\right)\left[\exp\left(C_{25} \sum_{k=1}^{r} \lambda_k^2\right) + 1\right] \leqq$$

$$\leqq 2C_{23} \exp\left(-C_{15}\sqrt{Nt_0} + C_{25}(Nt_0)^{1/5}\right) \leqq$$

$$\leqq 2C_{23} \exp\left(-C_{26}\sqrt{Nt_0}\right) \leqq \frac{C_{27}}{(Nt_0)^{1/2}}$$

From the estimates (2.28) and (2.29) we finally get

$$J_2 \leqq \frac{2}{(Nt_0)^{1/20}} + \frac{C_{27}}{(Nt_0)^{1/2}} \leqq \frac{C_{28}}{(Nt_0)^{1/20}}. \tag{2.30}$$

It is easy now to complete the proof of statement B. Let us assume that $\|\lambda\| \leqq$ $\leqq C_{29}(Nt_0)^{1/10}$, $Nt_0 \geqq C_{30}$, then (2.26) and (2.30) are valid and thus (2.24)—(2.30) give

$$\left| f_N^{(t_1, \ldots, t_r')}(\lambda_1, \ldots, \lambda_r) - \exp\left(-\frac{1}{2} \sum_{k=1}^{r} \lambda_k^2\right)\right| \leqq \frac{C_{31}}{(Nt_0)^{1/20}}. \tag{2.31}$$

On the other hand, it is possible to choose the constant C_{31} so large that $C_{31}(Nt_0)^{-1/20} > 2$ holds for $Nt_0 < C_{30}$ and in this case (2.31) will be valid also for $Nt_0 < C_{30}$ since the left side is at most 2 and the right side is greater than 2. This remark completes the proof of statement B and thus the proof of Theorem 3 is also completed.

3. Remarks

It is worth mentioning that the results of the preceding section yield not only the desired functional limit theorems but they also give estimates for the rate of convergence in the case of the ordinary central limit theorem. Applying statement

B with $r=1$, $t_1=0$, $t_1'=1$ and making use of ESSEEN's inequality ([10], p. 512) we easily obtain that under the conditions of Theorem 3 we have

$$\mathrm{P}\left(\frac{X_{n_1}+X_{n_2}+\ldots+X_{n_N}}{\sqrt{N}}<x\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-t^2/2}\,dt+O(N^{-\alpha}) \tag{3.1}$$

uniformly in x, where $\alpha=1/20$. The constant 1/20 could be improved to 1/6 by more careful estimates and some slight modifications of the method but we do not give the details here because the constant $\alpha=1/6$ does not seem to be the best one. It would be interesting to know whether (3.1) holds with $\alpha=1/2$ or at least with $\alpha=1/2-\varepsilon$ for each $\varepsilon>0$.

REFERENCES

[1] SALEM, R.—ZYGMUND, A.: On lacunary trigonometric series, *Proc. Nat. Acad. Sci. USA* **33** (1947), 333—338.

[2] WEISS, M.: On the law of the iterated logarithm for lacunary trigonometric series, *Trans. Amer. Math. Soc.* **91** (1959), 444—469.

[3] MORGENTHALER, G.: On Walsh-Fourier series, *Trans. Amer. Math. Soc.* **84** (1957), 472—507.

[4] RÉVÉSZ, P.—WSCHEBOR, M.: On the statistical properties of the Walsh functions, *MTA Mat. Kut. Int. Közl.* **9A** (1964), 543—554.

[5] GAPOSKIN, V. F.: Lacunary series and independent functions, *Uspehi Mat. Nauk* **21**/6 (1966), 3—82. (In Russian).

[6] MC.LEISH, D. L.: Dependent central limit theorems and invariance principles, Thesis, 1972. Uviversity of Montreal.

[7] BERKES, I.: The functional law of the iterated logarithm for dependent random variables, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* (to appear).

[8] BILLINGSLEY, P.: *Convergence of probability measures,* New York: Wiley, 1968.

[9] RÉVÉSZ, P.: A new law of the iterated logaritm for multiplicative systems and its applications, *Acta Sci. Math. Hung.* (to appear).

[10] FELLER, W.: *An introduction to probability theory and its applications,* New York: Wiley 1966.

[11] BERKES, I.: On Strassen's version of the loglog law for multiplicative systems, *Studia Sci. Math. Hung.* (to appear).

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received April 27, 1973)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 425—431.

ON STRASSEN'S VERSION OF THE LOGLOG LAW FOR MULTIPLICATIVE SYSTEMS

by

I. BERKES

1. Introduction

A sequence $X_1, X_2, \ldots$ of random variables is called a weakly multiplicative system if

$$\sum_{\substack{1 \leq i_1 < i_2 < \ldots < i_k \\ k=1,2,\ldots}} |\mathsf{E}(X_{i_1} X_{i_2} \ldots X_{i_k})| < +\infty.$$

In particular, if all the expectations $\mathsf{E}(X_{i_1} X_{i_2} \ldots X_{i_k})$ vanish, then the seequence $X_1, X_2, \ldots$ is called a multiplicative system. The properties of such sequences are very close to those of independent random variables, e.g. under certain conditions they obey the central limit theorem, the invariance principle and the law of the iterated logarithm (see [1]—[5]). The purpose of the present paper is to prove the functional (Strassen-type) law of the iterated logarithm for such sequences.

Definition. Let $X_1, X_2, \ldots$ be a sequence of r.v.-s, $S_n = X_1 + \ldots + X_n$ $(S_0 = 0)$ and let $\psi_n(t)$ denote the random function in [0, 1] which is linear in every interval $\left[\frac{k-1}{n}, \frac{k}{n}\right]$ $(1 \leq k \leq n)$ and $\psi_n\left(\frac{k}{n}\right) = (2n \log\log n)^{-1/2} S_k$ $(0 \leq k \leq n)$. The sequence $X_1, X_2, \ldots$ is said to obey the functional law of the iterated logarithm if the sequence $\psi_n(t)$ is equicontinuous with probability one and the set of its norm-limit points (in the norm $C[0, 1]$) coincides with the set

$$K = \left\{x(t) : x(t) \text{ is absolutely continuous in } [0, 1],\ x(0) = 0 \text{ and } \int_0^1 (\dot{x}(t))^2 dt \leq 1\right\}.$$

Our result can be formulated as follows:

THEOREM. *Let $X_1, X_2, \ldots$ be a uniformly bounded sequence of r.v.-s such that*

$$\sum_{\substack{1 \leq i_1 < i_2 < \ldots < i_k \\ k=1,2,\ldots}} |\mathsf{E}(X_{i_1} X_{i_2} \ldots X_{i_k})| < +\infty \tag{1.1}$$

and

$$\sum_{\substack{1 \leq i_1 < i_2 < \ldots < i_k \\ k=1,2,\ldots}} \left|\mathsf{E}\left((X_{i_1}^2 - 1)(X_{i_2}^2 - 1) \ldots (X_{i_k}^2 - 1)\right)\right| < +\infty. \tag{1.2}$$

Then the sequence $X_1, X_2, \ldots$ obeys the functional law of the iterated logarithm.

COROLLARY. *Let $X_1, X_2, \ldots$ be a uniformly bounded multiplicative system such that $\mathsf{E}(X_{i_1}^2 X_{i_2}^2 \ldots X_{i_k}^2)=1$ $(i_1<i_2<\ldots<i_k,\ k=1, 2, \ldots)$. Then the sequence $X_1, X_2, \ldots$ obeys the functional law of the iterated logarithm.*

Under the conditions of the corollary the ordinary law of the iterated logarithm was proved by TAKAHASHI and RÉVÉSZ ([4], [5]).

2. Proof of the theorem

The proof is based on a previous result of ours (see [7], Theorem 1). Besides, we are going to use the following

LEMMA. *Let $Y_1, Y_2, \ldots$ be a weakly multiplicative system such that $|Y_i| \leqq M$ $(i=1, 2, \ldots)$. Then for every sequence $a_1, a_2, \ldots$ of real numbers and for every $N \geqq 1$, $\lambda>0$, $t>0$ we have*

$$\mathsf{E}(\exp\{\lambda|a_1 Y_1+\ldots+a_N Y_N|\}) \leqq C_1 \exp\{C_2 \lambda^2 (a_1^2+\ldots+a_N^2)\} \tag{2.1}$$

and

$$\mathsf{P}(|a_1 Y_1+\ldots+a_N Y_N| \geqq t\sqrt{a_1^2+\ldots+a_N^2}) \leqq C_3 \exp(-C_4 t^2) \tag{2.2}$$

where C_1, C_2, C_3, C_4 are positive constants depending only on M.

For the proof of (2.1) see [3]; (2.2) follows from (2.1) by a routine application of the Markov-inequality.

Let now $X_1, X_2, \ldots$ be a sequence of r.v.-s satisfying the conditions of the theorem, let K_1 denote a positive constant such that $|X_n| \leqq K_1$ $(n=1, 2, \ldots)$. Put $S_n = X_1+\ldots+X_n$ $(S_0=0)$ and let us introduce the piecewise constant random function $\tilde{\varphi}_n(t)=n^{-1/2}S_{[nt]}$ $(0 \leqq t \leqq 1)$. Let $0 \leqq t_1<t_1' \leqq \ldots \leqq t_r<t_r' \leqq 1$ be arbitrary numbers and let $f_n^{(t_1, t_1', \ldots, t_r, t_r')}(\lambda_1, \ldots, \lambda_r)$ denote the characteristic function of the random vector

$$\left(\frac{\tilde{\varphi}_n(t_1')-\tilde{\varphi}_n(t_1)}{\sqrt{t_1'-t_1}}, \ \ldots, \ \frac{\tilde{\varphi}_n(t_r')-\tilde{\varphi}_n(t_r)}{\sqrt{t_r'-t_r}}\right).$$

Let us finally introduce the notation $\|\lambda\|$ for the absolute value of an r-dimensional vector $\lambda=(\lambda_1, \ldots, \lambda_r)$, i.e. put $\|\lambda\| = (\lambda_1^2+\ldots+\lambda_r^2)^{1/2}$.

We shall prove that for every $0 \leqq t_1<t_1' \leqq \ldots \leqq t_r<t_r' \leqq 1$, $n \geqq 1$ and for every $\lambda=(\lambda_1, \ldots, \lambda_r)$ satisfying $\|\lambda\| \leqq C_5(nt_0)^{1/10}$ we have

$$\left|f_n^{(t_1, \ldots, t_r')}(\lambda_1, \ldots, \lambda_r)-\exp\left\{-\frac{1}{2}\sum_{k=1}^{r}\lambda_k^2\right\}\right| \leqq \frac{C_6}{(nt_0)^{1/20}} \tag{2.3}$$

where $t_0 = \min(t_1'-t_1, \ldots, t_r'-t_r)$, C_5 and C_6 are positive constants depending only on r and the sequence $X_1, X_2, \ldots$. By Theorem 1 of [7], this statement implies our theorem.

To prove (2.3), let us first write the characteristic function $f_n^{(t_1,\dots,t_r)}(\lambda_1,\dots,\lambda_r)$ in a different form. Making use of the formula $e^z = (1+z)\exp\left(\frac{z^2}{2}+r(z)\right)$ where $|r(z)|\leqq C_7|z|^3$ for $|z|\leqq C_8$ (C_7, C_8 are absolute constants) we obtain

$$\exp\left(\frac{i\lambda_k}{\sqrt{n(t'_k-t_k)}}X_v\right)=$$

$$=\left(1+\frac{i\lambda_k}{\sqrt{n(t'_k-t_k)}}X_v\right)\exp\left\{-\frac{\lambda_k^2}{2n(t'_k-t_k)}X_v^2+r\left(\frac{i\lambda_k}{\sqrt{n(t'_k-t_k)}}X_v\right)\right\}$$

$$(1\leqq k\leqq r,\ [nt_k]+1\leqq v\leqq [nt'_k])$$

and thus

$$f_n^{(t_1,\dots,t'_r)}(\lambda_1,\dots,\lambda_r)=\mathsf{E}\left(\exp\left\{\sum_{k=1}^{r}\frac{i\lambda_k}{\sqrt{n(t'_k-t_k)}}(X_{[nt_k]+1}+\dots+X_{[nt'_k]})\right\}\right)=$$

$$=\mathsf{E}\left(\prod_{k=1}^{r}\prod_{v=[nt_k]+1}^{[nt'_k]}\left(1+\frac{i\lambda_k}{\sqrt{n(t'_k-t_k)}}X_v\right)\exp\left\{-\sum_{k=1}^{r}\frac{\lambda_k^2}{2n(t'_k-t_k)}P_k+\sum_{k=1}^{r}Q_k\right\}\right)$$

where

$$P_k=\sum_{v=[nt_k]+1}^{[nt'_k]}X_v^2$$

$$Q_k=\sum_{v=[nt_k]+1}^{[nt'_k]}r\left(\frac{i\lambda_k}{\sqrt{n(t'_k-t_k)}}X_v\right).$$

Therefore

(2.4)
$$f_n^{(t_1,\dots,t'_r)}(\lambda_1,\dots,\lambda_r)-\exp\left(-\frac{1}{2}\sum_{k=1}^{r}\lambda_k^2\right)=$$

$$=\exp\left(-\frac{1}{2}\sum_{k=1}^{r}\lambda_k^2\right)\mathsf{E}\left(\prod_{k=1}^{r}\prod_{v=[nt_k]+1}^{[nt'_k]}\left(1+\frac{i\lambda_k}{\sqrt{n(t'_k-t_k)}}X_v\right)-1\right)+$$

$$+\mathsf{E}\left(\prod_{k=1}^{r}\prod_{v=[nt_k]+1}^{[nt'_k]}\left(1+\frac{i\lambda_k}{\sqrt{n(t'_k-t_k)}}X_v\right)\left\{\exp\left(-\sum_{k=1}^{r}\frac{\lambda_k^2}{2n(t'_k-t_k)}P_k+\sum_{k=1}^{r}Q_k\right)-\right.\right.$$

$$\left.\left.-\exp\left(-\frac{1}{2}\sum_{k=1}^{r}\lambda_k^2\right)\right\}\right)=J_1+J_2.$$

The expression J_1 can easily be estimated. Indeed,

$$\mathsf{E}\left(\prod_{k=1}^{r}\prod_{v=[nt_k]+1}^{[nt'_k]}\left(1+\frac{i\lambda_k}{\sqrt{n(t'_k-t_k)}}X_v\right)-1\right)=\sum L_{v_1\dots v_s}\mathsf{E}(X_{v_1}\dots X_{v_s})$$

where the sum consists of finitely many terms and $L_{v_1\dots v_s}$ are complex constants.

It is obvious that $\left|\frac{i\lambda_k}{\sqrt{n(t_k'-t_k)}}\right| \leqq \frac{\|\lambda\|}{\sqrt{nt_0}}$ for $1 \leqq k \leqq r$, therefore we have $|L_{\nu_1 \dots \nu_s}| \leqq$ $\leqq \left(\frac{\|\lambda\|}{\sqrt{nt_0}}\right)^s \leqq \left(\frac{1}{(nt_0)^{4/10}}\right)^s \leqq \frac{1}{(nt_0)^{4/10}}$ provided that $\|\lambda\| \leqq (nt_0)^{1/10}$, $nt_0 \geqq 1$. Using this fact, we obtain (under the above conditions) that

$$|J_1| \leqq \exp\left(-\frac{1}{2}\sum_{k=1}^{r}\lambda_k^2\right)\frac{A}{(nt_0)^{4/10}} \leqq \frac{A}{(nt_0)^{4/10}} \tag{2.5}$$

where A is the sum of the series appearing in (1.1) Making use of the inequality $1+x \leqq e^x$, the estimation of J_2 proceeds as follows:

$$|J_2| \leqq \mathsf{E}\left(\left|\prod_{k=1}^{r}\prod_{\nu=[nt_k]+1}^{[nt_k']}\left(1+\frac{i\lambda_k}{\sqrt{n(t_k'-t_k)}}X_\nu\right)\right| \times \right. \tag{2.6}$$

$$\left. \times \left|\exp\left(-\sum_{k=1}^{r}\frac{\lambda_k^2}{2n(t_k'-t_k)}P_k + \sum_{k=1}^{r}Q_k\right) - \exp\left(-\frac{1}{2}\sum_{k=1}^{r}\lambda_k^2\right)\right|\right) =$$

$$= \mathsf{E}\left(\prod_{k=1}^{r}\prod_{\nu=[nt_k]+1}^{[nt_k']}\left(1+\frac{\lambda_k^2}{n(t_k'-t_k)}X_\nu^2\right)^{1/2}\left|\exp\left(-\sum_{k=1}^{r}\frac{\lambda_k^2}{2n(t_k'-t_k)}P_k + \sum_{k=1}^{r}Q_k\right) - \right.\right.$$

$$\left.\left. - \exp\left(-\frac{1}{2}\sum_{k=1}^{r}\lambda_k^2\right)\right|\right) \leqq$$

$$\leqq \mathsf{E}\left(\exp\sum_{k=1}^{r}\frac{\lambda_k^2}{2n(t_k'-t_k)}P_k\left|\exp\left(-\sum_{k=1}^{r}\frac{\lambda_k^2}{2n(t_k'-t_k)}P_k + \sum_{k=1}^{r}Q_k\right) - \exp\left(-\frac{1}{2}\sum_{k=1}^{r}\lambda_k^2\right)\right|\right) =$$

$$= \mathsf{E}\left(\left|\exp\left(\sum_{k=1}^{r}Q_k\right) - \exp\left\{\sum_{k=1}^{r}\frac{\lambda_k^2}{2n(t_k'-t_k)}(P_k - n(t_k'-t_k))\right\}\right|\right) \leqq J_3 + J_4$$

where

$$J_3 = \mathsf{E}\left(\left|\exp\left(\sum_{k=1}^{r}Q_k\right) - 1\right|\right)$$

$$J_4 = \mathsf{E}\left(\left|\exp\left\{\sum_{k=1}^{r}\frac{\lambda_k^2}{2n(t_k'-t_k)}(P_k - n(t_k'-t_k))\right\} - 1\right|\right).$$

By the assumptions we have $|X_\nu| \leqq K_1$, thus $\frac{\|\lambda\|}{\sqrt{nt_0}}K_1 \leqq C_8$ (i.e. $\|\lambda\| \leqq C_9\sqrt{nt_0}$) implies that the inequality $\left|\frac{i\lambda_k}{\sqrt{n(t_k'-t_k)}}X_\nu\right| \leqq C_8$ holds for every k, ν. On the other hand,

$nt_0 \geqq 2$ implies $[nt_k'] - [nt_k] \leqq n(t_k' - t_k) + 2 \leqq 2n(t_k' - t_k)$ and therefore we obtain for $\|\lambda\| \leqq C_9 \sqrt{nt_0}, \quad nt_0 \geqq 2$

$$(2.7) \qquad \left| \sum_{k=1}^{r} Q_k \right| \leqq \sum_{k=1}^{r} \sum_{v=[nt_k]+1}^{[nt_k']} \left| r\left(\frac{i\lambda_k}{\sqrt{n(t_k' - t_k)}} X_v \right) \right| \leqq$$

$$\leqq \sum_{k=1}^{r} \sum_{v=[nt_k]+1}^{[nt_k']} C_7 \left| \frac{i\lambda_k}{\sqrt{n(t_k' - t_k)}} X_v \right|^3 \leqq \sum_{k=1}^{r} ([nt_k'] - [nt_k]) C_7 \frac{\|\lambda\|^3 K_1^3}{(n(t_k' - t_k))^{3/2}} \leqq$$

$$\leqq \sum_{k=1}^{r} \frac{2C_7 \|\lambda\|^3 K_1^3}{(n(t_k' - t_k))^{1/2}} \leqq C_{10} \frac{\|\lambda\|^3}{(nt_0)^{1/2}}$$

where C_{10} depends on K_1, r. In the sequel the constants $C_{11}, C_{12}, \ldots$ will denote positive numbers depending (at most) on r and the sequence $X_1, X_2, \ldots$. Using now (2.7) and the inequality $|e^z - 1| \leqq 2|z|$, valid for $|z| \leqq C_{11}$, we obtain

$$(2.8) \qquad J_3 \leqq 2\mathsf{E}\left(\left| \sum_{k=1}^{r} Q_k \right| \right) \leqq 2C_{10} \frac{\|\lambda\|^3}{(nt_0)^{1/2}} \leqq 2C_{10} \frac{1}{(nt_0)^{2/10}}$$

provided that

$$\left| \sum_{k=1}^{r} Q_k \right| \leqq C_{11}, \quad \|\lambda\| \leqq C_9 \sqrt{nt_0}, \quad nt_0 \geqq 2, \quad \|\lambda\| \leqq (nt_0)^{1/10}.$$

According to (2.7), each of these conditions is satisfied if

$$\|\lambda\| \leqq C_{12}(nt_0)^{1/10}, \quad nt_0 \geqq 2.$$

To estimate J_4, we first note that

$$(2.9) \quad \mathsf{P}(|P_k - ([nt_k'] - [nt_k])| \geqq (n(t_k' - t_k))^{3/4}) \leqq C_{13} \exp\left(-C_{14}\sqrt{nt_0}\right) \quad (1 \leqq k \leqq r)$$

for $nt_0 \geqq 4$. In fact, by the assumptions of the theorem the sequence $X_n^2 - 1$ is a uniformly bounded weakly multiplicative system and thus the lemma gives

$$\mathsf{P}(|P_k - ([nt_k'] - [nt_k])| \geqq (n(t_k' - t_k))^{3/4}) = \mathsf{P}\left(\left| \sum_{v=[nt_k]+1}^{[nt_k']} (X_v^2 - 1) \right| \geqq (n(t_k' - t_k))^{3/4} \right) \leqq$$

$$\leqq \mathsf{P}\left(\left| \sum_{v=[nt_k]+1}^{[nt_k']} (X_v^2 - 1) \right| \geqq \frac{1}{2^{3/4}} ([nt_k'] - [nt_k])^{3/4} \right) \leqq C_{15} \exp\left(-C_{16}([nt_k'] - [nt_k])^{1/2}\right) \leqq$$

$$\leqq C_{15} \exp\left(-C_{17}(n(t_k' - t_k))^{1/2}\right) \leqq C_{15} \exp\left(-C_{17}(nt_0)^{1/2}\right)$$

provided that $nt_0 \geqq 4$. Let us now introduce the events

$$D_k = \left\{ |P_k - n(t_k' - t_k)| < 2(n(t_k' - t_k))^{3/4} \right\} \qquad k = 1, 2, \ldots, r$$

$$D = D_1 D_2 \ldots D_r$$

and let us split J_4 into two summands as follows: $J_4 = J_4^{(1)} + J_4^{(2)}$, where

$$J_4^{(1)} = \int_D |e^U - 1| \, d\mathsf{P} \quad J_4^{(2)} = \int_D |e^U - 1| \, d\mathsf{P}$$

$$U = \sum_{k=1}^{r} \frac{\lambda_k^2}{2n(t_k' - t_k)} \left(P_k - n(t_k' - t_k)\right)$$

On the set D we have

$$|U| \leqq \sum_{k=1}^{r} \frac{\lambda_k^2}{2n(t_k' - t_k)} 2\left(n(t_k' - t_k)\right)^{3/4} = \sum_{k=1}^{r} \frac{\lambda_k^2}{\left(n(t_k' - t_k)\right)^{1/4}} \leqq \frac{\|\lambda\|^2}{(nt_0)^{1/4}} \tag{2.10}$$

and hence we obtain (using the inequality $|e^z - 1| \leqq 2|z|$ again)

$$J_4^{(1)} \leqq \int_D 2|U| \, d\mathsf{P} \leqq \frac{2\|\lambda\|^2}{(nt_0)^{1/4}} \leqq \frac{2}{(nt_0)^{1/20}} \tag{2.11}$$

provided that $|U| \leqq C_{11}$, $\|\lambda\| \leqq (nt_0)^{1/10}$. According to (2.10), both of these conditions are satisfied if $\|\lambda\| \leqq C_{18}(nt_0)^{1/10}$, $nt_0 \geqq 1$. On the other hand, on the set $\overline{D}_k$ we have $|P_k - ([nt_k'] - [nt_k])| \geqq |P_k - n(t_k' - t_k)| - 2 \geqq 2\left(n(t_k' - t_k)\right)^{3/4} - 2 \geqq \left(n(t_k' - t_k)\right)^{3/4}$ (provided that $nt_0 \geqq 3$), thus (2.9) gives $\mathsf{P}(\overline{D}) \leqq \sum_{k=1}^{r} \mathsf{P}(\overline{D}_k) \leqq C_{19} \exp\left(-C_{14}\sqrt{nt_0}\right)$. Using this fact, we obtain for $\|\lambda\| \leqq (nt_0)^{1/10}$, $nt_0 \geqq C_{20}$:

$$J_4^{(2)} \leqq \int_{\overline{D}} (e^U + 1) \, d\mathsf{P} \leqq \int_{\overline{D}} \left(\exp\left\{ \sum_{k=1}^{r} \frac{\lambda_k^2}{2n(t_k' - t_k)} P_k \right\} + 1 \right) d\mathsf{P} \leqq \tag{2.12}$$

$$\leqq C_{19} \exp\left(-C_{14}\sqrt{nt_0}\right) \left[\exp\left\{ \sum_{k=1}^{r} \frac{\lambda_k^2}{2n(t_k' - t_k)} \left([nt_k'] - [nt_k]\right) K_1^2 \right\} + 1 \right] \leqq$$

$$\leqq C_{19} \exp\left(-C_{14}\sqrt{nt_0}\right) \left[\exp\left(C_{21} \sum_{k=1}^{r} \lambda_k^2 \right) + 1 \right] \leqq$$

$$\leqq 2C_{19} \exp\left(-C_{14}\sqrt{nt_0} + C_{21}(nt_0)^{1/5}\right) \leqq 2C_{19} \exp\left(-C_{22}\sqrt{nt_0}\right) \leqq \frac{C_{23}}{(nt_0)^{1/2}}.$$

From the estimates (2.11) and (2.12) we finally get

$$J_4 \leqq \frac{2}{(nt_0)^{1/20}} + \frac{C_{23}}{(nt_0)^{1/2}} \leqq \frac{C_{24}}{(nt_0)^{1/20}}. \tag{2.13}$$

It is easy now to complete the proof of the theorem. Let us assume that $\|\lambda\| \leqq C_{25}(nt_0)^{1/10}$, $nt_0 \geqq C_{26}$, then (2.5), (2.8) and (2.13) are valid and thus (2.4)—(2.13) give

$$\left| f_n^{(t_1, \ldots, t_r')}(\lambda_1, \ldots, \lambda_r) - \exp\left(-\frac{1}{2} \sum_{k=1}^{r} \lambda_k^2 \right) \right| \leqq \frac{C_{27}}{(nt_0)^{1/20}}. \tag{2.14}$$

On the other hand, it is possible to choose the constant C_{27} so large that $C_{27}(nt_0)^{-1/20} > 2$ holds for $nt_0 < C_{26}$ and in this case (2.14) will be valid also for $nt_0 < C_{26}$ since the left side is at most 2 and the right side is greater than 2. Hence (2.3) is valid and the proof of our theorem is completed.

Remark. Using relation (2.3), the second statement of the lemma and a well-known tightness criterion ([6] p. 59 and p. 94) we could easily deduce an invariance principle for the sequence $X_1, X_2, \ldots$. We do not present the details here because a more general invariance principle has been proved recently by McLEISH (see [2]).

REFERENCES

[1] GAPOSKIN, V. F.: The central limit theorem for strongly multiplicative systems of functions, *Mat. Zametki* **6** (1969) 443—450. (In Russian).
[2] MC LEISH, D. L.: Dependent central limit theorems and invariance principles, Thesis, 1972. University of Montreal.
[3] RÉVÉSZ, P.: A new law of the iterated logarithm for multiplicative systems and its applications *Acta Sci. Math. Hung.* (to appear).
[4] TAKAHASHI, S.: Notes on the law of the iterated logarithm, *Studia Sci. Math. Hung.* **7** (1972) 21—24.
[5] RÉVÉSZ, P.: The law of the iterated logarithm for multiplicative systems, *Indiana Univ. Math. J.* **21** (1972) 557—564.
[6] BILLINGSLEY, P.: *Convergence of probability measures*, New York: Wiley, 1968.
[7] BERKES, I.: The functional law of the iterated logarithm for dependent random variables, *Z. Wahrscheinlichkeitstheorie verw. Gebiete.* (to appear).

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received April 27, 1973)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 433—436.

INEQUALITIES OBTAINED ON THE BASIS OF THE SPECTRUM OF THE GRAPH

by

DRAGOŠ M. CVETKOVIĆ

Inequalities for the number of interior stability, for the number of vertices of the maximal complete subgraph and for the chromatic number of the graph are obtained on the basis of the spectrum of the graph.

We consider only finite, undirected graphs without loops or multiple edges.

Let G be a graph with vertices: $x_1, \ldots, x_n$. The adjacency matrix of G is the matrix $A=\|a_{ij}\|_1^n$, where $a_{ij}=1$ if x_i and x_j are adjacent vertices and $a_{ij}=0$ in the opposite case. The spectrum of G is the set of eigenvalues of A. Since A is, for the considered class of graphs, a symmetric matrix, the spectrum is a set of real numbers. The largest eigenvalue r in the spectrum of a graph is called the index of the graph. The whole spectrum lies on the segment $[-r, r]$.

Recently, dozens of papers dealing with connectivity between the graph spectrum and graph structure have been published. The fundamental one among the first papers is [1]. Papers [3]—[6], [9] also belong to this group.

Let us take the following theorem (see, for example [2], p. 119) as a starting point.

THEOREM 1. *Let A be a hermitian matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ $(\lambda_1 \geqq \lambda_2 \geqq \ldots \ldots \geqq \lambda_n)$ and B one of its principal submatrices; B has the eigenvalues $\mu_1, \ldots, \mu_m$ $(\mu_1 \geqq \mu_2 \geqq \ldots \geqq \mu_m)$. Then, the inequalities*

$$\lambda_{n-m+i} \leqq \mu_i \leqq \lambda_i \qquad (i=1, \ldots, m)$$

hold.

The number of interior stability $\alpha(G)$ of the graph G is equal to the maximal number of vertices which can be chosen in G so that no pair of them is joined with an edge of G.

THEOREM 2. *The number of interior stability $\alpha(G)$ of the graph G satisfies the inequality*

$$\alpha(G) \leqq p_0 + \min(p_-, p_+), \tag{1}$$

where p_-, p_0, p_+ denote the number of numbers smaller, equal and greater than zero, in the spectrum of G.

There are graphs for which equality holds in (1).

PROOF. Let $s = p_0 + \min(p_-, p_+)$. Suppose that there is a graph G for which $\alpha(G) > s$ holds. Then, there is a subgraph of G with $\alpha(G)$ vertices not containing

edges of G. Thus, a principal submatrix, of the order $\alpha(G)$, of the adjacency matrix of G is equal to zero-matrix. Since, the spectrum of a zero-matrix contains only numbers equal to zero, Theorem 1 gives for the spectrum $\{\lambda_1, \ldots, \lambda_n\}$ $(\lambda_1 \geqq \lambda_2 \geqq \ldots \geqq \lambda_n)$ of G inequalities: $\lambda_i \geqq 0$, $\lambda_{n-\alpha(G)+i} \leqq 0$ $(i=1, \ldots, \alpha(G))$. However, this cannot be fulfilled since $\alpha(G) > s$. Thus, (1) holds.

Equality holds in (1), for example, for complete graphs.

This completes the proof of Theorem 2.

A special case of this result is noted in [3]. This paper deals with an adjacency matrix of a somewhat different shape.

In proving the following theorem one uses the spectrum of a complete graph. It is known, that the spectrum of a complete graph with k vertices contains the number $k-1$ as well as $k-1$ numbers equal to -1 (see, for example, [1]).

THEOREM 3. *Let $p_{\lambda<-1}$, p_{-1} and $p_{\lambda>-1}$ denote the number of eigenvalues which are smaller, equal and greater than -1 in the spectrum of the graph G. λ represents the smallest eigenvalue greater than -1. Let us further have $p = p_{\lambda<-1}+p_{-1}+1$ and $s = \min(p, p_{\lambda>-1}+p_{-1}, r+1)$. If $K(G)$ denotes the number of vertices of the maximal complete subgraph of the graph G, then inequality*

$$K(G) \leqq \begin{cases} s, & s < p \\ s-\alpha_\lambda, & s = p, \end{cases} \tag{2}$$

holds, where

$$\alpha_\lambda = \begin{cases} 0, & \lambda \leqq p-1 \\ 1, & \lambda > p-1. \end{cases}$$

PROOF. If G contains a complete subgraph on k vertices, then Theorem 1, in a similar way as in the proof of Theorem 2, gives for the spectrum of G the following inequalities:

$$\lambda_{n-k+1} \leqq k-1 \leqq \lambda_1 = r,$$

$$\lambda_{n-k+i} \leqq -1 \leqq \lambda_i \qquad (i = 2, \ldots, k).$$

The greatest value of k, satisfying these inequalities, is given by the expression on the right side of inequality (2).

This completes the proof of Theorem 3.

Paper [4] deals, among other things, with the relation between the spectrum and the number of vertices of maximal complete subgraph. The proofs of the results mentioned in [4] are not published, as far as we know.

Now we shall derive a lower bound for the chromatic number $\gamma(G)$ of the graph G.

Note that in [5] the inequality $\gamma(G) \leqq r+1$ is derived. In [6] some inequalities for $\gamma(G)$ are derived for regular graphs, the complements of which are unconnected graphs.

First, we shall give the definition of the sum of graphs.

The sum G_1+G_2 of graphs G_1 and G_2 is a graph having, as the set of vertices, the Cartesian product of the sets of vertices of graphs G_1 and G_2, where the adjacency

relation of vertices is defined in the following way. The vertices (x, y) and (u, v) of G_1+G_2 are adjacent if and only if either $x=u$ and y is adjacent to v in G_2 or if x is adjacent to u in G_1 and $y=v$.

This opertaion is called the Cartesian product of graphs in [7] and the term "sum" is mentioned according to [8].

The following theorem is proved in [7].

THEOREM 4. *If G is a graph with n vertices and if K_p is a complete graph with p vertices, then the relation*

$$\gamma(G) = \min \{p | \alpha(G+K_p) = n\}$$

holds.

Among other things, the spectrum of the sum of graphs is determined in [9]. Let the graphs G_1 and G_2 have m_1 and m_2 vertices and let $\{\lambda_{ji_j} | i_j=1, \ldots, m_j\}$, $j=1, 2$, be their spectrums. Then the sum G_1+G_2 has the spectrum

$$\{\lambda_{1i_1}+\lambda_{2i_2} | i_j = 1, \ldots, m_j; j = 1, 2\}.$$

THEOREM 5. *For a given k let a, b, c, d, e denote the numbers of eigenvalues λ in the spectrum $\{\lambda_1, \ldots, \lambda_n\}$ of the graph G which satisfy the following relations respectively:*

$$\lambda < -k+1, \qquad \lambda = -k+1, \quad -k+1 < \lambda < 1, \quad \lambda = 1, \quad \lambda > 1.$$

Further, let s be the smallest of the natural numbers k $(k>1)$ for which the inequality

$$\min [b+c+k(d+e), k(a+b)+(k-1)(c+d)] \geqq n \tag{3}$$

is satisfied. Then

$$\gamma(G) \geqq s$$

holds.

PROOF. According to Theorem 4 we have

$$\gamma(G) = q = \min \{p | \alpha(G+K_p) = n\} \tag{4}$$

Let now p_-, p_0 and p_+ denote the numbers of negative, equal to zero and positive eigenvalues in the spectrum of $G+K_q$. According to what has already been said for the spectrum of the sum of graphs, we have $p_+ = c+d+e+(q-1)e = c+d+qe$, $p_0 = b+(q-1)d$, $p_- = a+(q-1)(a+b+c)$. According to Theorem 2 and to (4), inequality

$$\min [b+c+q(d+e), q(a+b)+(q-1)(c+d)] \geqq n \tag{5}$$

must be valid. The minimal value of q for which (5) can hold is equals to s. This completes the proof of Theorem 5.

REFERENCES

[1] COLLATZ, L.—SINOGOWITZ, U.: Spektren endlicher Graphen, *Abh. Math. Sem. Univ. Hamburg*, **21** (1957), 63—77.
[2] MARCUS, M.—MINC, H.: *A survey of matrix theory and matrix inequalities*, Boston *1964*.
[3] ВАХОВСКИ, Е. В.: О характеристических числах матриц сосевства для неособенных графов, *Сибир. матем. журн.* **6.** (1965), №1, 44—49.
[4] ЛИХТЕНБАУМ, Л. М.: Характеристические числа неособенного графа, *Труды 3-го Всес. матем. съезда*, **1**, 1956, 135—136.
[5] WILF, H. S.: The eigenvalues of a graph and its chromatic number, *J. London Math. Soc.* **42** (1967), 330—332.
[6] FINCK, H.—J.: Vollständiges Produkt, chromatische Zahl und characteristisches Polynom regulärer Graphen II, *Wiss. Z. Techn. Hochsch. Ilmenau* **11** (1965), 81—87.
[7] ВИЗИНГ, В.—Г. ПЛЕСНЕВИЧ, Г. С.: К проблеме минимальной раскраски вершин графа, *Сибир, матем, журн.* **6** (1965), № 1, 234—236.
[8] BERGE, C.: *Théorie des graphes et ses applications*, deuxiéme édition, Paris 1963.
[9] CVETKOVIČ, D.: Spectrum of the graph of *n*-tuples, *Univ. Beograd Publ. Elektrotehn. Fak. Ser. Mat. Fiz.* №274—№301 (1969). 91—95.

University of Beograd

(Received March 1, 1971)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 437—446.

APPROXIMATION OF CONTINUOUS FUNCTIONS BY POLINOMIALS INVOLVING FREUD'S INTERPOLATORY POLYNOMIALS

by

R. B. SAXENA

1. Introduction

In 1964 G. FREUD [1] constructed the interpolatory polynomial

$$J_n(f, x) = f(0) + \sum_{k=1}^{n} \Phi_{kn}(x)[f(x_{kn}) - f(0)] \tag{1.1}$$

of degree atmost $4n-3$, which coincides with the function $f(x)$ on the roots of $T_n(x)$ the Chebyshev polynomial of first kind, viz., the nodes

$$x_{kn} = \cos\frac{2k-1}{2n}\pi, \qquad k = 1, 2, \ldots, n. \tag{1.2}$$

Here

$$\Phi_{kn}(x) = v_{kn}(x) l_{kn}^4(x) + 2(x - x_{kn}) l_{kn}^3(x) \psi_n(x_{kn}, x), \tag{1.3}$$

where

$$v_{kn}(x) = 1 - \frac{x_{kn}}{1 - x_{kn}^2}(x - x_{kn}),$$

$$\psi_n(u, v) = \frac{2}{n} \sum_{r=1}^{n-1} T_r'(u) T_r(v),$$

and

$$l_{kn}(x) = \frac{(-1)^{k+1}\sqrt{1 - x_{kn}^2}}{n} \cdot \frac{T_n(x)}{x - x_{kn}},$$

are the fundamental polynomials of Lagrange interpolation. FREUD [1] showed that, for an absolute constant c,

$$|f(x) - J_n(f, x)| \leqq c\omega\left(f, \frac{1}{4n}\right), \quad -\frac{1}{2} \leqq x \leqq \frac{1}{2}, \tag{1.4}$$

where $\omega(f, \delta)$ is the modulus of continuity of $f(x)$. This method was modified by the author [4] to the case when the nodes are the zeros of Chebyshev polynomial of second kind $U_n(x)$. We replaced (1.1) by the corresponding formula

(1.5)

$$A_n(f, x) = \frac{1+x}{2} f(1) + \frac{1-x}{2} f(-1) + \sum_{k=1}^{n} \left[f(\xi_{kn}) - \left\{ \frac{1+x}{2} f(1) + \frac{1-x}{2} f(-1) \right\} \right] \lambda_{kn}(x)$$

built on the nodes

$$\xi_{kn} = \cos\frac{k\pi}{n}, \qquad k = 0, 1, \ldots, n \tag{1.6}$$

and proved [4] — as an improvement to (1.4) — the inequality

$$|f(x) - \Lambda_n(f, x)| \leqq 414\omega\left(f, \frac{1}{n}\right), \quad -1 \leqq x \leqq 1. \tag{1.7}$$

P. VÉRTESI and G. FREUD [2] used the formula (1.5) on the nodes (1.2). Denoting the corresponding polynomials by

$$J_n^*(f, x) = \frac{1+x}{2}f(1) + \frac{1-x}{2}f(-1) + \tag{1.8}$$

$$+ \sum_{k=1}^{n}\left[f(x_{kn}) - \left\{\frac{1+x}{2}f(1) + \frac{1-x}{2}f(-1)\right\}\right]\varphi_{kn}(x),$$

they proved the inequality [2]

$$|f(x) - J_n^*(f, x)| \leqq c\left[\omega\left(f; \frac{\sqrt{1-x^2}}{4n}\right) + \omega\left(f; \frac{1}{16n^2}\right)\right], \quad -1 \leqq x \leqq 1, \tag{1.9}$$

due to A. F. TIMAN [6]. Later, by using the same polynomials $\Lambda_n(f, x)$ in (1.5) the author [5] improved upon (1.9) by proving the inequality

$$|f(x) - \Lambda_n(f, x)| \leqq 840\omega\left(f, \frac{\sqrt{1-x^2}}{n}\right), \quad \text{for} \quad x \in [-1, 1]. \tag{1.10}$$

Our polynomials $\Lambda_n(f, x)$ are of degree atmost $4n+2$ which take the same values as $f(x)$ at the nodes (1.6). The inequality (1.10) was first given by S. A. TELYAKOVSKII [7] and I. E. GOPENGAUZ [3].

We now enquire whether same order of approximation, as in (1.10), can be obtained using the polynomials $J_n^*(f, x)$ instead of $\Lambda_n(f, x)$? The answer is in the negative as $J_n^*(f, \pm 1) \neq f(\pm 1)$. In this paper, we construct the polynomials $g_n(f, x)$ which involve the interpolatory polynomials $J_n^*(f, x)$ and, as we shall show, approximate any function $f(x)$ continuous on $[-1, 1]$ with an error of order $\omega\left(f, \frac{\sqrt{1-x^2}}{n}\right)$ for $x \in [-1, 1]$. In particular, we shall prove the following

THEOREM. *For all functions $f(x)$, continuous on $[-1, 1]$, and any natural number n, the following inequality holds for all $x \in [-1, 1]$:*

$$|f(x) - g_n(f, x)| \leqq c_1\omega\left(f; \frac{\sqrt{1-x^2}}{n}\right), \tag{1.11}$$

where c_1 is an absolute constant.

We remark that our polynomials $g_n(f, x)$ are no longer interpolatory polynomials. The corresponding problem for interpolatory polynomials on the nodes (1.2) appears to be of some interest.

2. The operator $g_n(f, x)$

For all functions $f(x)$, continuous on $[-1, 1]$ we have

$$r_n(f) = f - J_n^*(f) \tag{2.1}$$

where $J_n^*(f)$ is the interpolating proces (1.8) considered by G. FREUD and P. VÉRTESI [2];

$$L(f, x) = \frac{1+x}{2} f(1) + \frac{1-x}{2} f(-1); \tag{2.2}$$

$$g_n(f) = J_n^*(f) + L(r_n(f)). \tag{2.3}$$

It is obvious that $g_n(f, x)$ are algebraic polynomials of degree not exceding $4n-2$. In addition

$$g_n(f; \pm 1) = f(\pm 1).$$

3. Modulus of continuity and auxiliary results

We shall use the following properties of modulus of continuity.

(i) $\omega(\delta_1) \leqq \omega(\delta_2), \quad 0 < \delta_1 < \delta_2$

(ii) $\omega(\delta_1 + \delta_2) \leqq \omega(\delta_1) + \omega(\delta_2),$

(iii) $\omega(n\delta) \leqq n\omega(\delta),$ n a positive integer

(iv) $\dfrac{\omega(\delta_1)}{\delta_1} \leqq 2\dfrac{\omega(\delta_2)}{\delta_2}, \quad \delta_1 \geqq \delta_2 > 0.$

From these we at once have

$$\frac{1-x}{2}\omega(2) \leqq (1-x)\omega(1) \leqq 2\omega(1-x) \qquad x \geqq 0; \tag{3.1}$$

$$\frac{1+x}{2}\omega(2) \leqq \omega(1+x)\omega(1) \leqq 2\omega(1+x) \qquad x \leqq 0; \tag{3.2}$$

$$(1-x^2)\omega(1) \leqq 2\omega(1-x^2) \quad \text{for} \quad x \in [-1, 1]; \tag{3.3}$$

$$\omega(1-x_{kn}) \leqq \begin{cases} \omega(1-x) & x_{kn} \geqq x \\ 2\omega(1-x) + 2\dfrac{x-x_{kn}}{1-x}\omega(1-x), & x_{kn} \leqq x \end{cases} \tag{3.4}$$

and

$$\omega(1+x_{kn}) \leqq \begin{cases} \omega(1+x) & x_{kn} \leqq x \\ 2\omega(1+x) + 2\dfrac{x_{kn}-x}{1+x}\omega(1+x), & x_{kn} \geqq x. \end{cases} \tag{3.5}$$

4. Auxiliary lemmas

LEMMA 1. *For the polanomials* $\varphi_{kn}(x)$ *in* (1.3) *we have for* $-1 \leqq x \leqq 1$,

$$\left| 1 - \sum_{k=1}^{n} \varphi_{kn}(x) \right| \leqq 3, \tag{4.1}$$

$$\sum_{k=1}^{n} |\varphi_{kn}(x)| \leqq 142, \tag{4.2}$$

$$\sum_{k=1}^{n} |x - x_{kn}| |\varphi'_{kn}(x)| \leqq 1584. \tag{4.3}$$

The estimates (4.1) and (4.2) have been proved by P. VÉRTESI [8]. We prove (4.3). From (1.3) we have[1]

$$\varphi_{kn}(x) = \frac{(1 - xx_k)(1 - x_k^2)}{n^4} \cdot \frac{T_n^4(x)}{(x - x_k)^4} +$$

$$+ \frac{4(-1)^{k+1}(1 - x_k^2)\frac{3}{2}}{n^4} \cdot \frac{T_n^3(x)}{(x - x_k)^2} \sum_{r=1}^{n-1} T'_r(x_k) T_r(x).$$

On putting $x = \cos\theta$ and $x_k = \cos\theta_k$, we have

$$\varphi_{kn}(x) = \frac{(1 - \cos\theta\cos\theta_k)\sin^2\theta_k \cos^4 n\theta}{n^4(\cos\theta - \cos\theta_k)^4} + \tag{4.4}$$

$$+ \frac{4(-1)^{k+1}\sin^2\theta_k \cos^3 n\theta}{n^4(\cos\theta - \cos\theta_k)^2} \sum_{r=1}^{n-1} r \sin r\theta_k \cos r\theta = S_1(\theta) + S_2(\theta).$$

We can write

$$S_2(\theta) = \frac{2(-1)^{k+1}\sin^2\theta_k \cos^3 n\theta}{n^4(\cos\theta - \cos\theta_k)^2} \cdot \frac{d}{d\theta} \sum_{r=1}^{n-1} [\cos r(\theta - \theta_k) - \cos r(\theta + \theta_k)] =$$

$$= \frac{(-1)^{k+1}\sin^2\theta_k \cos^3 n\theta}{n^4(\cos\theta - \cos\theta_k)^2} \cdot \frac{d}{d\theta} \left[\frac{\sin\left(n - \frac{1}{2}\right)(\theta - \theta_k)}{\sin\frac{1}{2}(\theta - \theta_k)} - \frac{\sin\left(n - \frac{1}{2}\right)(\theta + \theta_k)}{\sin\frac{1}{2}(\theta + \theta_k)} \right] =$$

$$= \frac{2\sin^2\theta_k \cdot \cos^3 n\theta}{n^4(\cos\theta - \cos\theta_k)^2} \cdot \frac{d}{d\theta} \left[\frac{\cos n\theta \sin\theta}{\cos\theta - \cos\theta_k} - \sin n\theta \right],$$

[1] For simplicity, we write x_k for x_{kn} and θ_k for θ_{kn}

after putting $\cos n\theta_k = 0$ and $\sin n\theta_k = (-1)^{k+1}$. Hence,

$$S_2'(\theta) = \frac{2\sin^2\theta_k \cos^3 n\theta}{n^4(\cos\theta - \cos\theta_k)^2} \cdot \frac{d^2}{d\theta^2}\left[\frac{\cos n\theta \sin\theta}{\cos\theta - \cos\theta_k} - \sin n\theta\right] + \tag{4.5}$$

$$+ \frac{d}{d\theta}\left[\frac{2\sin^2\theta_k \cos^3 n\theta}{n^4(\cos\theta - \cos\theta_k)^2}\right] \cdot \frac{d}{d\theta}\left[\frac{\cos n\theta \sin\theta}{\cos\theta - \cos\theta_k} - \sin n\theta\right] = S_2'^{(1)}(\theta) + S_2'^{(2)}(\theta).$$

By actual differentiation and using the inequalities

$$|\sin n\theta| \leqq n|\sin\theta|, \quad |\cos\theta| \leqq 1, \quad |\cos n\theta| \leqq 1,$$

we get

$$|S_2'^{(1)}(\theta)| = \left| \frac{2\sin^2\theta_k \cos^3 n\theta}{n^4(\cos\theta - \cos\theta_k)^2}\left[\frac{2\cos n\theta \sin^3\theta}{(\cos\theta - \cos\theta_k)^3} + \right.\right.$$

$$+ \frac{3\cos n\theta \cos\theta \sin\theta - 2n \sin n\theta \sin^2\theta}{(\cos\theta - \cos\theta_k)^2} - \frac{(n^2-1)\cos n\theta \sin\theta - 2n \sin n\theta \cos\theta}{(\cos\theta - \cos\theta_k)} +$$

$$\left.\left. + n^2 \sin n\theta\right]\right| \leqq \sin^2\theta_k \sin\theta \left[\frac{4\cos^4 n\theta \sin^2\theta}{n^4|\cos\theta - \cos\theta_k|^5} + \frac{6\cos^4 n\theta|}{n^4(\cos\theta - \cos\theta_k)^4} + \right.$$

$$\left. + \frac{4|\cos^3 n\theta|\sin\theta}{n^3(\cos\theta - \cos\theta_k)^4} + \frac{6|\cos^3 n\theta|}{n^2|\cos\theta - \cos\theta_k|^3} + \frac{2|\cos^3 n\theta \sin n\theta}{n^2(\cos\theta - \cos\theta_n)^2 \sin\theta|}\right].$$

Therefore,

$$\frac{|\cos\theta - \cos\theta_k||S_2'^{(1)}(\theta)|}{\sin\theta} \leqq \frac{4\cos^4 n\theta \sin^2\theta_k \sin^2\theta}{n^4(\cos\theta - \cos\theta_k)^4} + \frac{6\cos^4 n\theta \sin^2\theta_k}{n^4|\cos\theta - \cos\theta_k|^3} +$$

$$+ \frac{4|\cos^3 n\theta| \sin^2\theta_k \sin\theta}{n^3|\cos\theta - \cos\theta_k|^3} + \frac{6\cos^2 n\theta \sin^2\theta_k}{n^2(\cos\theta - \cos\theta_k)^2} + \frac{2|\cos^3 n\theta \sin n\theta|}{n^2(\cos\theta - \cos\theta_k)^2 \sin\theta}.$$

Since,

$$\sin\theta_k \leqq \sin\theta_k + \sin\theta \leqq 2\sin\frac{\theta+\theta_k}{2}, \quad \sin\theta \leqq \sin\theta + \sin\theta_k \leqq 2\sin\frac{\theta+\theta_k}{2},$$

$$|\cos n\theta| = |\cos n\theta - \cos n\theta_k| \leqq 2\left|\sin n\frac{\theta_k - \theta}{2}\right| \quad \text{and} \quad \frac{1}{\sin\dfrac{\theta+\theta_k}{2}} \leqq \frac{1}{\sin\dfrac{|\theta-\theta_k|}{2}},$$

we get,

$$(4.6)\quad \frac{|\cos\theta-\cos\theta_k||S_2'^{(1)}(\theta)|}{\sin\theta}\leqq\frac{64\sin^4\frac{\theta-\theta_k}{2}}{n^4\sin^4\frac{1}{2}(\theta-\theta_k)}+\frac{48\sin^4\frac{1}{2}n(\theta-\theta_k)}{n^4\sin\frac{1}{2}(\theta+\theta_k)\sin^3\frac{1}{2}|\theta-\theta_k|}+$$

$$+\frac{32\left|\sin^3\frac{1}{2}n(\theta-\theta_k)\right|}{n^3\sin^3\frac{1}{2}|\theta-\theta_k|}+\frac{32\sin^2\frac{1}{2}n(\theta-\theta_k)}{n^2\sin^2\frac{1}{2}(\theta-\theta_k)}+\frac{8}{n}\leqq$$

$$\leqq\frac{112\sin^4\frac{1}{2}n(\theta-\theta_k)}{n^4\sin^4\frac{1}{2}(\theta-\theta_k)}+\frac{32\left|\sin^3\frac{1}{2}n(\theta-\theta_k)\right|}{n^3\sin^3\frac{1}{2}|\theta-\theta_k|}+\frac{32\sin^2\frac{1}{2}n(\theta-\theta_k)}{n^2\sin^2\frac{1}{2}(\theta-\theta_k)}+\frac{8}{n}.$$

Similarly, we have

$$(4.7)\quad \frac{|\cos\theta-\cos\theta_k||S_2'^{(2)}(\theta)|}{\sin\theta}=\left|\sin^2\theta_k\left\{\frac{4\cos^4 n\theta\sin^2\theta}{n^4(\cos\theta-\cos\theta_k)^4}+\frac{4\cos^4 n\theta\cos\theta}{n^4(\cos\theta-\cos\theta_k)^3}-\right.\right.$$

$$-\frac{10\cos^3 n\theta\sin n\theta\sin\theta}{n^3(\cos\theta-\cos\theta_k)^3}-\frac{4\cos^4 n\theta}{n^3(\cos\theta-\cos\theta_k)^2}-\frac{6\cos^3 n\theta\cos\theta\sin n\theta}{n^3(\cos\theta-\cos\theta_k)^2\sin\theta}+$$

$$\left.\left.+\frac{6\cos^2 n\theta\sin^2 n\theta}{n^2(\cos\theta-\cos\theta_k)^2}\right\}+\frac{6\cos^3 n\theta\sin n\theta(\cos\theta+\cos\theta_k)}{n^2\sin\theta}+\frac{6\sin^2\theta_k\cos^3 n\theta\sin n\theta\sin\theta}{n^2(\cos\theta-\cos\theta_k)}\right|\leqq$$

$$\leqq\frac{68\sin^4\frac{1}{2}n(\theta-\theta_k)}{n^4\sin^4\frac{1}{2}(\theta-\theta_k)}+\frac{80\left|\sin^3\frac{1}{2}n(\theta-\theta_k)\right|}{n^3\sin^3\frac{1}{2}|(\theta-\theta_k)|}+$$

$$+\frac{52\sin^2\frac{1}{2}n(\theta-\theta_k)}{n^2\sin^2\frac{1}{2}(\theta-\theta_k)}+\frac{12\left|\sin\frac{1}{2}n(\theta-\theta_k)\right|}{n\sin\frac{1}{2}|(\theta-\theta_k)|}+\frac{12}{n}.$$

Combining (4.6) and (4.7), we have

$$(4.8)\quad \frac{|\cos\theta-\cos\theta_k||S_2'(\theta)|}{\sin\theta}\leqq\frac{180\sin^4\frac{1}{2}n(\theta-\theta_k)}{n^4\sin^4\frac{1}{2}(\theta-\theta_k)}+\frac{112\left|\sin^3\frac{1}{2}n(\theta-\theta_k)\right|}{n^3\sin\frac{1}{2}|\theta-\theta_k|}+$$

$$+\frac{84\sin^2\frac{1}{2}n(\theta-\theta_k)}{n^2\sin^2\frac{1}{2}(\theta-\theta_k)}+\frac{12\sin\left|\frac{1}{2}n(\theta-\theta_k)\right|}{n^2\sin\frac{1}{2}|\theta-\theta_k|}+\frac{20}{n}.$$

We shall now estimate the quantity

$$\frac{|\cos\theta-\cos\theta_k|\,|S_1'(\theta)|}{\sin\theta}.$$

We have

$$S_1'(\theta)=\frac{\sin^2\theta_k}{n^4}\left[\frac{4(1-\cos\theta\cos\theta_k)\cos^4 n\theta\sin\theta}{(\cos\theta-\cos\theta_k)^5}-\right.$$

$$\left.-\frac{4n\cos^3 n\theta\sin n\theta(1-\cos\theta\cos\theta_k)-\sin\theta\cos\theta_k\cos^4 n\theta}{(\cos\theta-\cos\theta_k)^4}\right].$$

Hence

$$\text{(4.9)}\qquad \frac{|\cos\theta-\cos\theta_k|\,|S_1'(\theta)|}{\sin\theta}=\left|\frac{\sin^2\theta_k}{n^4}\left[\frac{4\sin^2\theta\cos^4 n\theta}{(\cos\theta-\cos\theta_k)^4}+\frac{4\cos\theta\sin\theta\cos^4 n\theta}{(\cos\theta-\cos\theta_k)^3}-\right.\right.$$

$$\left.\left.-\frac{4n\cos^3 n\theta\sin n\theta\sin\theta}{(\cos\theta-\cos\theta_k)^3}-\frac{4n\cos^3 n\theta\sin n\theta\cos\theta}{\sin\theta(\cos\theta-\cos\theta_k)^2}+\frac{\cos\theta_k\cos^4 n\theta}{(\cos\theta-\cos\theta_k)^3}\right]\right|\leqq$$

$$\leqq\frac{64\sin^4\frac{1}{2}n(\theta-\theta_k)}{n^4\sin^4\frac{1}{2}(\theta-\theta_k)}+\frac{64\left|\sin^3\frac{1}{2}n(\theta-\theta_k)\right|}{n^3\sin^3\frac{1}{2}|\theta-\theta_k|}+\frac{16\sin^2\frac{1}{2}n(\theta-\theta_k)}{n^2\sin^2\frac{1}{2}(\theta-\theta_k)}.$$

From (4.8) and (4.9), we have on account of (4.4)

$$\text{(4.10)}\qquad |x-x_k|\,|\Phi_k'(x)|\leqq\frac{244\sin^4\frac{1}{2}n(\theta-\theta_k)}{n^4\sin^4\frac{1}{2}(\theta-\theta_k)}+\frac{176\left|\sin^3\frac{1}{2}n(\theta-\theta_k)\right|}{n^3\sin^3\frac{1}{2}|\theta-\theta_k|}+$$

$$+\frac{100\sin^2\frac{1}{2}n(\theta-\theta_k)}{n^2\sin^2\frac{1}{2}(\theta-\theta_k)}+\frac{12\left|\sin\frac{1}{2}n(\theta-\theta_k)\right|}{n^2\sin^2\frac{1}{2}|\theta-\theta_k|}+\frac{20}{n}.$$

To prove (4.3), let

$$\frac{j-1}{n}\Pi\leqq\theta\leqq\frac{j}{n}\Pi,$$

then from (4.10), we have on account of

$$\frac{1}{\sin\frac{1}{2}|\theta-\theta_k|}\leqq\frac{n}{2i-1},\quad k=j\pm i,$$

$$|x-x_k|\,|\varphi_k'(x)|\leqq\begin{cases}552 & k=j,\\ \dfrac{244}{(2i-1)^4}+\dfrac{176}{(2i-1)^3}+\dfrac{100}{(2i+1)^2}+\dfrac{32}{n} & \begin{cases}j<k=j+i\leqq n\\ i\leqq k=j-i<j.\end{cases}\end{cases}$$

Consequently,

$$\sum_{k=1}^{n}|x-x_k||\varphi_k'(x)| \leqq 552+\sum_{k\neq j}\leqq 536+\sum_{i=1}^{n}\left[\frac{244}{(2i-1)^4}+\frac{176}{(2i-1)^3}+\frac{100}{(2i-1)^2}+\frac{32}{n}\right]<$$

$$< 584+\sum_{i=1}^{\infty}\left[\frac{244}{(2i-1)^4}+\frac{176}{(2i-1)^3}+\frac{100}{(2i-1)^2}\right] < 1584,$$

and this completes the proof of (4.3).

LEMMA 2. *For* $-1\leqq x\leqq 1$,

$$|J_n^*(f,1)-J_n^*(f,x)| \leqq 4026\omega(1-x).$$

PROOF. From (1.8), we have

$$J_n^*(f,1)-J_n^*(f,x)=f(1)+\sum_{k=1}^{n}[f(x_{kn})-f(1)]\varphi_{kn}(1)-$$

$$-\frac{1+x}{2}f(1)-\frac{1-x}{2}f(-1)-\sum_{k=1}^{n}[f(x_{kn})-f(1)]\varphi_{kn}(x)- \tag{4.11}$$

$$-\sum_{k=1}^{n}\left[\frac{1-x}{2}f(1)-\frac{1-x}{2}f(-1)\right]\varphi_{kn}(x)=$$

$$=\frac{1-x}{2}[f(1)-f(-1)]-\frac{1-x}{2}[f(1)-f(-1)]\sum_{k=1}^{n}\varphi_{kn}(x)+$$

$$+\sum_{k=1}^{n}[f(x_{kn})-f(1)][\varphi_{kn}(1)-\varphi_{kn}(x)]=$$

$$=\sum_{k=1}^{n}[f(x_{kn})-f(1)][\varphi_{kn}(1)-\varphi_{kn}(x)]+\frac{1-x}{2}[f(1)-f(-1)]\left[1-\sum_{k=1}^{n}\varphi_{kn}(x)\right].$$

Hence, on account of (3.1)—(3.5) and lemma 1.

$$|J_n^*(f,1)-J_n^*(f,x)| \leqq \frac{1-x}{2}\omega(2)\left|1-\sum_{k=1}^{n}\varphi_{kn}(x)\right|+ \tag{4.12}$$

$$+\sum_{k=1}^{n}\omega(1-x_{kn})|\varphi_{kn}(x)-\varphi_{kn}(1)| \leqq$$

$$\leqq 6\omega(1-x)+\sum_{x_k\leqq x}\omega(1-x_{kn})|\varphi_{kn}(x)-\varphi_{kn}(1)|+\sum_{x_k\leqq x}\omega(1-x_{kn})|\varphi_{kn}(x)-\varphi_{kn}(1)| \leqq$$

$$\leqq 6\omega(1-x)+284\omega(1-x)+568\omega(1-x)+$$

$$+2\omega(1-x)\sum_{x_k\leqq x}(x-x_{kn})\frac{|\varphi_{kn}(x)-\varphi_{kn}(1)|}{1-x}\leqq$$

$$\leqq 858\omega(1-x)+2\omega(1-x)\sum_{x_k\leqq x}(x-x_{kn})|\varphi_{kn}'(\xi)|\, x<\xi<1 \leqq$$

$$\leqq 858\omega(1-x)+2\omega(1-x)\sum_{k=1}^{n}|\xi-_{kn}||\varphi_{kn}'(\xi)| \leqq 4026\omega(1-x).$$

Similarly we have

LEMMA 3. *For* $-1 \leqq x \leqq 1$

$$|J_n^*(f,-1)-J_n^*(f,x)| \leqq 4026\omega(1+x). \tag{4.13}$$

5. Proof of the theorem

First, let $\sqrt{1-x^2} \geqq \frac{1}{n}$. By virtue of the definition of g_n and the inequality (1.9) we have

$$|f(x)-g_n(f,x)| = \left| r_n(f,x) - \left\{ \frac{1+x}{2} r_n(f,1) + \frac{1-x}{2} r_n(f,-1) \right\} \right| \leqq \tag{5.1}$$

$$\leqq c\left[\omega\left(f; \frac{\sqrt{1-x^2}}{4n}\right) + \omega\left(f; \frac{1}{16n^2}\right)\right] + 2c\omega\left(f; \frac{1}{16n^2}\right) \leqq 5c\omega\left(f; \frac{\sqrt{1-x^2}}{n}\right).$$

Next, let $\sqrt{1-x^2} < \frac{1}{n}$. In this case, on account of (1.8) and (2.1)—(2.3),

$$f(x)-g_n(f,x) = [f(x)-J_n^*(f,x)] - \frac{1+x}{2}[f(1)-J_n^*(f,1)] -$$

$$-\frac{1-x}{2}[f(-1)-J_n^*(f,-1)] =$$

$$= \frac{1+x}{2}[f(x)-f(1)] + \frac{1-x}{2}[f(x)-f(-1)] +$$

$$+\frac{1+x}{2}[J_n^*(f,1)-J_n^*(f,x)] + \frac{1-x}{2}[J_n^*(f,-1)-J_n^*(f,x)].$$

Thus, oving to lemma 2 and lemma 3, we have

$$|f(x)-g_n(f,x)| \leqq \frac{1+x}{2}\omega(1-x) + \frac{1-x}{2}\omega(1+x) +$$

$$+4026\frac{1+x}{2}\omega(1-x) + 4026\frac{1-x}{2}\omega(1+x) \leqq$$

$$\leqq 2014[(1+x)\omega(1-x)+(1-x)\omega(1+x)].$$

If $x<0$, then owing to (3.1)—(3.3),

$$(1+x)\omega(1-x)+(1-x)\omega(1+x) \leqq (1-x^2)\omega(2)+2\omega(1-x^2) \leqq \tag{5.2}$$

$$\leqq 4\omega(1-x^2)+2\omega(1-x^2) = 6\omega(1-x^2).$$

If $x > 0$, then

$$(1+x)\omega(1-x)+(1-x)\omega(1+x) \leqq 2\omega(1-x^2)+(1-x^2)\omega(2) \leqq 6\omega(1-x^2).$$

Hence, in every case

$$|f(x)-g_n(f,x)| \leqq 12\,084\omega(1-x^2) < 12\,084\omega\left(\frac{\sqrt{1-x^2}}{n}\right). \tag{5.3}$$

(5.1) and (5.3) complete the proof of the theore.

REFERENCES

[1] FREUD, G.: Egy Jackson-féle interpolációs eljárásról, *Mat. Lapok*, **4** 1964, 330—336.
[2] FREUD, G. and VÉRTESI, P.: A new proof of A. F. Timan's approximation theoram, *Studia Sci. Math. Hungar.* **2** (1967), 403—414.
[3] GOPENGAUZ, I. E.: On the theorem of A. F. Timan on approximation of continuous functions on a finite segment, *Math. Zametski*, **1** No. 2, 163—172, 1967 (Russian).
[4] SAXENA, R. B.: On a polynomial of interpolation, *Studia Sci. Math. Hungar.* **2** (1967), 167—183.
[5] SAXENA, R. B.: A new proof of S. A. Telyakovskii's theorem on the approximation of continuous functions by algebraic polynomials, *Studia Sci. Math. Hungar.* *7* (1972), 3—9.
[6] TIMAN, A. F.: Strengthening of Jackson's theorem on the best approximation of continuous functions by polynomials on the finite segment of real axis. *Dokl. Akad. Nauk. SSSR* **78** (1951), 17—20, (Russian).
[7] TELYAKOVSKII, S. A.: Two theorems on the approximation of functions by algebraic polynomials, *Math. Sbornik*, **70**, No. 2, 252—265, 1966 (Russian).
[8] VÉRTESI, P.: Jackson tételének bizonyítása interpolációs úton. *Mat. Lapok*, **18**, (1966), 83—92.

Lucknow University Lucknow, India.

(Received February 7, 1972; revised February 12, 1973)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 447—456.

ON THE CLASS OF MATRICES WHICH ARE CONVERGENT FOR A GIVEN MATRIX

by

V. N. JOSHI

1. Introduction

It is the purpose of this paper to study certain properties and a characterization of a class of matrices which are convergent for a given matrix. Such matrices play an important role in connection with the most general iterative schemes which one sets up to solve a system of linear equations with a possibly singular coefficient matrix. In such a situation the problem of the convergence of the iterations requires special considerations. In case of a nonsingular coefficient matrix, the convergence of iterations is equivalent to the convergence of the so-called iteration matrix associated with the scheme. The fact that this equivalence is no longer valid for a singular coefficient matrix was pointed out by H. B. Keller [7]. He determined a 'condition for convergence of an iteration matrix' which replaced that of the matrix being 'convergent'. In the case of rectangular matrices, this condition has been further modified by the author [5]. These considerations have motivated the present investigation.

2. Preliminaries and notation

All matrices considered here have complex entries unless stated otherwise. We denote by C_m the m-dimensional vector space. For any matrix V, we denote respectively by $n(V)$ and $\bar{n}(V)$ the null space of V and any complement of the null space of V. Thus

$$C_m = n(V) \oplus \bar{n}(V),$$

where $\oplus$ denotes the direct sum. The phrase 'V is on C_m' will be used to signify that V operates on the points of C_m. If C^1 is the entire complex plane and Z is the set of complex numbers $Z=\{e^{i\theta}, \theta \neq 2n \ n=0, 1, \ldots\}$, then the deleted set C^1-Z will be denoted by C_0^1. If $\lambda(V)$ is an eigenvalue of V, then we will use $\hat{\lambda}(V)$ to denote $\lambda(V)$, if and only if, it belongs to C_0^1. We will denote the 'deleted spectral radius' of V, viz. $\hat{\varrho}(V)$, as follows:

$$\hat{\varrho}(V) = \max_i |\hat{\lambda}_i(V)|.$$

It is well known that, for any matrix V to be convergent, a necessary and sufficient condition is that $\max_i |\lambda_i(V)|$ be less than unity. If $n(V)=n(V^*)$, where V^* denotes the hermitian conjugate of V, then we will call V an EP matrix [6], and if in addition, V has rank r, then we will call V an EP_r matrix. The Moore—Penrose generalised inverse V^+ of V, is the unique matrix satisfying (i) $VV^+V=V$ (ii) $V^+VV^+=V^+$ (iii) $(VV^+)^*=VV^+$ and (iv) $(V^+V)^*=V^+V$ [8].

3. The class of semiconvergent matrices

Let A and B be $m\times m$ matrices on C_m. A is said to be convergent for B, if

$$\lim_{k\to\infty} A^k = P_{n(B)}, \tag{3.1}$$

where $P_{n(B)}$ is a projection, not necessarily orthogonal, on to $n(B)$ [7].

We define the class of semiconvergent matrices to be those matrices A such that there exist a matrix B for which it is convergent.

For a given matrix B, we will denote the class of semiconvergent matrices A by $C(B)$, and the notation $A\in C(B)$ will be used to signify that a matrix A belongs to $C(B)$.

The first lemma, which we state and prove for the sake of completeness, is due to H. B. KELLER [7].

LEMMA (3.1). *If A is an identity on $n(B)$ and is convergent on $\bar{n}(B)$, then $A\in C(B)$.*

PROOF. Any vector $u\in C_m$ can be expressed uniquely as

$$u = v+w, \tag{3.2}$$

where $v\in n(B)$ and $w\in\bar{n}(B)$. Now, in view of the fact that A is an identity on $n(B)$ we get

$$Au = v+Aw;$$

which implies that, for any non negative integer k,

$$A^k u = v+A^k w.$$

Proceeding to the limit as $k\to\infty$, and using the hypothesis, we have

$$\lim_{k\to\infty} A^k u = v = P_{n(B)}u. \tag{3.3}$$

Since u is an arbitrary element of C_m, the lemma follows from (3.3) and (3.1).

LEMMA (3.2). *If $A\in C(B)$, then $\hat{\varrho}(A)\leqq 1$.*

PROOF. If u_λ is an eigenvector of A corresponding to an eigenvalue $\lambda=\lambda(A)$ of A, then for any integer $k\geqq 0$,

$$A^k u_\lambda = \lambda^k u_\lambda,$$

and in the limit as k tends to infinity,

$$\lim_{k\to\infty} A^k u_\lambda = \lim_{k\to\infty} \lambda^k u_\lambda. \tag{3.4}$$

If possible, let λ be a complex eigenvalue with $|\lambda|=1$. We can write $\lambda=e^{i\theta}$. It can be easily verified that the limit in (3.4) exists for $\lambda=e^{i\theta}$, if and only if, $\theta=2n\pi$, $n=0, 1, \ldots$. Hence $\lambda\in C_0^1$, and we rewrite (3.4) in the notation of Section 2, as follows

$$\lim_{k\to\infty} A^k u_{\hat{\lambda}} = \lim_{k\to\infty} \hat{\lambda}^k u_{\hat{\lambda}}. \tag{3.5}$$

In view of the fact that $A \in C(B)$, the limit in (3.5) exists finitely for all $u_{\hat{\lambda}}$ and the corresponding $\hat{\lambda}$. This implies that $|\hat{\lambda}| \not> 1$. In other words $|\hat{\lambda}| \leqq 1$. Since $\hat{\varrho}(A) = \max_i |\hat{\lambda}_i(A)|$, we get $\hat{\varrho}(A) \leqq 1$. This completes the proof of the lemma.

If $A \in C(B)$, and B has nullity p, then A has an eigenvalue $\lambda = 1$ of multiplicity p. This fact is brought out in the following lemma.

LEMMA (3.3). *If $A \in C(B)$, then $n(B)$ is identical with the invariant subspace, S, of A spanned by the eigenvectors of A with eigenvalue $\lambda = 1$, and A is an identity on $n(B)$.*

PROOF. By Lemma (3.2), we have $\hat{\varrho}(A) \leqq 1$. Let v be an eigenvector of A corresponding to a unit eigenvalue. If v is of grade 2 [4], then there exists an eigenvector $w \neq 0$, such that

$$(A - I)v = w \quad \text{and} \quad (A - I)w = 0.$$

In view of this it follows that for any integer $k \geqq 1$

$$A^k v = v + kw.$$

Since $w \neq 0$, we conclude from this that $\lim_{k \to \infty} A^k v$ does not exist. This contradiction to the hypothesis implies that all the eigenvectors of A corresponding to a unit eigenvalue are simple. Further, if p is the multiplicity of the unit eigenvalue, then there are p linearly independent eigenvectors corresponding to it. We denote these eigenvectors by $v_1, v_2, \ldots, v_p$ so that $Av_i = v_i$ $(i = 1, 2, \ldots, p)$. If S is the subspace of C_m which is spanned by $v_1, v_2, \ldots, v_p$ then it follows that S has dimension equal to p.

Any vector $\eta \in S$ can be expressed uniquely as

$$\eta = \sum_{i=1}^{p} \alpha_i v_i,$$

where α_i are scalars and it follows that $A\eta = \eta$; since η is arbitrary vector in S, A is an identity on S.

Let ξ be a vector in S, hence it is in C_m. Now ξ can be expressed uniquely as

$$\xi = \mu + \beta, \quad \mu \in n(B) \quad \text{and} \quad \beta \in \bar{n}(B).$$

In view of the fact that A is an identity on S, we have

$$\lim_{k \to \infty} A^k \xi = \xi = \mu + \beta.$$

On the other hand, by hypothesis

$$\lim_{k \to \infty} A^k \xi = \mu.$$

From this it follows that $\beta = 0$ or $\xi = \mu$. Since $B\xi = B\mu = 0$, we conclude that

$$S \subseteqq n(B). \tag{3.6}$$

If $z \in n(B)$ so that $P_{n(B)}z=z$, then using hypothesis it follows that

$$\lim_{k\to\infty} A^k z = z. \tag{3.7}$$

In view of (3.7), we get

$$Az=z. \tag{3.8}$$

If this implication were not true, then $\hat{\varrho}(A)\neq 1$ and hence $\hat{\varrho}(A)<1$. It follows from well known results that $(I-A)^{-1}$ exists. Also there exists a vector $\gamma\neq 0$ such that

$$Az = z+\gamma$$

and recursively, for any positive integer k,

$$A^k z = z+\sum_{i=0}^{k-1} A^i\gamma .$$

It follows that as $k\to\infty$

$$\lim_{k\to\infty} A^k z = z+(I-A)^{-1}\gamma.$$

Comparing this with (3.7), we have

$$(I-A)^{-1}\gamma = 0,$$

which implies that $\gamma=0$. This contradiction shows that (3.8) is true; hence

$$n(B)\subseteq S. \tag{3.9}$$

Combining this with (3.6), we get

$$n(B)=S.$$

Since A is an identity on S, it is also an identity on $n(B)$. This completes the proof of the lemma.

Now we are in a position to put down the first necessary and sufficient condition for a matrix A to be convergent for a given matrix B.

THEOREM (3.1). *A set of necessary and sufficient conditions for $A\in C(B)$, is that both $n(B)$ and $\bar{n}(B)$ are invariant under A; A is an identity on $n(B)$ and is convergent on $\bar{n}(B)$.*

PROOF. If $A\in C(B)$, then from Lemma (3.3) it follows that $n(B)$ is identical with S and A is an identity on $n(B)$; since S is invariant under A, $n(B)$ is invariant under A. Any complement $\bar{n}(B)$ of B is contained in the union of the invariant subspaces of A corresponding to eigenvalues with absolute value less than unity. We conclude the necessary part of the proof by noting that such a union is invariant under A.

The proof of the sufficiency follows directly from the Lemma (3.1).

COROLLARY (3.1). *If $A\in C(B)$, then a necessary and suffcient condition for A to be convergent is that B be nonsingular.*

PROOF. If B is nonsingular, its null space is the null vector, and the sufficiency of the proof follows on using the above theorem.

If u is any vector in C_m, then we can write $u = v+w$, where $v \in n(B)$ and $w \in \bar{n}(B)$. By the definition of A convergent for B, we have

$$\lim_{k\to\infty} A^k u = v.$$

On the other hand, A is convergent implies that

$$\lim_{k\to\infty} A^k u = 0.$$

It follows that v is the null vector and hence B is nonsingular.

From now we shall assume that A leaves invariant the null space of B and some one of its complements. The following theorem reveals a structure of a matrix A which is convergent for B. This is the first characterization of the class of semi-convergent matrices which is useful in the discussion of the convergence of iterative schemes.

THEOREM (3.2). *A set of necessary and sufficient conditions for $A \in C(B)$ is that for some nonsingular matrix C of the appropriate size, $A = I - C^{-1}B$ and that A be convergent on $\bar{n}(B)$.*

PROOF. Let $v \in n(B)$. This implies that $Bv = 0$. Since $A = I - C^{-1}B$, we have $Av = v$; and v being arbitrary, A is an identity on $n(B)$. By hypothesis, A is convergent on $\bar{n}(B)$, therefore, in view of the Theorem (3.1), the sufficiency follows.

From Lemma (3.3) we have $S = n(B)$. In order words $n(B)$ is identical with $n(A-I)$. Let $v_1, v_2, \ldots, v_p$ be the basis of the $n(B)$. It follows that each v_i, $i = 1, 2, \ldots, p$, is orthogonal to the rows of B and also to the rows of $A-I$. Therefore, the rows of $(A-I)$ are linear combinations of the rows of B. Equivalently, there exists an $m \times m$ nonsingular matrix D, such that

$$D(A-I) = B.$$

Taking $D = -C$, we get

$$A = I - C^{-1}B,$$

and it follows from Theorem (3.1) that A is convergent on $\bar{n}(B)$.

4. Further properties of semiconvergent matrices

THEOREM (4.1). *If $A \in C(B)$ and if A' and B' are similar to A and B respectively under the same similarity transformation, then $A' \in C(B')$.*

PROOF. By hypothesis, there exists a nonsingular matrix L of appropriate size such that

$$A' = LAL^{-1} \quad \text{and} \quad B' = LBL^{-1}. \tag{4.1}$$

It follows from this that $n(B)$ is isomorphic to $n(B')$. If $C_m=n(B')\oplus\bar{n}(B')$, then $\bar{n}(B)$ is isomorphic to $\bar{n}(B')$. Let p be the dimension of $n(B)$. If $l_i=Lw_i$, $i=1, 2, \ldots, \ldots, m-p$, where $w_1,, w_2, \ldots, w_{m-p}$ are the basis vectors for $\bar{n}(B)$, then it is easy to see that the vectors $l_1, l_2, \ldots, l_{m-p}$ form a basis for $\bar{n}(B')$. Since $A\in C(B)$, from Theorem (3.2) we have

$$A = I-C^{-1}B,$$

for some nonsingular C. Now, using (4.1), we can write

$$A' = I-D^{-1}B', \tag{4.2}$$

where $D=LCL^{-1}$ and is clearly nonsingular.

If $z\in\bar{n}(B')$, we have $z=\sum_{i=1}^{m-p} d_i l_i$. Consequently, $L^{-1}z=\sum_{i=1}^{m-p} d_i w_i$ and therefore,

$$(A')^k z = L\sum_{i=1}^{m-p} d_i A^k w_i.$$

Proceeding to the limit as $k\to\infty$, we get

$$\lim_{k\to\infty}(A')^k z = 0,$$

since $w_i\in\bar{n}(B)$, $i=1, 2, \ldots, m-p$, and by hypothesis $A\in C(B)$. This implies that $A'\in C(B')$. In view of the Theorem (3.2), the proof of this theorem is complete.

We recall the definition of an *EP* matrix and of the Moore—Penrose generalised inverse of a given matrix B to state and prove

THEOREM (4.2). *If B is given $m\times m$ EP matrix, then the following statements are equivalent:*

(i) *A is convergent for B.*
(ii) *A is convergent for B^*.*
(iii) *A is convergent for B^+.*

PROOF. (i)⇒(ii): A is convergent for B implies, in view of Lemma (3.3), that A is an identity on $n(B)$. Since B is an *EP* matrix, we have, $n(B)=n(B^*)$. Therefore, A is an identity on $n(B^*)$. Now $n(B)=n(B^*)$, if and only if, $R(B)=R(B^*)$, where $R(V)$ denotes the range of an $m\times m$ matrix V. Let $\bar{n}(B)=R(B^*)$. Now, it follows that

$$\bar{n}(B)=R(B^*)=R(B)=\bar{n}(B^*).$$

By noting the fact that A is convergent on $\bar{n}(B)$, (i)⇒(ii) follows.

(ii)⇒(iii): We note that this is true even if B is not an *EP* matrix. From the definition of the generalised inverse of B, it follows that $n(B^*)=n(B^+)$. Since A is convergent for B^*, (ii)⇒(iii).

(iii)⇒(i): We have $n(B^+)=n(B^*)$ by the definition of the generalised inverse, and by hypothesis, $n(B^*)=n(B)$; whence $n(B^+)=n(B)$, completing the proof.

THEOREM (4.3). *Let $A\in C(B)$. Then a necessary and sufficient condition for $A\in C(B^*)$ is that B is an EP matrix.*

PROOF. Sufficiency of the proof follows from the Theorem (4.2).

If $A \in C(B)$ and also $A \in C(B^*)$, then in view of Lemma (3.3), we have $n(B)=S$ and $n(B^*)=S$ which implies that $n(B)=n(B^*)$; whence B is an EP matrix.

In view of the Theorem (3.2), if $A \in C(B)$, then $A = I-C^{-1}B$ for some nonsingular matrix C; therefore, B can be written as $B = C(I-A)$. We will need the following

DEFINITION (4.1). For a given matrix A, the matrices of the form $C(I-A)$, with nonsingular C, will be called complementary matrices.

It is easy to see that a matrix A is not necessarily convergent for any of its complementary matrices. However, we can construct matrices which are convergent for their complementary matrices.

THEOREM (4.4). *Let M be an $m\times n$ $(m\leqq n)$ matrix whose r rows are linearly independent, and N be $l\times n$ $(l\leqq n)$ matrix whose q rows are linearly independent. Let $A=MN^+NM^+$ and $\hat{A}=NM^+MN^+$, then A and $\hat{A}$ are convergent for their complementary matrices.*

PROOF. It can be proved that eigenvalues of A and $\hat{A}$ lie between 0 and 1 [1, 2, 9]. If B is a complementary matrix of A, then by definition

$$B = C(I-A)$$

for some nonsingular C. If $x \in n(B)$ we have, since C is nonsingular, $Ax=x$. In other words $n(B)=S$. Since $0\leqq\lambda_i(A)\leqq 1$, the proof of the theorem follows immediately. Similarly, A is convergent for its complementary matrices.

In the remainder of this section we shall give a characterization of matrices which converge for their complementary matrices.

Let A be an $m\times m$ matrix. Since A and $(I-A)$ commute, the $n(I-A)$ is invariant under A [3]. For a nonsingular matrix C of appropriate size, $n(I-A)$ is identical with the $n(C[I-A])$ and, therefore, the latter is invariant under A. If the dimension of $n(C[I-A])=p$, the restriction of A to $n(C[I-A])$ can be characterizes as follows:

THEOREM (4.5). *The $n(C[I-A])$, for any nonsingular matrix C, is an invariant subspace of C_m under A and matrix A can be represented as*

$$\begin{pmatrix} I_p & L \\ 0 & M_{m-p} \end{pmatrix},$$

where 0 is $(m-p)\times p$ zero matrix, L is $p\times(m-p)$ and M_{m-p} and I_p are square matrices of the order $m-p$ and p respectively.

PROOF. Let $\xi_1, \xi_2, \ldots, \xi_p$ be the basis for $n(C[I-A])$. (From now on we shall denote $n(C[I-A])$ by n only.) Let $C_m=n\oplus\bar{n}$. The dimension of $\bar{n}$, therefore, is $m-p$. We denote by $\beta_1, \beta_2, \ldots, \beta_{m-p}$ the basis vectors of $\bar{n}$. In terms of $\xi_1, \xi_2, \ldots, \xi_p$ and $\beta_1, \beta_2, \ldots, \beta_{m-p}$, we can uniquely represent any $x \in C_m$ as

$$x = \begin{pmatrix} \xi \\ \beta \end{pmatrix},$$

where ξ denotes the p components in n, and β denotes the $m-p$ components in $\bar{n}$. Let A be represented by the matrix

$$\begin{pmatrix} X_p & L \\ N & M_{m-p} \end{pmatrix},$$

where X_p is $p \times p$ nonsingular matrix. The vector Ax can be written as

$$\begin{pmatrix} X_p\xi + L\beta \\ N\xi + M_{m-p}\beta \end{pmatrix}.$$

Suppose $x \in n$; i.e. $\beta=0$. Since A is an identity on n, $Ax=x$, and therefore, we get

$$X_p\xi = \xi \quad \text{and} \quad N\xi = 0.$$

From the arbitrariness of ξ, the result immediately follows.

If A is convergent for its complementary matrix, $\bar{n}$ is also invariant under A. Proceeding on similar lines, we see that A can be represented as,

$$\begin{pmatrix} I_p & 0_{p\times(m-p)} \\ 0_{(m-p)\times p} & M_{m-p} \end{pmatrix}, \tag{4.3}$$

which can be rewritten as

$$\begin{pmatrix} I_p & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & M_{m-p} \end{pmatrix} = A_1 + A_2, \text{ say,}$$

where 0 denotes zero matrix of appropriate dimension. A_1 denotes the effect of A on n only, and A_2 that on $\bar{n}$ only. This is evident, since $A_1A_2=0=A_2A_1$.

5. Convergent splitting

We first demonstrate the rather trivial fact that $LB=B$, where L is nonsingular and B is singular matrix, need not always imply $L=I$. Consider

$$L = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

It is easy to verify that $LB=B$, but L is not an identity.

If $A \in C(B)$, then in view of the Theorem (3.2), we have

$$A = I - C^{-1}B.$$

This expression is not necessarily unique. If D is a nonsingular matrix of the appropriate size such that

$$A = I - D^{-1}B,$$

then we have $(DC^{-1})B=B$, where $L=DC^{-1}$ is nonsingular. Using above remark we conclude that D is not necessarily the same as C. This noninqueness of the expression can be effectively used to construct a convergent splitting, which is very useful

in setting up a convergent iterative scheme. We restrict our considerations to the following problem.

Let B be an $m\times m$ positive semidefinite matrix of rank r. Construct an $m\times m$ matrix A with the following properties:

(i) $0\leqq\lambda(A)\leqq 1$.
(ii) Multiplicity of the eigenvalue $\lambda(A)=1$ is the same as the nullity of B.
(iii) For some nonsingular matrix C, $B=C(I-A)$.

In what follows we give a very simple construction of a matrix A possessing the above properties.

Since B is positive semidefinite, $\lambda(B)$ are all real and nonnegative. Define a positive number α, viz.,

$$\alpha \geqq k = \max_{1\leqq i\leqq m} \lambda_i(B).$$

Construct a nonsingular matrix $C=(c_{ij})$, where $c_{ij}=\alpha\delta_{ij}$, $i,j=1,2,\ldots,m$. It is easy to see that C is positive definite and that $C^{-1}=(\alpha^{-1}\delta_{ij})$. Let $A = I-C^{-1}B = = \left(I-\frac{B}{\alpha}\right)$. We shall show that this A meets the requirements.

By construction, it is evident that A satisfies (iii) and B and C^{-1} commute. The matrix A is, therefore, hermitian. Consequently $\lambda(A)$ are all real. Now

$$\lambda(A) = 1-\frac{\lambda(B)}{\alpha} \geqq 1-\frac{k}{\alpha} \geqq 0$$

and

$$\lambda(A) = 1-\frac{\lambda(B)}{\alpha} \leqq 1.$$

Hence $0\leqq\lambda(A)\leqq 1$ and (i) holds. In view of the fact that A is herimitian, all its eigenvectors are simple [3]; therefore, from (iii) we get $S=n(B)$ and (ii) holds.

6. Concluding remarks

The construction of C described in Section 5 requires the knowledge of only the maximum eigenvalue of B. This can be obtained by using the power method in its simplest form [4]. If this is not desirable, then we can select $\alpha\geqq\|B\|_1$, or $\alpha\geqq\|B\|_\infty$ where $\|\cdot\|_1$ and $\|\cdot\|_\infty$ denote the one and the maximum matrix norms [4].

Suppose that we want to solve the following system of linear equations, by iteration.

$$Bx=b, \tag{6.1}$$

where B is positive semidefinite $m\times m$ matrix, x and b are $m\times 1$ vectors. We select C as shown above, then split B as,

$$B = C-P.$$

Equation (6.1) can be written as

$$Cx = Px+b.$$

Let $x^{(0)}$ be arbitrary and we set up an iterative scheme

$$Cx^{(k+1)} = Px^{(k)} + b, \qquad k = 0, 1, \ldots$$

or

$$x^{(k+1)} = Ax^{(k)} + C^{-1}b, \quad \text{where} \quad A = I - C^{-1}B. \tag{6.2}$$

Since A is convergent for B (see Section 5), the sequence of iterates $\{x^{(k)}\}$, given by (6.2), converges to a solution of (6.1) [5, 7].

We conclude this section with a final remark that if B were not positive semidefinite to start with, or if the vector b does not lie in the column space of B, then we can multiply by B^* on the both sides of the equation (6.1) to get

$$B^*Bx = B^*b = d, \text{ say.}$$

Now B^*B is positive semidefinite and the residual $b - Bx$ is in the orthogonal complement of the $R(B)$ giving the least squares solution of (6.1).

I would like to thank DR. A. S. HOUSEHOLDER, and DR. R. P. TEWARSON for their most valuable comments and suggestions in preparing this paper.

REFERENCES

[1] BEN-ISRAEL, A.: On the geometry of subspaces in Euclidean n-spaces, *SIAM J. App. Math.* **15** No. 5. (1967) 1184—1198.

[2] DAHLQUIST, G., SJÖBERG, B. and SVENSSON, P.: Comparison of the method of averages with the method of least squares, *Math. of Comp.*, **22,** No. 104, (1968) 833—845.

[3] FRIEDMAN, B.: *Principles and Techniques of Applied Mathematics,* John Wiley and Sons, (1966).

[4] HOUSEHOLDER, A. S.: *The Theory of Matrices in Numerical Analysis,* Blaisdell, New York, (1964).

[5] JOSHI, V. N.: A note on the solution of rectangular linear systems by iteration, *SIAM Rev.,* **12** No. 3, (1970), 463—466.

[6] KATZ, I. J. and PEARL, M. H.: On EP_r and normal EP_r matrices, *J. Res. NBS 70B (Math. and Math. Phys.),* **2** (1966) 47—77.

[7] KELLER H. B.: On the solution of singular and semidefinite linear systems by iteration, *J. SIAM Numer. Anal., Ser. B,* **2.** No. 2, (1965). 281—290.

[8] PENROSE, R.: A generalised inverse for matrices, *Proc. Cam. Phil. Soc.,* **51.** No. 3 (1955) 406—413.

[9] ZASSENHAUS, H.: Angles of inclination in correlation theory, *Amer. Math. Monthly,* **71** (1964) 218—219.

Department of Mathematics, Hunter College of CUNY
695 Park Avenue, New York, NY. 10021

(Received May 26, 1972)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 457—461.

REMARKS ON ITERATIVE METHODS FOR COMPUTING THE GENERALISED INVERSE

by

V. N. JOSHI

The genaralised inverse A^+ of an $m\times n$ matrix A is the unique solution of the following equations [5]

$$AXA = A \tag{1}$$

$$XAX = X \tag{2}$$

$$(AX)^* = AX \tag{3}$$

$$(XA)^* = XA. \tag{4}$$

If A is nonsingular, then $A^+ = A^{-1}$, otherwise A^+ still possesses properties which make it a central concept in matrix theory and in numerical analysis. Quite a few direct methods have appeared in the literature for the computation of A^+. For large sparse matrices such methods are expensive and cumbersome. However, in such cases iterative methods are preferred on high speed electronic digital computers mainly because a large number of operations are highly repetitive in character. One such iterative method is discussed in [1, 2]. This is a generalization of the Newton's method for the computation of the inverse of a nonsingular matrix.

We denote by A^* the conjugate transpose of A, $P_{R(A)}$ the orthogonal projection of E^m on $R(A)$, the range of A, and $P_{R(A^*)}$ the orthogonal projection of E^n on $R(A^*)$. We state, for the sake of completeness, the iterative method discussed in [1, 2] as the following

THEOREM. *The sequence of matrices defined by*

$$X_{n+1} = X_n(2P_{R(A)} - AX_n), \qquad (n = 0, 1, \ldots) \tag{5}$$

where X_0 is an $n\times m$ complex matrix satisfying

(6) $X_0 = A^*B_0$ *for some nonsingular* $m\times m$ *matrix* B_0,

(7) $X_0 = C_0A^*$ *for some nonsingular* $n\times n$ *matrix* C_0,

(8) $\|AX_0 - P_{R(A)}\| < 1$, ($\|\cdot\|$ *is a multiplicative matrix norm*)

$$\|X_0A - P_{R(A^*)}\| < 1, \tag{9}$$

converges to A^+.

In the proof of this theorem, it is assumed that A^+ is characterized as the unique solution of the matrix equations

$$AX = P_{R(A)} \quad \text{and} \quad XA = P_{R(A^*)}. \tag{10}$$

In fact the equations (10) do not have a unique solution and, hence, do not characterize A^+. In this connection we quote an example from [4]. Let

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix},$$

and define a matrix X as

$$X = \begin{pmatrix} \frac{1}{2}+\beta & \beta \\ \beta & \frac{1}{2}+\beta \end{pmatrix},$$

where β is any real number.

It is easy to verify that this X satisfies (10) for all the values of β, and $X=A^+$ if and only if $\beta = -\frac{1}{4}$. However, the equations (2) and (10) characterize A^+ [4]. The proof of the above theorem as given in [1] is, therefore, incomplete. In order to complete the proof one must prove that $X_nAX \to X_n$ or equivalently,

$$\operatorname*{Lt}_{n\to\infty} \|X_nAX_n - X_n\| = 0. \tag{11}$$

We shall prove the relation (11) for the iterative scheme given in (5). From (5) and (7) it follows [1] that

$$X_n = C_nA^*, \qquad (n = 0, 1, \ldots) \tag{12}$$

where

$$C_{n+1} = C_n(2P_{R(A^*)} - A^*AC_n).$$

Let $E_n = P_{R(A)} - AX_n$, $n=0, 1, \ldots$. In view of (5) this can be written as

$$AX_n = P_{R(A)} - E_n = P_{R(A)} - E_0^{2^n},$$

which implies that

$$X_nAX_n = X_nP_{R(A)} - X_nE_0^{2^n}. \tag{13}$$

Using (12) and the fact that $A^*=A^*P_{R(A)}$ we can rewrite equation (13) as

$$X_nAX_n = C_nA^*P_{R(A)} - X_nE_0^{2^n} = C_nA^* - X_nE_0^{2^n} = X_n - X_nE_0^{2^n},$$

from which it follows that

$$\|X_nAX_n - X_n\| = \|X_nE_0^{2^n}\| \leqq \|X_n\| \, \|E_0^{2^n}\| \leqq \|X_n\| \, \|E_0\|^{2^n}.$$

From this and the equation (8), the relation (11) now follows.

For the iterative scheme given in (5), it can be very easily verified that $E_n = E_{n-1}^2$, therefore, at each step an error matrix is multiplied by itself. In what follows, we will give an iterative scheme wherein an error matrix is multiplied by a constant matrix. In this case, the convergence can not be as rapid as in (5), however, the scheme to be discussed has an advantage over the other scheme in the sense that at each step just one matrix multiplication, instead of two, is required which is an important consideration.

THEOREM. *If B is an $n \times m$ matrix with its columns in the column space of A^*, then the sequence of matrices X_n defined by*

$$X_n = B + X_{n-1}(P_{R(A)} - AB), \qquad (n = 1, 2, 3, \ldots) \tag{14}$$

with $X_0 = 0$ or B, converges to A^+ provided

$$\|P_{R(A)} - AB\| < 1$$

and

$$\|P_{R(A^*)} - BA\| < 1. \tag{16}$$

PROOF. Since the columns of B are in the column space of A^*, we have

$$P_{R(A^*)} B = B. \tag{17}$$

Also from (1)—(4) it follows that

$$P_{R(A)} A = AA^+ A = A = AP_{R(A^*)}. \tag{18}$$

Now, in view of (14) and (18), we have

$$P_{R(A)} - AX_n = (P_{R(A)} - AX_{n-1})(P_{R(A)} - AB) = (P_{R(A)} - AX_{n-2})(P_{R(A)} - AB)^2 \tag{19}$$

$$= \ldots = (P_{R(A)} - AX_0)(P_{R(A)} - AB)^n = \begin{cases} (P_{R(A)} - AB)^{n+1} & \text{if} \quad X_0 = B \\ (P_{R(A)} - AB)^n & \text{if} \quad X_0 = 0. \end{cases}$$

Again using (14), (17) and (18) we get

$$P_{R(A^*)} - X_n A = P_{R(A^*)} - BA - X_{n-1} A + X_{n-1} ABA = \tag{20}$$

$$= P_{R(A^*)}^2 - P_{R(A^*)} BA - X_{n-1} AP_{R(A^*)} + X_{n-1} ABA =$$

$$= (P_{R(A^*)} - X_{n-1} A)(P_{R(A^*)} - BA) = \ldots = (P_{R(A^*)} - X_0 A)(P_{R(A^*)} - BA)^n =$$

$$= \begin{cases} (P_{R(A^*)} - BA)^{n+1} & \text{if} \quad X_0 = B \\ (P_{R(A^*)} - BA)^n & \text{if} \quad X_0 = 0. \end{cases}$$

If we write

$$C = P_{R(A)} - AB, \tag{21}$$

and

$$D = P_{R(A^*)} - BA, \tag{22}$$

then (14) can be rewritten as

$$X_n = B + X_{n-1}C, \qquad (n = 1, 2, \ldots) \tag{23}$$

and recursively,

$$X_n = B \sum_{r=0}^{s} C^r, \tag{24}$$

where $s = n$ or $n-1$ depending upon whether X_0 is B or 0. In view of (20) and (22) we get

$$X_n A = P_{R(A^*)} - D^n$$

and, therefore,

$$X_n A X_n = P_{R(A^*)} X_n - D^n X_n.$$

Using (17) and (24) the last equation becomes

$$X_n A X_n = P_{R(A^*)} B \sum_{r=0}^{s} C^r - D^n X_n = B \sum_{r=0}^{s} C^r - D^n X_n = X_n - D^n X_n,$$

whence,

$$\|X_n A X_n - X_n\| \leqq \|X_n\| \, \|D\|^n. \tag{25}$$

In view of (15) and (16) we get from (19), (20) and (25) that

$$\operatorname*{Lt}_{n \to \infty} \|P_{R(A)} - AX_n\| \to 0$$

$$\operatorname*{Lt}_{n \to \infty} \|P_{R(A^*)} - X_n A^*\| \to 0$$

and

$$\operatorname*{Lt}_{n \to \infty} \|X_n A X_n - X_n\| \to 0.$$

This completes the proof of the theorem.

COROLLARY. If $B = \alpha A^*$ and $P_{R(A)}$ is replaced by I_m, an identity matrix of order m, then the sequence

$$X_{n+1} = \alpha A^* + X_n(I_m - \alpha A A^*) = X_n + \alpha(I_m - X_n A)A^*, \qquad (n = 0, 1, \ldots)$$

with $X_0(=B)=\alpha A^*$ converges to A^+. This result is presented in [3]. The scheme (14) is, therefore, more general.

Similar to (14) we can have a sequence

$$X_n = B+(P_{R(A^*)}-BA)X_{n-1}, \qquad (n = 1, 2, \ldots) \tag{26}$$

converging to A^+ under appropriate conditions.

We conclude with a final remark that the scheme (14) or (26) can be selected according to whether AB or BA has smaller dimension and/or more sparseness.

REFERENCES

[1] BEN-ISRAEL, A.: An iterative method for computing the generalized inverse of an arbitrary matrix, *Math. of Comp.*, **19** (1965) 452—455.

[2] BEN-ISRAEL, A.: A note on an iterative method for generalized inversion of matrices, *Math. of Comp.* **20** (1966), 439—440.

[3] BEN-ISRAEL, A. and COHEN, DAN: On iterative computation of generalized inverses and associated projections, *J. SIAM. Num. Anal.*, **3** No. 3 (1966) 410—419.

[4] PETRYSHYN, W. V.: On generalized inverses and on the uniform convergence of $(I-K)^n$ with application to iterative methods. *J. Math. Anal. Appl.*, **18** (1967) 417—439.

[5] PENROSE, R.: A generalized inverse for matrices, *Proc. Cam. Phil. Soc.*, **51** (1955) 406—413.

Department of Mathematics, Hunter College of CUNY
695 Park Avenue, New York NY. 10021

(Received May 26, 1972)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 463—467.

ON THE EQUIVALENCE OF SETS DEFINABLE BY SATISFACTION AND ULTRAFILTERS

by

H. ANDRÉKA and I. NÉMETI

The preordering of first order formulae defined by the evaluations of the variables is considered. We prove that the property of being a maximal satisfiable set of formulae coincides with the property of being an ultrafilter of the preordering mentioned above. (A set of formulae is satisfiable if there is an evaluation, which satisfies all members of the set.)

All the investigations are relativised to an arbitrary class of models.

There are some well known theorems which are natural consequences of this one. This might have a significant effect on those approaches where the open formulae are just as important as the closed ones.

1. Universal algebraic notations

By an algebra we mean a fourtuple $\langle \Omega, a, A, o \rangle$, where

Ω	is a set of symbols (the set of operation symbols)
$a:\Omega\to\omega$	the arity function (where ω is the set of nonnegativ integers)
A	is a set (the carrier of the algebra)
	$\Omega_n \stackrel{d}{=} a^{-1}(n)$ (We sometimes use relations as many-valued functions, e.g. $\equiv(b) \stackrel{d}{=} \{a \mid \langle b,a\rangle \in \equiv\}$. Similarly if the inverse of a function is many-valued.)
$o:\Omega\to\bigcup_{i=0}^{\infty} {}^{(i_A)}A$	such, that $r\in\Omega_n \Rightarrow o(r):{}^nA\to A$ (where ${}^BA$ stands for the B-th Cartesian-power of A).

Algebras are identified with their carrier when there is no danger of ambiguity.

If X is a set of symbols (disjoint from Ω), then

$$W_\Omega(X) \stackrel{d}{=} \bigcap \left\{ A \mid A = X \cup \bigcup_{n=0}^{\infty} \Omega_n \times {}^nA \right\}$$

The $\langle \Omega, a, W_\Omega(X), o_W \rangle$ algebra is called the word-algebra of the set Ω over X, if

$$(\forall b\in\Omega_n,\ \alpha_i\in W_\Omega(x),\ i = 1, \ldots, n) o_w(b)(\alpha_1, \ldots, \alpha_n) \stackrel{d}{=} b\alpha_1, \ldots, \alpha_n$$

The elements of the word-algebra are sometimes cited in infix notation instead of prefix notation in which they defined above.

A transitive and reflexive relation is called a preordering. Let $\leqq$ be a preordering on the set H, and $D \subseteqq H$, now

$$\inf D \stackrel{d}{=} \{a | (\forall d \in D) a \leqq d \& (\forall d \in D) b \leqq d \Rightarrow b \leqq a)\}$$

For the shorthand of the sentence "U is an ultrafilter of the set preordered by $\leqq$" we shall write $\overline{US}\leqq, H, U$. That is,

$$\overline{US} \leqq, H, U \stackrel{d}{\Leftrightarrow} ((x \in U \& x \leqq y) \Rightarrow y \in U) \& ((x \in U \& y \in U) \Rightarrow \inf\{x, y\} \subseteqq U) \&$$

$$\& ((x \notin U \& \inf\{x, y\} = \inf H \& \sup\{x, y\} = \sup H) \Rightarrow y \in U) \& U \neq H.$$

πA stands for the power of A.

2. Logical notations

$$V \stackrel{d}{=} \{x_i | i \in \omega\}$$

$$\Lambda \stackrel{d}{=} \{\wedge, \neg, \exists_i | i \in \omega\}$$

R_n and K_n are arbitrary but fixed sets of symbols with the constraint that $\stackrel{1}{=} \in R_2$ and the elements of the set $\{V, \Lambda, K_0, K_n, R_n | n \in \omega\}$ are pairwise disjoint.

$$R \stackrel{d}{=} \bigcup_{i=1}^{\infty} R_i \quad K \stackrel{d}{=} \bigcup_{i=0}^{\infty} K_i$$

$$T \stackrel{d}{=} W_K(V) \quad P \stackrel{d}{=} \bigcup_{i=1}^{\infty} R_i \times {}^iT \quad F \stackrel{d}{=} W_\Lambda(P)$$

Now we specify how the elements of the sets just defined will be referred to:

$$\varrho_i \in R_i \quad \varkappa_i \in K_i \quad \tau \in T \quad \varphi, \psi \in F$$

We introduce symbols derived from Λ as

$$0 \stackrel{d}{=} \varphi \wedge \neg\varphi$$

(where φ is arbitrary)

$$1 \stackrel{d}{=} \neg 0$$

Interpretation:

$$I \stackrel{d}{=} \langle R, K, a, M_I, o_I \rangle$$

where

$$\langle K, a|_K, M_I, o_I|_K \rangle$$

is an algebra, and

$$o_I(\varrho_n) \stackrel{d}{\in} \pi({}^nM_I)$$

with the constraint that

$$o_I(\doteq) \overset{d}{=} \{\langle a, a\rangle | a \in M_I\}$$

$$Q \overset{d}{=} \{\langle R, K, a, M_I, o_I\rangle | M_I \text{ and } o_I \text{ are arbitrary}\}$$

In the following N stands for an arbitrary subset of Q.

Let $g(I, s): T \to M_I$ be the homomorphic extension, from the set V to the word-algebra $W_K(V)$, of the function $s: V \to M_I$. The brackets of the unary functions will be omitted, if no danger of ambiguity arises.

Let

$$k(I, s)\varrho_n\tau_1 \dots \tau_n = \uparrow \overset{d}{\leftrightarrow} \langle g(I, s)\tau_1, \dots, g(I, s)\tau_n\rangle \in o_I\varrho_n$$

$$k(I, s)\varphi \wedge \psi = \uparrow \overset{d}{\leftrightarrow} k(I, s)\varphi = \uparrow \& k(I, s) = \uparrow$$

$$k(I, s)\neg\varphi = \uparrow \overset{d}{\leftrightarrow} k(I, s)\varphi = \downarrow$$

$$k(I, s)\exists_i\varphi = \uparrow \overset{d}{\leftrightarrow} (\exists a \in M_I)k\big(I, s(x_i/a)\big)\varphi = \uparrow$$

where

$$s(x_i/a) \overset{d}{=} s \setminus \{\langle x_i, sx_i\rangle\} \cup \{\langle x_i, a\rangle\}$$

$$\varphi \leqq_N \psi \overset{d}{\leftrightarrow} (\forall I \in N, s)\big(k(I, s)\varphi = \uparrow \Rightarrow k(I, s)\psi = \uparrow\big)$$

Note, that the relation $\leqq_N$ is a preordering on F.

Let $\Sigma \subseteqq F$.

$$I \models \varphi \overset{d}{\leftrightarrow} (\forall s)k(I, s)\varphi = \uparrow$$

$$N^* \overset{d}{=} \{\varphi | (\forall I \in N) I \models \varphi\}$$

$$\Sigma^* \overset{d}{=} \{I | (\forall \varphi \in \Sigma) I \models \varphi\}$$

The meaning of N^{**} is obvious.

Let L be an index-set and $\lambda \in L$.

$$I_\lambda \overset{d}{=} \langle R, K, a, M_\lambda, o_\lambda\rangle$$

$$\langle K, a, M, b\rangle \overset{d}{=} \prod_{\lambda \in L} \langle K, a, M_\lambda, o_\lambda\rangle$$

where Π denotes direct product of algebras.

Let $\overline{US} \subseteqq, \pi L, D$ and

$$d: R \to \bigcup_{n=1}^{\infty} {}^nM \text{ such that if } \langle a_1, \dots, a_n\rangle \overset{d}{=} a \in {}^nM,$$

then

$$a \in d\varrho_n \overset{d}{\leftrightarrow} \{\lambda | \langle \varepsilon_\lambda a_1, \dots, \varepsilon_\lambda a_n\rangle \in o_\lambda \varrho_n\} \in D$$

where ε_λ is the λ-th projection function (projection to the λ-th dimension of a cartesian space)

$$J_D \stackrel{d}{=} \langle R, K, a, M, d \cup b \rangle$$

$$I_D \stackrel{d}{=} \langle R, K, a, M_D, o_D \rangle \stackrel{d}{=} J_D/d(\doteq)$$

Let $s: V \to M$

$$s_D(x_i) \stackrel{d}{=} d(\doteq)(s(x_i)) \quad s_\lambda(x_i) \stackrel{d}{=} \varepsilon_\lambda(s(x_i))$$

$$Us(\{\langle I_\lambda, s_\lambda \rangle | \lambda \in L\}, D) \stackrel{d}{=} \langle I_D, s_D \rangle$$

THEOREM. (Ultraproduct theorem (FRAYNE, MOREL & SCOTT [2], KOCHEN [3], COHN [1]))

$$k(I_D, s_D)\varphi = \uparrow \Leftrightarrow \{\lambda \in L | k(I_\lambda, s_\lambda)\varphi = \uparrow\} \in D$$

3. Equivalence of the sets definable by satisfaction and the ultrafilters

THEOREM:

$$\overline{US} \leqq_N, F, U \Leftrightarrow (\exists I \in N^{**}, s) k(I, s)^{-1}(\uparrow) = U$$

PROOF:

Step 1: $(\forall I \in N^{**}, s)\overline{US} \leqq_N, F, k(I, s)^{-1}(\uparrow)$

For: this follows immediately from the definition of function $k(I, s)$ and the preordering $\leqq_N$.

▲

Step 2: $\overline{US} \leqq_N, F, U \Rightarrow (\exists I \in N^{**}, s) k(I, s)^{-1}(\uparrow) = U$

For:

Let $\overline{US} \leqq_N, F, U$

Step 2.1: $(\forall \varphi \in U)(\exists I_\varphi \in N, s_\varphi) k(I_\varphi, s_\varphi)\varphi = \uparrow$

For:

$$\varphi \in U \Rightarrow \varphi \not\equiv_N 0 \Rightarrow \text{ it is not true, that } (\forall I \in N, s) k(I, s)\varphi = \uparrow$$

▲

For any $\varphi \in U$ we choose an $I_\varphi \in N$ and s_φ such that $k(I_\varphi, s_\varphi)\varphi = \uparrow$.

$$l_\varphi \stackrel{d}{=} \{\psi | \psi \leqq_N \varphi \& \psi \in U\}$$

$$E \stackrel{d}{=} \{l_\varphi | \varphi \in U\}$$

Step 2.2: $l_{\varphi_1}, l_{\varphi_2} \in E \Rightarrow l_{\varphi_1} \cap l_{\varphi_2} \neq \emptyset$

For:

$$l_{\varphi_1} \cap l_{\varphi_2} = \{\psi | \psi \leqq_N \varphi_1 \& \psi \in U\} \cap \{\psi | \psi \leqq_N \varphi_2 \& \psi \in U\} = \{\psi | \psi \leqq \varphi_1 \& \psi \leqq \varphi_2 \& \psi \in U\},$$

and so $\varphi_1 \wedge \varphi_2 \in l_{\varphi_1} \cap l_{\varphi_2}$.

▲

It follows from step 2.2, that $(\exists \overline{US} \subseteqq, \pi U, D)\ D \supseteqq E$. In the following D stands for such an ultrafilter.

Step 2.3: $(\forall \psi \in U)\{\varphi \in U | k(I_\varphi, s_\varphi)\psi = \uparrow\} \in D$

For:

$$\varphi \in l_\psi \Rightarrow \varphi \leqq_N \psi \& \varphi \in U \Rightarrow k(I_\varphi s_\varphi)\psi = \uparrow \& \varphi \in U$$

and so $l_\psi \subseteqq \{\varphi \in U | k(I_\varphi, s_\varphi)\psi = \uparrow\}$

▲

$$\langle I_D, s_D \rangle \stackrel{d}{=} Us(\{\langle I_\varphi, s_\varphi \rangle | \varphi \in U\}, D)$$

According to the ultraproduct theorem

$$k(I_D, s_D)\psi = \uparrow \Leftrightarrow \{\psi \in U | k(I_\varphi, s_\varphi)\psi = \uparrow\} \in D$$

and so it follows from step 2.3, that

$$(\forall \psi \in U) k(I_D, s_D)\psi = \uparrow,$$

and because U is an ultrafilter,

$$(\forall \psi \notin U) k(I_D, s_D)\psi = \downarrow.$$

▲

Corollary:

$$\overline{US} \leqq_Q, F, U \Leftrightarrow (\exists I, s)^{-1}(\uparrow) = U$$

That is, the sets of formulas definable by means of an appropriately chosen pair $\langle I, s \rangle$ are just the ultrafilters of the preordering structure $\langle \leqq_Q, F \rangle$. If we restrict ourselves to choose I from N^{**}, then what we get are exactly the ultrafilters of $\langle \leqq_N, F \rangle$.

REFERENCES

[1] COHN, P. M.: *Universal Algebra.* New York, 1965.
[2] FRAYNE, T., MOREL, A., & SCOTT, D.: Reduced direct products, *Fund. Math.* **51** (1962) 195—228.
[3] KOCHEN, SIMON B.: Ultraproducts in the theory of models, *Ann of Math.* **74** (1961) 221—261.

(Received October 17, 1972)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 469—472.

ROBBINS-MONRO PROCEDURE IN A HILBERT SPACE. II

by

P. RÉVÉSZ

Introduction

Let $Y_1, Y_2, \ldots$ be a sequence of r.v.'s taking values from a separable Hilbert space H, for which

$$\mathsf{E}(Y_1) = \mathsf{E}(Y_{k+1}|Y_k, \ldots, Y_1) = 0 \qquad (k = 1, 2, \ldots) \tag{1}$$

(with probability 1) and

$$\mathsf{P}(\|Y_k\| \leqq K) = 1 \qquad (k = 1, 2, \ldots). \tag{2}$$

Further let $M_n(x)$ be a sequence of operators (mapping H into H) for which

$$\|M_n(x)\| \leqq A\|x\| + B \qquad (A > 0;\ B > 0;\ n = 1, 2, \ldots) \tag{3}$$

and

$$(x, M_n(x)) \geqq \alpha\|x\| \quad (n \geqq n_0;\ \alpha > 0) \tag{4}$$

if $\|x\| \geqq \varepsilon/2$ where ε is a given positive number and n_0 is a positive integer (may depend on ε).

Consider the sequence $\{X_n\}$ of H valued r.v's defined by the recursion

$$X_{n+1} = X_n - \frac{M_n(X_n)}{n} - \frac{n^\beta Y_n}{n} \qquad \left(n = 1, 2, \ldots;\ 0 \leqq \beta < \frac{1}{2}\right) \tag{5}$$

where X_1 is an arbitrary element of H.

The sequence $\{X_n\}$ will be called a Robbins—Monro procedure (RMP).

Remark 1. The meaning of condition (4) is the following: we assume that cosine of the angle of x and $M_n(x)$ is greater than a fixed positive number (if x is not very near to the origin) and $\|M_n(x)\|$ is also greater than a fixed positive number (if $\|x\| \geqq \varepsilon/2$). In [1] the case was investigated when the cosine of the angle is greater than a positive number and $\|M_n(x)\| \geqq \varrho\|x\|$ i.e. we assumed than $M_n(x)$ has a linear lower bound. In this paper this condition is dropped.

Our aim is to prove the following:

THEOREM. *Under conditions* (1)—(4) *we have*

$$\mathsf{P}\{\|X_n\| \geqq \varepsilon\} \leqq \begin{cases} \exp\left(-C(\log n)^{1/\beta}\right) & \text{if} \quad 0 < \beta < \dfrac{1}{2} \\ \exp\left(-Cn^{\alpha/K}\right) & \text{if} \quad \beta = 0 \quad \text{and} \quad K \geqq \alpha \\ 0 & \text{if} \quad \beta = 0,\ K < \alpha \text{ and } n \text{ is large enough} \end{cases}$$

where C is a positive constant and ε is the positive number of condition (4).

§ 1. Lemmas

The following lemma was proved in [2].

LEMMA 1. ([2]) *Let $\eta_1, \eta_2, \ldots$ be a sequence of real valued r.v.'s for which*

$$\mathsf{E}(\eta_1) = \mathsf{E}(\eta_{n+1}|\eta_n, \ldots, \eta_1) = 0 \qquad (n = 1, 2, \ldots) \tag{6}$$

(with probability 1) and

$$\mathsf{P}(|\eta_n| \leqq K) = 1 \qquad (n = 1, 2, \ldots). \tag{7}$$

Further let $\{a_n\}$ be a monotonically decreasing sequence of positive numbers for which $\sum_{i=1}^{\infty} a_i^2 < \infty$. Then

$$\mathsf{P}\left\{\sum_{i=k}^{n} a_i \eta_i \geqq C\right\} \leqq \exp\left\{-\frac{1}{\vartheta} \frac{C^2}{2K^2 \sum_{i=k}^{\infty} a_i^2}\right\}$$

provided that

$$0 < C \leqq \frac{\vartheta \log \vartheta K \sum_{i=k}^{\infty} a_i^2}{|a_k|} \qquad (\vartheta > 0).$$

In [2] as a simple consequence of Lemma 1 we obtained:

LEMMA 2. ([2]) *Let $\eta_1, \eta_2, \ldots$ be a sequence of r.v.'s obeying conditions* (6) *and* (7). *Then*

$$\mathsf{P}\left(\sum_{i=k}^{n} \frac{\eta_i}{i} > R \sum_{i=k}^{n} \frac{1}{i} - S\right) \leqq 2 \exp\left(-An^{R/K}\right)$$

if $1 \leqq k \leqq \gamma n$ where $A > 0$ and $0 < \gamma < 1$ are constants depending on K, R and S only.

LEMMA 3. *Let $\eta_1, \eta_2, \ldots$ be a sequence of r.v.'s obeying conditions* (6) *and* (7). *Then*

$$\mathsf{P}\left(\sum_{i=\gamma n}^{n} \frac{\eta_i}{i} > R\right) \leqq \exp\left(-An\right)$$

for any $0 < \gamma < 1$ and $R > 0$ where A is a positive constant.

This lemma is a well-known big deviation theorem but it can be obtained as a consequence of Lemma 1 too. Hence the proof will be omitted.

LEMMA 4. *Let* $\eta_1, \eta_2, \ldots$ *be a sequence of r.v.'s for which* (6) *and* (7) *hold. Then*

$$\mathsf{P}\left\{\sum_{i=k}^{n} \frac{1}{i^{1-\beta}} \eta_i \geqq R \log \frac{n}{k} + S\right\} \leqq \exp\left(-T(\log n)^{1/\beta}\right) \tag{8}$$

$$\left(0 < \beta < \frac{1}{2};\ k = 1, 2, \ldots, n\right)$$

where R, S *and* T *are positive constants not depending on* k *and* n.

Remark 2. Clearly (8) is true even for negative S if $n-k$ is large enough.

PROOF. Put

$$l_n = \left(\frac{R\beta}{4K} \log n + 1\right)^{1/\beta}$$

$$C = \frac{R}{2} \log \frac{n}{k} + S, \quad a_i = \frac{1}{i^{1-\beta}}, \quad \lambda_n = \max(l_n, k),$$

$$\vartheta = \max\left(\frac{4(1-2\beta)K}{\beta}, e\right).$$

Then

$$\sum_{i=1}^{l_n} \frac{1}{i^{1-\beta}} K \leqq K \frac{(l_n+1)^\beta}{\beta} \leqq \frac{R}{2} \log n$$

(if n is large enough) and since

$$C = \frac{R}{2} \log \frac{n}{k} + S \leqq \frac{4K}{\beta} l_n^\beta \leqq \frac{\vartheta \log \vartheta \sum_{i=l_n}^{\infty} a_i^2}{a_{l_n}}$$

and

$$\frac{C^2}{\sum_{i=k}^{n} a_i^2} \geqq T_1 (\log n)^{1/\beta}$$

(for a suitable constant T_1 if n is large enough and $k \geqq l_n$) we have our Lemma 4. by Lemma 1.

LEMMA 5. *Under conditions* (1)—(4) *we have*

$$\|X_{n+1}\| \leqq \|X_n\| \left(1 + \frac{1}{2n^2}\left(A^2 + \frac{4AB}{\varepsilon}\right)\right) - \frac{\alpha}{n} - \frac{n^\beta}{n}\left(\frac{X_n}{\|X_n\|}, Y_n\right) + O\left(\frac{n^{2\beta}}{n^2}\right)$$

provided that $\|X_n\| \geqq \varepsilon/2$.

This lemma can be proved on the same way as Lemma 2 of [1] was proved, hence the proof will be omitted.

§ 2. Proof of the theorem

Introduce the following notations

$$\eta_n = -\left(\frac{X_n}{\|X_n\|},\ Y_n\right),$$

$$\mathfrak{A}_k^{(n)} = \mathfrak{A}_k = \left\{\|X_k\| \leqq \frac{\varepsilon}{2},\ \|X_{k+1}\| > \frac{\varepsilon}{2},\ \ldots,\ \|X_{n-1}\| > \frac{\varepsilon}{2},\ \|X_n\| > \varepsilon\right\},$$

$$\mathfrak{A}_0 = \left\{\|X_1\| > \frac{\varepsilon}{2},\ \|X_2\| > \frac{\varepsilon}{2},\ \ldots,\ \|X_{n-1}\| > \frac{\varepsilon}{2},\ \|X_n\| > \varepsilon\right\},$$

$$D = \frac{1}{2}\left(A^2 + \frac{4AB}{\varepsilon}\right).$$

Then clearly we have

$$\mathsf{P}\{\|X_n\| > \varepsilon\} = \sum_{k=0}^{n-1} \mathsf{P}(\mathfrak{A}_k) \tag{9}$$

and by Lemma 5.

$$\|X_n\| \leqq \|X_{l+1}\| \prod_{j=l+1}^{n-1}\left(1+\frac{D}{j^2}\right) + \sum_{j=l+1}^{n-1} \frac{1}{j^{1-\beta}}\eta_j - \sum_{j=l+1}^{n-1}\frac{\alpha}{j} + o_l(1)$$

provided that

$$\|X_{l+1}\| \geqq \frac{\varepsilon}{2},\ \|X_{l+2}\| \geqq \frac{\varepsilon}{2},\ \ldots,\ \|X_{n-1}\| \geqq \frac{\varepsilon}{2}.$$

Hence

$$\mathsf{P}(\mathfrak{A}_l) \leqq \mathsf{P}\left\{\sum_{j=l+1}^{n-1}\frac{1}{j^{1-\beta}}\eta_j \geqq \alpha\sum_{j=l+1}^{n-1}\frac{1}{j} + \|X_n\| - \|X_{l+1}\|\sum_{j=l+1}^{n-1}\left(1+\frac{D}{j^2}\right) + \right.$$

$$\left. + o_l(1);\ \|X_n\| \geqq \varepsilon,\ \|X_l\| \leqq \frac{\varepsilon}{2}\right\}, \tag{10}$$

$$\mathsf{P}(\mathfrak{A}_0) \leqq \mathsf{P}\left\{\sum_{j=1}^{n-1}\frac{1}{j^{1-\beta}}\eta_j \geqq \alpha\sum_{j=1}^{n-1}\frac{1}{j} - U\right\}$$

where U is a suitable positive constant.

Then our statements follows from (9), (10), and Lemmas 2, 3, 4. (including Remark 2).

REFERENCES

[1] RÉVÉSZ, P.: Robbins-Monro procedure in a Hilbert space I. In print.
[2] KOMLÓS, J.—RÉVÉSZ, P.: On the rate of convergence of the Robbins-Monro method. *Zeitschrift für Wahrscheinlichkeitstheorie* **25** (1972) 39—47.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received May 17, 1973)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 473—477.

CHARACTERIZATION OF CONTINUOUS RANDOM PROCESSES AND SIGNED MEASURES

by

O. KALLENBERG

Roughly stated, PROHOROV's well known theorem ([1], p. 37) implies that a sequence of random processes converges in distribution if and only if it is tight and convergence holds on some suitable determining class of sets ([1], p. 15). Establishing tightness is usually the most difficult part, and much effort has therefore been spent on searching for efficient tightness criteria (cf. [1], §§ 12, 15), while the determining class is often chosen without any discussion ([1], Th's 8.1, 15.1). However, simple characterizations of distributions may sometimes yield essential simplifications of the general criteria for convergence in distribution. Results in this direction were obtained for simple point processes in [4, 5] (independently of MÖNCH [8] and extending a result by RÉNYI [9]), and for processes with interchangeable increments in [7]. In the present paper, we shall show that the characterization of simple point processes in [4], Th. 2.1, and [8] essentially carries over to fairly general classes of continuous random processes and of non-atomic random signed measures described by certain mild restrictions on the local behavior as well as on the tails of the finite-dimensional distributions. Corresponding convergence criteria will follow as simple corollaries.

1. Continuous random processes

Consider the space $D_0[0, 1]$ of left continuous functions x with right-hand limits on $[0, 1]$ satisfying $x(0)=0$. Clearly, the relation

$$m[s, t) = x(t)-x(s), \quad 0 \leqq s \leqq t \leqq 1, \tag{1}$$

defines a unique correspondence between functions $x \in D_0[0, 1]$ and finitely additive set functions m on the ring $\mathcal{U}$ of finite unions of intervals $[s, t) \subset [0, 1]$. In particular, m extends to a (countably additive) signed measure on the Borel algebra in $[0, 1]$ if and only if x has bounded variation. In what follows, x and m will be identified and the two modes of notation in (1) will be used interchangeably. Furthermore, we write x_t for $x(t)$ and $\underset{h}{\Delta} x_t$ for $x_{t+h}-x_t$, and we use $\overset{\mathrm{d}}{=}$ to denote equality in distribution, which for random processes in $D_0[0, 1]$ means equality of the corresponding finite-dimensional distributions (cf. Th. 14.5 in [1]).

For any random process ξ on $[0, 1]$ we introduce the conditions

(C_1) The moment functions of ξ are bounded and the product moments determine the finite-dimensional distributions of ξ.

(C_2) For each $t \in [0, 1)$ there exists some $a>1$ such that

$$\lim_{h \downarrow 0} h^{-a} \mathsf{E} |\underset{h}{\Delta} \xi_t|^{2a} = 0.$$

Note that the product moments determine the finite-dimensional distributions e.g. provided the one-dimensional distributions have analytic (at zero) characteristic functions (cf. [2], pp. 514, 529). As for (C_2), note that the limit involved is positive but finite for the Brownian motion; hence the sample paths of ξ are allowed to vary almost as irregularly.

THEOREM 1. *Let ξ and η be random processes in $D_0[0, 1]$, and suppose that ξ satisfies (C_1) and (C_2). Then $\xi U \overset{d}{=} \eta U$, $U \in \mathcal{U}$, implies $\xi \overset{d}{=} \eta$.*

The point here is that only a simple type of one-dimensional distributions are needed for a full description of ξ (cf. [1], pp. 20, 121, and [4], Lemma 1.1). Though (C_2) forces a.s. continuity on [0, 1] by [6], the present formulation in terms of $D_0[0, 1]$ is preferable for applications. Note also that it is possible to replace ξ by $\xi \circ g$ in (C_2), where g is any continuous and strictly increasing function onto [0, 1] (cf. [6]).

PROOF. Suppose that $\xi U \overset{d}{=} \eta U$, $U \in \mathcal{U}$, and consider the relation

$$\mathsf{E}(\xi U)^m \xi_{t_1} \dots \xi_{t_n} = \mathsf{E}(\eta U)^m \eta_{t_1} \dots \eta_{t_n}, \quad U \in \mathcal{U}, \quad m \in Z_+, \tag{2}$$

which is trivially true for $n=0$. Suppose that (2) holds for some fixed $n \in Z_+$ and any $t_1 \leqq \dots \leqq t_n$ to the right of U. Given U and $\{t_j\}$, let t, $h>0$ be such that $[t, t+h) \cap U = \emptyset$ and $t+h \leqq t_1$. Writing $\xi = \xi_{t_1} \dots \xi_{t_n}$, $\boldsymbol{\eta} = \eta_{t_1} \dots \eta_{t_n}$, we get

$$\mathsf{E}(\xi U + \underset{h}{\Delta} \xi_t)^m \xi = \mathsf{E}(\eta U + \underset{h}{\Delta} \eta_t)^m \boldsymbol{\eta},$$

so by subtracting (2),

$$\mathsf{E}\{(\xi U + \underset{h}{\Delta} \xi_t)^m - (\xi U)^m\} \xi = \mathsf{E}\{(\eta U + \underset{h}{\Delta} \eta_t)^m - (\eta U)^m\} \boldsymbol{\eta},$$

or equivalently

$$\sum_{k=1}^{m} h^{-1} \binom{m}{k} \{\mathsf{E}(\underset{h}{\Delta} \xi_t)^k (\xi U)^{m-k} \xi - \mathsf{E}(\underset{h}{\Delta} \eta_t)^k (\eta U)^{m-k} \boldsymbol{\eta}\} = 0. \tag{3}$$

For fixed $k=2, \dots, m$, choose $b, c, d_1, \dots, d_n > 1$ with $a^{-1}+b^{-1}+c^{-1}+\sum_j d_j^{-1} = 1$ and apply Hölder's inequality to obtain

$$|h^{-1} \mathsf{E}(\underset{h}{\Delta} \eta_t)^k (\eta U)^{m-k} \boldsymbol{\eta}| \leqq \tag{4}$$

$$\leqq \{h^{-a} \mathsf{E}|\underset{h}{\Delta} \eta_t|^{2a}\}^{1/a} \{\mathsf{E}|\underset{h}{\Delta} \eta_t|^{(k-2)b}\}^{1/b} \{\mathsf{E}|\eta U|^{(m-k)c}\}^{1/c} \prod_{j=1}^{n} \{\mathsf{E}|\eta_{t_j}|^{d_j}\}^{1/d_j} =$$

$$= \{h^{-a} \mathsf{E}|\underset{h}{\Delta} \xi_t|^{2a}\}^{1/a} \{\mathsf{E}|\underset{h}{\Delta} \xi_t|^{(k-2)b}\}^{1/b} \{\mathsf{E}|\xi U|^{(m-k)c}\}^{1/c} \prod_{j=1}^{n} \{\mathsf{E}|\xi_{t_j}|^{d_j}\}^{1/d_j}.$$

Clearly, the same estimate is obtained with η replaced by ξ. Now the first factor in (4) tends to zero by (C_2), while the remaining ones are bounded by (C_1), so letting $h \downarrow 0$ in (3), we get

$$\frac{d}{dt} \{\mathsf{E}(\xi U)^{m-1} \xi_t \xi - \mathsf{E}(\eta U)^{m-1} \eta_t \boldsymbol{\eta}\} = 0, \tag{5}$$

where d/dt stands for the right-hand derivative. Thus (cf. [1], p. 155) the expression within curly brackets in (5) is constant on any open subinterval of $[0, t_1] \setminus U$, and this result extends by continuity to closed intervals. In fact, continuity of product moments follows by a.s. path continuity [6] and the uniform integrability implied by (C_1). For any $t \leqq t_1$ to the right of U we thus obtain with $V = [0, t) \setminus U$

$$\mathsf{E}(\xi U)^{m-1} \xi V \xi = \mathsf{E}(\eta U)^{m-1} \eta V \eta,$$

and adding (2), we get

$$\mathsf{E}(\xi U)^{m-1} \xi_t \xi = \mathsf{E}(\eta U)^{m-1} \eta_t \eta.$$

This proves (2) by induction for arbitrary n and $t_1, \ldots, t_n$, and in particular, ξ and η have the same product moments of all orders, so $\xi \overset{\mathrm{D}}{=} \eta$ by (C_1).

In the particular case when ξ (and hence η) is a.s. bounded by some real number, (C_1) holds automatically while it may be seen from the proof that Theorem 1 remains true with $a = 1$ in (C_2) (which is weaker than $a > 1$ by Jensen's inequality).

Let us finally indicate how Theorem 1 yields corresponding convergence criteria. Writing $\overset{\mathrm{d}}{\to}$ for convergence in distribution [1] and proceeding as in the proof of Theorem 15.1 in [1], we get

COROLLARY 1. *Let $\xi, \xi_1, \xi_2, \ldots$ be random processes in $D_0[0, 1]$, and suppose that ξ satisfies (C_1) and (C_2). Then $\xi_n \overset{\mathrm{d}}{\to} \xi$ in the Skorohod topology if and only if $\{\xi_n\}$ is tight and $\xi_n U \overset{\mathrm{d}}{\to} \xi U$, $U \in \mathcal{U}$.*

2. Non-atomic random signed measures

In the particular case of processes of bounded variation, the hypotheses in Theorem 1 may be relaxed, and they may even be completely removed whenever the processes are monotone. Results of this type carry over to general random measures, so let us consider some fixed locally compact second countable Hausdorff space S, and let $\mathcal{B}$ be the ring of bounded Borel sets in S. By a *random signed measure* on S we mean a mapping ξ of some abstract probability space into the space of locally finite signed measures on $\mathcal{B}$ such that ξB is a (measurable) random variable for each $B \in \mathcal{B}$. This means that ξ is the difference between two non-negative random measures as defined in [4].

We shall need the condition

(C_3) For any $n \in N$ and $B_1, \ldots, B_n \in \mathcal{B}$, the distribution of $(\xi B_1, \ldots, \xi B_n)$ is determined by those of $\sum_j t_j \xi B_j$, $t_1, \ldots, t_n \in R_+$.

By considering Laplace tranforms, it is seen as in [1], p. 49, that (C_3) holds automatically provided ξ is non-negative. In fact, by analytic continuation it even suffices that the Laplace transform of ξB be finite in some sub-interval of R_+ (or of R_-) for each $B \in \mathcal{B}$. Alternatively, we may assume the distribution of any vector $(\xi B_1, \ldots, \xi B_n)$ to be determined by its product moments, since the latter appear as coefficients in the polynomials $\mathsf{E}\{\sum_j t_j \xi B_j\}^k$, $k \in N$.

Generalizing the ring of interval unions considered in § 1, let $\mathcal{U} \subset \mathcal{B}$ be any DC-ring [4], i.e. a ring such that any $B \in \mathcal{B}$ may be covered by finitely many sets in $\mathcal{U}$ with arbitrarily small diameter (in any fixed metric).

THEOREM 2. *Let ξ and η be random signed measures on S, and suppose that ξ is non-atomic and satisfies* (C_3). *Then $\xi U \overset{d}{=} \eta U$, $U \in \mathscr{U}$, implies $\xi \overset{d}{=} \eta$.*

PROOF. For fixed $U \in \mathscr{U}$ and any $n \in N$, let $U_{n1}, \ldots, U_{nr_n} \in \mathscr{U}$ be a disjoint partitioning of U, chosen such that $\max_j |U_{nj}| \to 0$. Construct random variables $\xi_{n1}, \ldots, \xi_{nr_n}$ from $\xi U_{n1}, \ldots, \xi U_{nr_n}$ by random permutation independently of ξ, and similarly for $\eta_{n1}, \ldots, \eta_{nr_n}$. As in [7], define the processes

$$X_n(t) = \sum_{j < r_n t} \xi_{nj}, \quad Y_n(t) = \sum_{j < r_n t} \eta_{nj}, \quad t \in [0, 1], \quad n \in N.$$

Since ξ is non-atomic and U is bounded, it is seen by a simple compactness argument that $\max_j |\xi_{nj}| \to 0$ a.s., so we get

$$\sum_j \xi_{nj}^2 \leqq \max_j |\xi_{nj}| \sum_j |\xi_{nj}| \leqq \max_j |\xi_{nj}| \cdot |\xi| U \to 0 \text{ a.s.}, \quad n \to \infty,$$

and it follows by Th. 2.2 in [7] that $X_n(t) \overset{d}{\to} t\xi U$, $t \in [0, 1]$. On the other hand, $\{Y_n\}$ is tight in the Skorohod topology by Lemma 2.1 in [7] since $\sum_j \eta_{nj}^2 \leqq (|\eta| U)^2 < \infty$. If $Y_n \overset{d}{\to} Y$, $n \in N' \subset N$, then $Y(t) \overset{d}{=} t\xi U$, $t \in [0, 1]$, since $Y_n(t) \overset{d}{=} X_n(t) \overset{d}{\to} t\xi U$. We shall prove that Y must be linear. Indeed, let (α, σ, β) be the canonical random elements [7] of Y, and note that

$$\mathsf{E}[e^{iuY(t)} | \alpha, \sigma, \beta] = e^{iu\alpha t} \mathsf{E}[e^{iu(Y(t) - \alpha t)} | \sigma, \beta], \quad u \in R, \ t \in [0, 1],$$

where for $t = 1/2$ the second factor on the right is always real, and is < 1 for small $u > 0$ provided $(\sigma, \beta) \neq 0$. Taking real parts and integrating, we obtain for small u

$$\mathsf{E}\cos\left(uY\left(\tfrac{1}{2}\right)\right) < \mathsf{E}\cos(u\alpha/2), \tag{6}$$

unless $(\sigma, \beta) = 0$ a.s. Since (6) contradicts the fact that $Y(t) \overset{d}{=} \alpha t$, we must have $Y(t) = \alpha t$ a.s. By Th. 2.3 in [1], $Y_n \overset{d}{\to} Y$, $n \in N$, where $Y(t) = t\xi U$, so by Th. 2.2 in [7], $\sum_j \eta_{nj}^2 \overset{d}{\to} 0$, proving that η has a.s. no atoms in U. Hence η is a.s. non-atomic.

Conditionally on ξ and η respectively, we now obtain as for ξ above

$$X_n(t) \xrightarrow{d} t\xi U, \quad Y_n(t) \xrightarrow{d} t\eta U, \quad t \in [0, 1], \tag{7}$$

and since the limits are non-random, (6) remains true in the sense of convergence in probability. Therefore, it extends to the multi-dimensional case of arbitrary disjoint $U_1, \ldots, U_s \in \mathscr{U}$ and corresponding processes X_{kn}, Y_{kn}, $k = 1, \ldots, s$, $n \in N$, and in particular,

$$\sum_k X_{kn}(t_k) \xrightarrow{d} \sum_k t_k \xi U_k, \quad \sum_k Y_{kn}(t_k) \xrightarrow{d} \sum_k t_k \eta U_k, \quad t_1, \ldots, t_s \in [0, 1],$$

which by Fubini's theorem remains true unconditionally. But

$$\sum_k X_{kn}(t_k) \stackrel{d}{=} \sum_k Y_{kn}(t_k), \quad t_1, \ldots, t_s \in [0, 1], \quad n \in N,$$

so

$$\sum_k t_k \xi U_k \stackrel{d}{=} \sum_k t_k \eta U_k, \quad t_1, \ldots, t_s \in [0, 1],$$

and (C_3) completes the proof.

As in § 1, we may now write down corresponding convergence criteria. For brevity, we confine ourselves to the particularly important case of non-negative random measures. Proceeding as in the proof of Th. 1.1 in [4], we get

COROLLARY 2. *Let $\xi, \xi_1, \xi_2, \ldots$ be non-negative random measures on S and suppose that ξ is non-atomic with $\xi \partial U = 0$ a.s., $U \in \mathcal{U}$. Then $\xi_n \xrightarrow{d} \xi$ in the vague topology if and only if $\xi_n U \xrightarrow{d} \xi U$, $U \in \mathcal{U}$.*

Another case where one-dimensional convergence is enough is that of simple point processes, as shown in [4], § 2.

Note. A preliminary version of Theorem 1 was briefly announced in my lecture at the Point Process Conference at Göteborg in December 1971. For the particular case of non-negative random measures, GRANDELL [3] recently found a simpler proof based on subordination, showing that (C_1) and (C_2) are superfluous in that case. The above extension to random signed measures is new as are the present methods of proof.

REFERENCES

[1] BILLINGSLEY, P.: *Convergence of probability measures*. Wiley, New York 1968.

[2] FELLER, W.: *An introduction to probability theory and its applications* II (2nd ed.), Wiley, New York 1971.

[3] GRANDELL, J.: A note on characterization and convergence of nonatomic random measures. Akstract communicated at the Int. Conf. on Prob. Theory and Math. Stat. at Vilnius 1973.

[4] KALLENBERG, O.: Characterization and convergence of random measures and point processes, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **27** (1973), 9—21

[5] KALLENBERG, O.: Extremality of Poisson and sample processes, *J. Stoch. Proc. Appl.*, to appear.

[6] KALLENBERG, O.: Conditions for continuity of random processes without discontinuities of the second kind, *Ann. Probab.* **1** (1973), 519—526.

[7] KALLENBERG, O.: Canonical representations and convergence criteria for processes with interchangeable increments, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **27** (1973), 23—36.

[8] MÖNCH, G.: Verallgemeinerung eines Satzes von A. Rényi, *Studia Sci. Math. Hungarica* **6** (1971), 81—90.

[9] RÉNYI, A.: Remarks on the Poisson process, *Studia Sci. Math. Hungarica* **2** (1967), 119—123.

Göteborg

(Received May 2, 1973)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 479—483.

ASYMPTOTISCH OPTIMALE ABSCHÄTZUNGEN ZWISCHEN HALBNORMEN

von

R. GÜNTTNER

§ 1. Einleitung und Ergebnisse

Bezeichnungen: Im folgenden seien r_n und q_n $(n=1, 2, 3, \ldots)$ Halbnormen, d.h. nichtnegative, symmetrische, subadditive Funktionale, die auf einem Vektorraum X definiert sind. Sie sollen also den Bedingungen

$$p(f) \geqq 0, \quad p(\alpha f) = |\alpha| \cdot p(f),$$

$$p(f+g) \leqq p(f)+p(g),$$

für alle $f, g \in X$ und $\alpha \in \mathbf{R}$ genügen. Insbesondere folgt hieraus

$$p(f+g) \geqq p(f)-p(g).$$

Sei X ein normierter Raum. Ein Funktional p heiße beschränkt, wenn es eine Zahl M gibt, so daß für alle $f \in X$ gilt

$$p(f) \leqq M \cdot \|f\|.$$

Unter der Norm $\|p\|$ eines solchen Funktionals werde das Infimum über alle möglichen M verstanden; insbesondere gilt dann $p(f) \leqq \|p\| \cdot \|f\|$ für alle $f \in X$, und die Norm des Funktionals $q(f)=p(f)/\|p\|$ ist $\|q\|=1$.

Problemstellung: In der Analysis und speziell in der Approximationstheorie treten oft Abschätzungen eines (Rest-) Funktionals r_n durch ein anderes Funktional q_n auf. Sei $\|r_n\|=\|q_n\|=1$. Dann haben die Abschätzungen häufig die Form, daß für alle $f \in X$ gilt:

$$r_n(f) \leqq c \cdot q_n(f) \qquad (n = 1, 2, 3, \ldots) \tag{1.1}$$

($c \geqq 1$ eine Konstante; Beispiele siehe § 2). Die Abschätzung (1.1) ist für jedes feste n unverbesserbar, wenn zu jedem $\delta > 0$ und zu jedem $n \in \mathbf{N}$ ein $f_n \in X$ existiert mit

$$r_n(f_n) > (1-\delta) c \cdot q_n(f_n) \qquad (n=1, 2, 3, \ldots) \tag{1.2}$$

Die hier gestellte Frage lautet nun, unter welchen Voraussetzungen die Abschätzung (1.1) auch asymptotisch optimal ist: Gibt es ein (nicht mehr von n abhängiges) $f \in X$, so daß (1.2) für unendlich viele Werte von n gültig ist? — Daß dies nicht immer gilt, zeigt schon das einfache Beispiel $X=C[a, b]$ mit der üblichen sup-Norm, $r_n(f) = \frac{1}{2}\omega\left(f; \frac{1}{n}\right)$, $q_n(f)=\|f\|$, $c=1$. Andererseits gilt es zumindest für den trivialen Fall $r_n(f)=q_n(f)$.

Ergebnis: Im Falle $c=1$ gilt der einfache

SATZ 1. *Seien* r_n *und* q_n $(n=1, 2, 3, \ldots)$ *Halbnormen mit* $\|r_n\|=\|q_n\|=1$, *die auf dem Banachraum* X *definiert sind und die Ungleichung* $r_n(f)\leqq q_n(f)$ *erfüllen. Gilt nun* $q_n(f)\to 0$ $(n\to\infty)$ *für jedes* $f\in X$, *dann gibt es zu jedem* $\varepsilon>0$ *eine unendliche Teilfolge* n_k *natürlicher Zahlen und ein* $f\in X$ *mit* $q_{n_k}(f)>0$ *und*

$$r_n(f) > (1-\varepsilon)q_n(f) \qquad (n = n_1, n_2, n_3, \ldots)$$

Für den allgemeinen Fall $c\geqq 1$ wird folgendes vorausgesetzt: Zu jedem $\delta>0$ und zu jedem $n\in\mathbf{N}$ gebe es ein Element $f_n\in X$, $\|f_n\|=1$, das der Ungleichung (1.2) genügt; außerdem seien die f_n so wählbar, daß sie die Bedingungen

$$q_n(f_n) \nrightarrow 0 \quad (\text{für } n\to\infty) \tag{1.3}$$

für jedes feste δ, und

$$q_n(f_k) \to 0 \quad (\text{für } n\to\infty) \tag{1.4}$$

für jedes feste k und δ erfüllen. (In § 2 werden Beispiele angegeben, die den Bedingungen (1.1)—(1.4) genügen.) Es gilt nun

SATZ 2. *Seien* r_n *und* q_n $(n=1, 2, 3, \ldots)$ *Halbnormen mit* $\|r_n\|=\|q_n\|=1$, *die auf einem Banachraum* X *definiert sind und die Voraussetzungen* (1.1)—(1.4) *erfüllen: Dann gibt es zu jedem* $\varepsilon>0$ *eine unendliche Teilfolge* n_k *und ein* $f\in X$ *mit* $q_{n_k}(f)>0$ *und*

$$r_n(f) > (1-\varepsilon)c\cdot q_n(f) \qquad (n = n_1, n_2, n_3, \ldots)$$

Bemerkungen:

a) Voraussetzung (1.3) ist äquivalent mit der Bedingung

$$r_n(f_n) \nrightarrow 0 \quad (\text{für } n\to\infty), \tag{1.5}$$

die in manchen Fällen einfacher zu handhaben ist. Die Äquivalenz folgt aus (1.1) und (1.2).

b) Satz 1 ist ein Spezialfall von Satz 2: Ist $c=1$ in der Abschätzung (1.1), so ist (1.2) trivialerweise erfüllbar, denn es gibt zu jedem $\delta>0$ und n ein f_n mit $\|f_n\|=1$ und $r_n(f_n)>1-\delta$. Ebenso ist (1.5) und damit (1.3) erfüllt. Aus $q_n(f)\to 0$ $(n\to\infty)$ für jedes $f\in X$ folgt (1.4).

c) Für die Gültigkeit der Sätze 1 und 2 kann die Voraussetzung (1.1) abgeschwächt werden, indem in (1.1) die Konstante c durch einen größeren Wert d ersetzt wird; nur der Wert der Konstanten in (1.2) überträgt sich auf die Ungleichung des Satzes. Auch muß (1.2) nur für alle $n\geqq n_0(\delta)$ erfüllt sein.

§ 2. Beispiele

In den folgenden Beispielen i)—v) sind die Voraussetzungen von Satz 1 erfüllt; Beispiel vi) gibt eine Anwendung von Satz 2 für den Fall $c>1$.

Zunächst sei $X=C[-1, 1]$ der Vektorraum der stetigen Funktionen auf dem Intervall $I=[-1, 1]$. Dieser ist mit der üblichen Norm

$$\|f\| := \max_{x\in I} |f(x)|$$

ein Banachraum. Weiter bezeichne $E_n(f)$ die beste Approximation von f durch Polynome vom höchsten Grade n, und $\omega(f;\cdot)$ sei der Stetigkeitsmodul von f.

i) Sei $P_n[f]$ das Interpolationspolynom von f zu $n+1$ Stützstellen. Da Polynome vom Grade n durch P_n auf sich abgebildet werden, folgt

$$\|P_n[f]-f\| \leqq (\|P_n\|+1)\cdot E_n(f) \tag{2.1}$$

Hier ist $r_n(f) = \|P_n[f]-f\|/(\|P_n\|+1)$ und $q_n(f) = E_n(f)$. (Zu $\|P_n-I\| = \|P_n\|+1$ mit I identische Abb., $\|P_n\|$ Norm des Interpolationsoperators, siehe z. B. [5].)

ii) Es seien $-1 \leqq x_1^{(n)} < x_2^{(n)} < \ldots < x_{n+1}^{(n)} \leqq 1$ die zu $P_n[f]$ gehörigen Interpolationsknoten. Setzt man zur formelmäßigen Vereinfachung $x_0^{(n)} := -1$, $x_{n+2}^{(n)} := 1$, so gilt die in [3] bewiesene Abschätzung

$$\|P_n[f]-f\| \leqq \frac{1}{2}(\|P_n\|+1)\cdot\omega(f;\delta_n) \tag{2.2}$$

mit

$$\delta_n = \max_{0\leqq i\leqq n+1} |x_{i+1}^{(n)} - x_i^{(n)}|.$$

In diesem Fall ist $r_n(f)$ wie oben und $q_n(f) = \frac{1}{2}\omega(f;\delta_n)$

iii) Das Integral $I(f) = \int_{-1}^{1} f(x)\,dx$ werde approximiert durch die Riemannschen Summen

$$S(U_n, V_n; f) = \sum_{i=1}^{n} (y_{i+1}^{(n)} - y_i^{(n)})\cdot f(x_i^{(n)})$$

mit den Zerlegungen

$$U_n = \{y_1^{(n)} \leqq y_2^{(n)} \leqq \ldots \leqq y_{n+1}^{(n)}\}$$

$$V_n = \{x_1^{(n)} \leqq x_2^{(n)} \leqq \ldots \leqq x_n^{(n)}\}$$

wobei noch gelten soll $y_1^{(n)} = -1$, $y_{n+1}^{(n)} = 1$, $y_i^{(n)} \leqq x_i^{(n)} \leqq y_{i+1}^{(n)}$. Eine Anzahl von Quadraturformeln können durch solche Riemannschen Summen dargestellt werden, [1]. Ist Δ_n die Feinheit der Zerlegung $U_n \cup V_n$, so ist (vgl. [1])

$$|I(f)-S(U_n, V_n; f)| \leqq 2\omega(f;\Delta_n) \tag{2.3}$$

Hier ist $r_n(f) = \frac{1}{4}|I(f)-S(U_n, V_n; f)|$ und $q_n(f) = \frac{1}{2}\omega(f;\Delta_n)$.

Sei nun $X = \boldsymbol{C}_{2\pi}$ der Raum der stetigen 2π-periodischen Funktionen und

iv) $$a_n(f) = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos nx\,dx, \quad b_n(f) = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin nx\,dx$$

die Fourierkoeffizienten von f. Eine Abschätzung wie in [2] ergibt

$$|a_n(f)| \leqq \frac{2}{\pi}\omega\left(f;\frac{\pi}{n}\right) \tag{2.4}$$

[analog für $b_n(f)$.] Dabei ist jetzt $r_n(f) = \frac{\pi}{4}|a_n(f)|$ und $q_n(f) = \frac{1}{2}\omega\left(f;\frac{\pi}{n}\right)$.

v) Besitze f eine stetige Ableitung f'. Partielle Integration liefert

$$|a_n(f)| \leqq \frac{2}{n\pi}\,\omega\left(f';\ \frac{\pi}{n}\right) \tag{2.5}$$

In diesem Beispiel ist $X=C_{2\pi}^{(1)}$ der Banachraum der stetig differenzierbaren Funktionen mit der Norm

$$\|f\| = \max\left[\max_x |f(x)|,\ \max_x |f'(x)|\right]$$

Jetzt ist

$$r_n(f) = \frac{n\pi}{4}\,|a_n(f)|, \quad q_n(f) = \frac{1}{2}\,\omega\left(f';\ \frac{\pi}{n}\right).$$

vi) Bezeichne $E_n^*(f)$ die beste Approximation von $f\in C_{2\pi}$ durch trigonometrische Polynome vom höchsten Grade n. Dann gilt

$$E_n^*(f) \leqq \omega\left(f;\ \frac{\pi}{n+1}\right). \tag{2.6}$$

Hierzu vgl. [4] oder [6]. In diesem Beispiel ist also $r_n(f)=E_n^*(f), q_n(f)=\frac{1}{2}\,\omega\left(f;\ \frac{\pi}{n+1}\right)$, es gilt (1.1) für $c=2$. In [4] und [6] ist die Gültigkeit von (1.2) [für $n\geqq n_0(\delta)$, vgl. Bem. c) § 1] und (1.3) gezeigt. Wegen $q_n(f)\to 0$ für jedes f ist (1.4) trivialerweise erfüllt.

§ 3. Beweis der Sätze

Es genügt, nur Satz 2 zu beweisen [vgl. Bem. b) § 1]. Seien f_n die in (1.2)—(1.4) genannten Elemente, die außer von n auch noch von δ abhängen können. Wegen (1.3) gibt es eine Teilfolge m_l mit

$$q_{m_l}(f_{m_l}) > c_\delta > 0, \tag{3.1}$$

wobei c_δ nur noch von δ abhängt. Sei $\varepsilon(0<\varepsilon<1)$ vorgegeben. Mit

$$\delta = \frac{1}{2}\,\varepsilon \tag{3.2a}$$

ist auch c_δ festgelegt; damit sei

$$\gamma = \frac{1}{12}\,\varepsilon\cdot c_\delta \tag{3.2b}$$

Wegen $c_\delta\leqq 1$ ist $\gamma < \frac{1}{12}$. Aus der Teilfolge m_l wird nun eine Teilfolge m_{l_k} rekursiv ausgewählt und abkürzend mit n_k bezeichnet. n_0 ist frei wählbar und sei bereits festgelegt ($n_0\geqq n_0(\delta)$, vgl. Bem. c) § 1). Auf Grund von (1.4) und (1.1) kann für $k=1, 2, 3, \ldots$ die Folge n_k so gewählt werden, daß die Ungleichungen

$$r_{n_k}(f_{n_i}) \leqq \gamma^{k+1}, \qquad i = 0, 1, 2, \ldots, k-1 \tag{3.3}$$

$$q_{n_k}(f_{n_i}) \leqq \gamma^{k+1}, \qquad i = 0, 1, 2, \ldots, k-1 \tag{3.4}$$

erfüllt sind. Damit sei nun

$$f := \sum_{i=0}^{\infty} \gamma^i f_{n_i}. \tag{3.5}$$

Diese Reihe konvergiert wegen $\|f_{n_i}\|=1$ und $0<\gamma<\frac{1}{12}$, da X Banachraum ist. Aus (3.4) folgt nun für $k\geqq 1$

$$q_{n_k}(f) \leqq q_{n_k}(\gamma^k f_{n_k}) + \sum_{i=0}^{k-1} q_{n_k}(\gamma^i f_{n_i}) + q_{n_k}\left(\sum_{i=k+1}^{\infty} \gamma^i f_{n_i}\right) <$$

$$< \gamma^k q_{n_k}(f_{n_k}) + 2 \cdot \sum_{i=0}^{\infty} \gamma^{k+i+1},$$

also wegen $0<\gamma<\frac{1}{12}$

$$q_{n_k}(f) < \gamma^k [q_{n_k}(f_{n_k}) + 3\gamma]. \tag{3.6}$$

Ganz analog folgt aus

$$r_{n_k}(f) \geqq r_{n_k}(\gamma^k f_{n_k}) - \sum_{i=0}^{k-1} r_{n_k}(\gamma^i f_{n_i}) - r_{n_k}\left(\sum_{i=k+1}^{\infty} \gamma^i f_{n_i}\right)$$

zusammen mit (3.3) die Abschätzung

$$r_{n_k}(f) > \gamma^k [r_{n_k}(f_{n_k}) - 3\gamma].$$

Unter Berücksichtigung von (1.2) ergibt sich hieraus

$$r_{n_k}(f) > \gamma^k [c(1-\delta) q_{n_k}(f_{n_k}) - 3\gamma].$$

Die im Satz behauptete Ungleichung $r_{n_k}(f) > c(1-\varepsilon) q_{n_k}(f)$ ist also wegen (3.6) richtig, wenn gilt

$$c(1-\delta) q_{n_k}(f_{n_k}) - 3\gamma > c(1-\varepsilon)[q_{n_k}(f_{n_k}) + 3\gamma]$$

$$(\varepsilon - \delta) q_{n_k}(f_{n_k}) > 3\gamma \left[1 - \varepsilon + \frac{1}{c}\right].$$

Die rechte Seite ist wegen $c\geqq 1$ nicht größer als 6γ; auf Grund von (3.2) und (3.1) gilt die letzte Ungleichung und damit die im Satz behauptete Abschätzung. Hieraus folgt insbesondere $r_{n_k}(f)>0$; aus $q_{n_k}(f)=0$ würde aber wegen (1.1) auch $r_{n_k}(f)=0$ folgen, also gilt $q_{n_k}(f)>0$. Damit ist der Satz vollständig bewiesen.

LITERATUR

[1] BAKER, C. T. H.: On the Nature of certain Quadrature Formulas and their Errors; *SIAM J. Numer. Anal.*, Vol. 5, No. 4, 1968.

[2] BARY, N. K.: *A Treatise on Trigonometric Series;* Vol. I, p. 70/71; (Pergamon Press, Oxford, 1964).

[3] BRASS, H., GÜNTTNER, R.: Eine Fehlerabschätzung zur Interpolation stetiger Funktionen, *Studia Sci. Math. Hungar.* **8** (1973) 363—367.

[4] CHENEY, E. W.: *Introduction to Approximation Theory;* p. 144/45; (Mc Graw-Hill, New York, 1966).

[5] CHENEY, E. W., PRICE, K. H.: *Minimal Projections;* (in A. Talbot (Ed.)): Approximation Theory; Academic Press, London, 1970)

[6] KORNEICUK, N. P.: Die exakte Konstante in dem Theorem von D. Jackson über die gleichmäßige Approximation stetiger periodischer Funktionen (Russisch); *Dokl. Akad. Nauk*, **145**, 1962.

Techn. Universität Clausthal, BRD

(Eingegangen: 1. November, 1973)

Studia Scientiarum Mathematicarum Hungarica 8 (1973) 485—495.

GENERALIZATION OF A PROBLEM OF P. TURÁN

by

J. SZABADOS

1. Introduction

It is a well-known result of FABER that for an arbitrary system of nodes, there exists a continuous function for which the Lagrange interpolating polynomials do not converge uniformly to the function considered. On the other hand, as L. FEJÉR proved, the so-called Hermite—Fejér interpolating polynomials (prescribing the values of the function and 0 derivatives at the nodes) uniformly converge for arbitrary continuous functions, for a wide class of nodes. Working with n nodes, the degree of the Lagrange and Hermite—Fejér interpolating polynomials are $n-1$ and $2n-1$, respectively. Thus, by increasing the degree by n, we obtained a better interpolating polynomial. It is natural to ask whether the increment of the degree by n is optimal or not. One is inclined to think that the optimal degree is somewhere in the middle of $n-1$ and $2n-1$. Therefore it is rather surprising that if we do not prescribe the 0 derivative *in one single point*, the resulting procedure (a modified Hermite—Fejér interpolation) is no more convergent for all continuous functions. This question was initiated by E. FELDHEIM, and later investigated by P. TURÁN. The latter considered two special systems of nodes. In order to formulate the purpose of the present paper, we have to introduce some notations, and to quote the TURÁN's results [1].

Let

$$x_k = x_{kn} = \cos\theta_k = \cos\theta_{kn}, \quad \theta \leqq \theta_1 < \theta_2 < \ldots < \theta_n \leqq \pi \tag{1}$$

and

$$\omega_n(x) = c \prod_{k=1}^{n} (x - x_k) \quad (c > 0).$$

If $f(x)$ is an arbitrary function defined in $[-1, 1]$ then the so-called Hermite—Fejér interpolating polynomials

$$H_n(f, x) = \sum_{k=1}^{n} f(x_k) \left[1 - \frac{\omega_n''(x_k)}{\omega_n'(x_k)} (x - x_k)\right] \frac{\omega_n(x)^2}{\omega_n'(x_k)^2 (x - x_k)^2}$$

of degree $\leqq 2n-1$ based on (1) satisfy the relations

$$H_n(f, x_k) = f(x_k), \quad H_n'(f, x_k) = 0 \qquad (k = 1, \ldots, n). \tag{2}$$

It is well-known that if (1) is a so-called "strongly normal system" of nodes, i.e.

$$\min_{k=1,\dots,n} \min_{|x|\leqq 1} \left[1-\frac{\omega_n''(x_k)}{\omega_n'(x_k)}(x-x_k)\right] > 0 \tag{3}$$

then

$$\lim_{n\to\infty} \|f(x)-H_n(f,x)\|_{[-1,1]} = 0 \tag{4}$$

for all continuous function $f(x)$. (4) holds e.g. if $\omega_n(x)$ is the generalized Jacobi polynomial $P_n^{(\alpha,\beta)}(x)$ with $-1 \leqq \alpha, \beta < 0$.

Starting from an earlier result of E. FELDHEIM, P. TURÁN [1] investigated the following problem. Choose a $\nu(n)$, $1 \leqq \nu(n) \leqq n$, omit the condition $H_n'(f, x_{\nu(n)})=0$ in (2), and look for a polynomial $H^*_{\nu(n),n}(f,x)$ of degree $\leqq 2n-2$, satisfying the remaining conditions. It is easily seen that (cf. [1], (4.1))

$$H^*_{\nu(n),n}(f,x) = H_n(f,x) + \frac{\omega_n(x)^2}{x-x_{\nu(n)}} \sum_{k=1}^{n} f(x_k) \frac{\omega_n''(x_k)}{\omega_n'(x_k)^3}. \tag{5}$$

P. TURÁN [1] proved (among others) that if

$$|x_{\nu(n)}| \leqq 1-\varepsilon \quad \left(0 < \varepsilon \leqq \frac{1}{4}\right) \tag{6}$$

then the condition

$$\int_{-1}^{1} \frac{xf(x)}{\sqrt{1-x^2}}\,dx = 0 \tag{7}$$

when $\omega_n(x)=P_n^{(-1/2,\,-1/2)}(x)$ is the nth Chebyshev polynomial and the condition

$$f(-1) = f(1) \tag{8}$$

when $\omega_n(x)=P_n^{(-1,\,-1)}(x)$ is the integral of the Legendre polynomial, is necessary and sufficient to

$$\lim_{n\to\infty} \|f(x)-H^*_{\nu(n),n}(f,x)\| = 0 \tag{9}$$

where the norm is taken in $[-1, 1]$ or $[-1+\varepsilon, 1-\varepsilon]$ in the first and in the second case, resp. It is somewhat curious that the conditions (7) and (8) are of entirely different character.

The aim of the present paper is three-fold:

a) to generalize Turán's result for $\omega_n(x)=P_n^{(\alpha,\beta)}(x)$, $-1 \leqq \alpha, \beta < 0$;

b) to show the connection between the local and global conditions (7) and (8);* and

c) to generalize the results for the case when (6) is not assumed.

---

* This idea was raised by R. ASKEY.

2. The sufficient condition for the convergence

Let

$$w_{\alpha,\beta}(x) = (1-x)^{\alpha}(1+x)^{\beta}, \quad -1 < \alpha, \beta < 0, \quad |x| < 1$$

and $E_n(f)$ the best approximation of $f(x)$ in $[-1, 1]$ by polynomials of degree $\leqq n$. We prove the following

THEOREM 1. *If*

$$E_n(f) = \begin{cases} o(\theta_{\nu(n)}^{2\alpha+2}) & \text{for} \quad \theta_{\nu(n)} \leqq \dfrac{\pi}{2} \\ \\ o\big((\pi-\theta_{\nu(n)})^{2\beta+2}\big) & \text{for} \quad \theta_{\nu(n)} \geqq \dfrac{\pi}{2} \end{cases} \qquad (-1 \leqq \alpha, \beta < 0) \tag{10}$$

and

$$\lim_{\substack{\gamma\to\alpha+0 \\ \delta\to\beta+0}} \int_{-1}^{1} w'_{\gamma+1,\delta+1}(x) f(x)\, dx = 0 \tag{11}$$

then

$$\lim_{n\to\infty} \|f(x) - H^{*}_{\nu(n),n}(f, x)\|_{[-1,1]} = 0. \tag{12}$$

PROOF. By (4) and (5), (12) is equivalent to

$$\lim_{n\to\infty} \left[\left\| \frac{P_n^{(\alpha,\beta)}(x)^2}{x - x_{\nu(n)}} \right\| \sum_{k=1}^{n} f(x_k) \frac{P_n^{(\alpha,\beta)''}(x_k)}{P_n^{(\alpha,\beta)'}(x_k)^3} \right] = 0. \tag{12}$$

We now prove that

$$\left\| \frac{P_n^{(\alpha,\beta)}(x)^2}{x - x_{\nu(n)}} \right\| = \begin{cases} O(\theta_{\nu(n)}^{-2\alpha-2}) & \text{if} \quad \theta_{\nu(n)} \leqq \dfrac{\pi}{2} \\ \\ O\big((\pi-\theta_{\nu(n)})^{-2\beta-2}\big) & \text{if} \quad \theta_{\nu(n)} > \dfrac{\pi}{2}. \end{cases} \tag{13}$$

Without loss of generality we may assume that

$$x \quad \text{and} \quad x_{\nu(n)} > \frac{\beta-\alpha}{\alpha+\beta+1}. \tag{14}$$

Let the nearest node to x be $x_{\mu(n)}$. First we prove (13) for $\mu(n) \geqq \nu(n)$. Using the relations

$$\left(\frac{d}{dx}\right)^k P_n^{(\alpha,\beta)}(x) \sim \theta_{\mu(n)}^{-\alpha-k-1/2} n^{k-1/2}, \quad k = 0, 1$$

(cf. SZEGŐ [2], (7.32.10)), we get

$$\frac{P_n^{(\alpha,\beta)}(x)^2}{|x-x_{\nu(n)}|} \sim \begin{cases} \theta_{\nu(n)}^{-\alpha-3/2} n^{1/2} \theta_{\nu(n)}^{-\alpha-1/2} n^{-1/2} = O(\theta_{\nu(n)}^{-2\alpha-2}) & \text{if} \quad \mu(n) = \nu(n) \\ \dfrac{P_n^{(\alpha,\beta)}(x)^2}{|x-x_{\mu(n)}|} \left| \dfrac{x-x_{\mu(n)}}{x-x_{\nu(n)}} \right| = O(\theta_{\mu(n)}^{-2\alpha-2}) = O(\theta_{\nu(n)}^{-2\alpha-2}) & \text{if} \quad \mu(n) > \nu(n). \end{cases}$$

Assume now that $\mu(n) \leqq \nu(n)-1$. We may presume that $-\frac{1}{2} \leqq \alpha < 0$, because in the case $-1 \leqq \alpha < -\frac{1}{2}$ the local maxima of $P_n^{(\alpha,\beta)}(x)$ are decreasing in $[x_{\nu(n)}, 1]$. But then we can use the relation

$$\theta_k = \frac{k\pi + O(1)}{n} \tag{15}$$

as well (cf. [2], (8.9.1)), and obtain

$$\frac{P_n^{(\alpha,\beta)}(x)^2}{|x-x_{\nu(n)}|} = O(n^{2\alpha+2}) \frac{\mu(n)^{-2\alpha-1}}{\nu(n)^2-\mu(n)^2} = O\left(\left(\frac{n}{\nu(n)}\right)^{2\alpha+2}\right) = O(\theta_{\nu(n)}^{-2\alpha-2})$$

which proves (13).

Thus, if we show that

$$\left| \sum_{k=1}^{n} f(x_k) \frac{P_n^{(\alpha,\beta)''}(x_k)}{P_n^{(\alpha,\beta)'}(x_k)^3} \right| \leqq c_1 E_{n-2}(f)^* \tag{16}$$

then (12) follows by (10) and (13). Applying the differential equation

$$(1-x^2) P_n^{(\gamma,\delta)''}(x) + [\gamma - \delta - (\gamma+\delta+2)x] P_n^{(\gamma,\delta)'}(x) + n(n+\gamma+\delta+1) P_n^{(\gamma,\delta)}(x) = 0$$

(see [2], (4.2.1)), and the Cotes numbers

$$\lambda_k = \lambda_{k,n}^{(\gamma,\delta)} = \frac{\sigma_n}{P_n^{(\gamma,\delta)'}(x_k)^2 (1-x_k^2)} \qquad (k = 1, \ldots, n;\ \gamma, \delta > -1)$$

where

$$\sigma_n = \sigma_n^{(\gamma,\delta)} = 2^{\gamma+\delta+1} \frac{\Gamma(n+\gamma+1)\Gamma(n+\delta+1)}{\Gamma(n+1)\Gamma(n+\gamma+\delta+1)}, \qquad 0 < c_2 \leqq \sigma_n \leqq c_3 < \infty \tag{17}$$

(cf. [2], (15.3.1)) we obtain with the notation $g(x) = f(x) \cdot [(\gamma+\delta+2)x + \gamma - \delta]$,

$$\sum_{k=1}^{n} f(x_k^{(\gamma,\delta)}) \frac{P_n^{(\gamma,\delta)''}(x_k^{(\gamma,\delta)})}{P_n^{(\gamma,\delta)'}(x_k^{(\gamma,\delta)})^3} = \sum_{k=1}^{n} f(x_k^{(\gamma,\delta)}) \frac{(\gamma+\delta+2)x_k + \gamma - \delta}{P_n^{(\gamma,\delta)'}(x_k^{(\gamma,\delta)})^2 (1-x_k^{(\gamma,\delta)2})} = \tag{18}$$

$$= \frac{1}{\sigma_n^{(\gamma,\delta)}} \sum_{k=1}^{n} g(x_k^{(\gamma,\delta)}) \lambda_n^{(\gamma,\delta)} = \frac{1}{\sigma_n^{(\gamma,\delta)}} Q_n^{(\gamma,\delta)}(g)$$

---

* Here and in what follows $c_1, c_2, c_3, \ldots$ denote positive constants depending only on α and β.

where $Q_n^{(\gamma,\delta)}(g)$ is the quadrature sum to $g(x)$. If $q_{n-2}(x)$ is the best approximating polynomial of degree $\leqq n-2$ to $f(x)$ and $p_{n-1}(x) = [(\gamma+\delta+2)x+\gamma-\delta]\, q_{n-2}(x)$, then

$$\left| \sum_{k=1}^{n} f(x_k^{(\gamma,\delta)}) \frac{P_n^{(\gamma,\delta)''}(x_k^{(\gamma,\delta)})}{P_n^{(\gamma,\delta)'}(x_k^{(\gamma,\delta)})^3} \right| \leqq$$

$$\leqq \frac{1}{\sigma_n^{(\gamma,\delta)}} \left[|Q_n^{(\gamma,\delta)}(g-p_{n-1})| + \left| \int_{-1}^{1} w_{\gamma,\delta}(x)[p_{n-1}(x)-g(x)]\,dx \right| + \right.$$

$$\left. + \left| \int_{-1}^{1} w_{\gamma,\delta}(x) g(x)\,dx \right| \right] \leqq$$

$$\leqq \frac{1}{\sigma_n^{(\gamma,\delta)}} \left[2E_{n-2}(f) \int_{-1}^{1} |w'_{\gamma+1,\delta+1}(x)|\,dx + \left| \int_{-1}^{1} w'_{\gamma+1,\delta+1}(x) f(x)\,dx \right| \right] =$$

$$= \frac{1}{\sigma_n^{(\gamma,\delta)}} \left[E_{n-2}(f) 2^{\gamma+\delta+3} \frac{(\gamma+1)^{\gamma+1}(\delta+1)^{\delta+1}}{(\gamma+\delta+2)^{\gamma+\delta+2}} + \left| \int_{-1}^{1} w'_{\gamma+1,\delta+1}(x) f(x)\,dx \right| \right] \leqq$$

$$\leqq \frac{1}{\sigma_n^{(\gamma,\delta)}} 2\sqrt[e]{e} E_{n-2}(f) + \left| \int_{-1}^{1} w'_{\gamma+1,\delta+1}(x) f(x)\,dx \right| \right] \quad (-1 < \gamma, \delta < 0).$$

Hence and by (11), (17) we get

$$\left| \sum_{k=1}^{n} f(x_k^{(\alpha,\beta)}) \frac{P_n^{(\alpha,\beta)''}(x_k)}{P_n^{(\alpha,\beta)'}(x_k)^3} \right| \leqq \frac{2\sqrt[e]{e}}{\sigma_n^{(\alpha,\beta)}} E_{n-2}(f) \leqq \frac{2\sqrt[e]{e}}{c_2} E_{n-2}(f)$$

which proves (16). Q.e.d.

3. Remarks to Theorem 1

First we note that in case $-1<\alpha, \beta<0$ an equivalent form of condition (10) can be

$$E_n(f) = \begin{cases} o\left(\left(\dfrac{\nu(n)}{n}\right)^{2\alpha+2}\right) & \text{for} \quad \theta_{\nu(n)} \leqq \dfrac{\pi}{2} \\ o\left(\left(\dfrac{n-\nu(n)}{n}\right)^{2\beta+2}\right) & \text{for} \quad \theta_{\nu(n)} > \dfrac{\pi}{2}. \end{cases}$$

This follows from the relation (15).

From the proof of Theorem 1 it is also obvious that *if for a certain subsequence* $\{n_i\}$, $\theta_{\nu(n_i)} \leqq \dfrac{\pi}{2}$, $E_{n_i}(f) = o(\theta_{\nu(n_i)}^{2\alpha+2})$ *and* (11) *hold then*

$$\lim_{i\to\infty} \|f(x) - H^*_{\nu(n_i),n_i}(f,x)\|_{[-1,1]} = 0.$$

As for the second condition (11), we prove that

$$\lim_{\substack{\gamma\to\alpha+0\\ \delta\to\beta+0}} \int_{-1}^{1} w'_{\gamma+1,\,\delta+1}(x)f(x)\,dx = \tag{19}$$

$$= \begin{cases} \displaystyle\int_{-1}^{1} w'_{\alpha+1,\,\beta+1}(x)f(x)\,dx & \text{if} \quad -1<\alpha,\beta<0 \\ 2^{\alpha+1}f(-1)+(\alpha+1)\displaystyle\int_{-1}^{1}(1-x)^{\alpha}f(x)\,dx & \text{if} \quad -1<\alpha<0, \quad \beta=-1 \\ 2^{\beta+1}f(1)+(\beta+1)\displaystyle\int_{-1}^{1}(1+x)^{\beta}f(x)\,dx & \text{if} \quad \alpha=-1, \quad -1<\beta<0 \\ f(-1)-f(1) & \text{if} \quad \alpha=\beta=-1. \end{cases}$$

Notice that in the case $-1<\alpha<0$, $\beta=-1$ or $\alpha=-1$, $-1<\beta<0$ the condition (11) is of "mixed" type: it is of global and local character at the same time.

The first relation is obvious. In connection with the second, we get (using the previous notations)

$$\int_{-1}^{1} w'_{\gamma+1,\,\delta+1}(x)f(x)\,dx = \int_{-1}^{1} w'_{\gamma+1,\,\delta+1}(x)[f(x)-q_n(x)]\,dx + \tag{20}$$

$$+\int_{-1}^{1} w'_{\gamma+1,\,\delta+1}(x)q_n(x)\,dx.$$

Here (see the proof of Theorem 1)

$$\left|\int_{-1}^{1} w'_{\gamma+1,\,\delta+1}(x)[f(x)-q_n(x)]\,dx\right| \leqq \sqrt[e]{e}\,E_n(f) \tag{21}$$

and integrating by parts twice

$$\lim_{\substack{\gamma\to\alpha+0\\ \delta\to-1+0}} \int_{-1}^{1} w'_{\gamma+1,\,\delta+1}(x)q_n(x)\,dx = \lim_{\substack{\gamma\to\alpha+0\\ \delta\to-1+0}} \left(-\int_{-1}^{1} w_{\gamma+1,\,\delta+1}(x)q'_n(x)\,dx\right) =$$

$$= -\int_{-1}^{1}(1-x)^{\alpha+1}q'_n(x)\,dx = 2^{\alpha+1}q_n(-1)+(\alpha+1)\int_{-1}^{1}(1-x)^{\alpha}q_n(x)\,dx.$$

Thus

$$\left|\lim_{\substack{\gamma\to\alpha+0\\ \delta\to-1+0}} \int_{-1}^{1} w'_{\gamma+1,\,\delta+1}(x)f(x)\,dx - 2^{\alpha+1}f(-1)-(\alpha+1)\int_{-1}^{1}(1-x)^{\alpha}f(x)\,dx\right| \leqq$$

$$\leqq \sqrt[e]{e}\,E_n(f)+2^{\alpha+1}|q_n(-1)-f(-1)|+(\alpha+1)\left|\int_{-1}^{1}(1-x)^{\alpha}[q_n(x)-f(x)]\,dx\right| \leqq$$

$$\leqq \left(\sqrt[e]{e}+2^{\alpha+2}\right)E_n(f)$$

and here the right hand side tends to 0 as $n\to\infty$. If $\alpha=\beta=-1$, then (20)—(21) again hold, and

$$\lim_{\gamma,\delta\to -1+0}\int_{-1}^{1} w'_{\gamma+1,\delta+1}(x)q_n(x)\,dx = -\lim_{\gamma,\delta\to -1+0}\int_{-1}^{1} w_{\gamma+1,\delta+1}(x)q'_n(x)\,dx =$$

$$= -\int_{-1}^{1} q'_n(x)\,dx = q_n(-1)-q_n(1)$$

which proves the last relation in (19).

Now let us see how the TURÁN's results (7)—(8) are contained in Theorem 1. If we assume (6), then (10) is equivalent to the continuity of $f(x)$. This shows that in case (6) only the integral condition (11) appears in Theorem 1. Putting $\alpha=\beta=-\frac{1}{2}$ or $\alpha=\beta=-1$ in (11), (19) gives (7) or (8), respectively. In the case $\alpha=\beta=-1$ our result is even sharper than (9) because (6) is not necessary, further the convergence is uniform in the entire interval $[-1, 1]$.

As a generalization of FELDHEIM's result, if the conditions (10) and (11) do not hold, but $x_{v(n)}\in[a, b]\subset[-1, 1]$ $(n\geqq n_0)$ *then*

$$\lim_{n\to\infty} H^*_{v(n),n}(f, x) = f(x)$$

uniformly in $[-1, 1]\setminus[a-\varepsilon, b+\varepsilon)$ $(\varepsilon>0)$ *for all continuous function* $f(x)$. This can be seen from the proof of Theorem 1.

4. The necessity of the conditions (10) and (11)

First of all we show that (11) is necessary, not only to the uniform, but even to the pointwise convergence of $H^*_{v(n),n}(f, x)$.

THEOREM 2. *Let* $-1\leqq\alpha,\beta<0$, $-1\leqq\xi\leqq 1$, $\xi\neq-\alpha^{-1},\beta^{-1}$, *and* $f(x)$ *an arbitrary continuous function in* $[-1, 1]$ *for which* (11) *does not hold. Then there exists a sequence* $\{v(n)\}$ *such that* $\lim_{n\to\infty} H^*_{v(n),n}(f,\xi)$ *does not exist.*

REMARK. This result has been proved by P. TURÁN [1] in the special case $\alpha=\beta=-\frac{1}{2}$, $\xi=\cos\frac{\pi}{5}$, $f(x)=x$.

PROOF. (18) implies that

$$\lim_{n\to\infty}\sum_{k=1}^{n} f(x_k^{(\alpha,\beta)})\frac{P_n^{(\alpha,\beta)''}(x_k^{(\alpha,\beta)})}{P_n^{(\alpha,\beta)'}(x_k^{(\alpha,\beta)})^3} = c(\alpha,\beta)\lim_{\substack{\gamma\to\alpha+0\\ \delta\to\beta+0}}\int_{-1}^{1} w'_{\gamma+1,\delta+1}(x)f(x)\,dx \neq 0.$$

Thus, by (5), we only have to prove that with a suitable choice of $\{v(n)\}$,

$$\lim_{n\to\infty}\frac{P_n^{(\alpha,\beta)}(\xi)^2}{\xi-x_{v(n)}}$$

does not exist. At first let $\xi=1$. Then with $\theta_{\nu(n)}\geqq n^{\alpha}$ we have

$$\frac{P_n^{(\alpha,\beta)}(1)^2}{1-x_{\nu(n)}}\sim\frac{n^{2\alpha}}{\theta_{\nu(n)}^2}=O(1),$$

while choosing $0<\theta_{\nu(n)}\leqq n^{(\alpha-1)/2}$ (this is possible because by assumption $\alpha>-1$) we get

$$\frac{P_n^{(\alpha,\beta)}(1)^2}{1-x_{\nu(n)}}\geqq c_4 n^{\alpha+1}\to\infty\quad\text{as}\quad n\to\infty.$$

The case $\xi=-1$ can be similarly treated.

Now let $-1<\xi=\cos\theta<1$ and apply the asymptotic formula

(22)
$$P_n^{(\alpha,\beta)}(\xi)=(\pi n)^{-1/2}\left(\sin\frac{\theta}{2}\right)^{-\alpha-1/2}\left(\cos\frac{\theta}{2}\right)^{-\beta-1/2}\left[\cos(N\theta+\gamma)+O\left((n\sin\theta)^{-1}\right)\right]$$

$$\left(N=n+\frac{\alpha+\beta+1}{2},\quad\gamma=-\frac{(2\alpha+1)\pi}{4},\quad-1<\xi<1,\quad\alpha,\beta\geqq-1\right)$$

(cf. [2], (8.21.18)). Evidently, there exist sequences $\{n_j\}$ and $\{\nu(n_j)\}$ such that

$$|\cos(N_j\theta+\gamma)|\geqq c_5>0\quad\text{and}\quad 0<\theta-\theta_{\nu(n_j)}\leqq c_6 n_j^{-1}.$$

Then we have by (22)

$$\frac{P_n^{(\alpha,\beta)}(\xi)^2}{\xi-x_{\nu(n)}}\leqq -c_1(\alpha,\beta,\xi)\frac{n_j^{-1}}{(\theta^2-\theta_{\nu(n_j)}^2)}\leqq -c_2(\alpha,\beta,\xi)<0$$

while choosing $0<\theta_{\nu(n_j)}-\theta\leqq c_6 n_j^{-1}$ we get

$$\frac{P_n^{(\alpha,\beta)}(\xi)^2}{\xi-x_{\nu(n_j)}}\geqq c_3(\alpha,\beta,\xi)>0.$$

Q.E.D.

Next we show that if $f(x)$ has certain structural properties then the above oscillation of $H^*_{\nu(n),n}(f,x)$ is *bounded.*

THEOREM 3. *If*

$$E_n(f)=\begin{cases}O(\theta_{\nu(n)}^{2\alpha+2}) & \text{for}\quad\theta_{\nu(n)}\leqq\dfrac{\pi}{2}\\[2ex] O\left((\pi-\theta_{\nu(n)})^{2\beta+2}\right) & \text{for}\quad\theta_{\nu(n)}>\dfrac{\pi}{2}\end{cases}\qquad(-1\leqq\alpha,\beta<0)\tag{23}$$

and (11) *holds then*

$$\|H^*_{\nu(n),n}(f,x)\|_{[-1,1]}=O(1).\tag{24}$$

This can be easily verified from the proof of Theorem 1.

COROLLARY. *If* (6) *holds then* (24) *is true for all bounded functions.*

Finally, we prove that *the "o" in* (10) *cannot be replaced by "O".* For the sake of simplicity, we restrict ourselves to the case when $\theta_{v(n)}$ is bounded away from π, but is would be easy to handle the general case, too.

THEOREM 4. *Given* $-1 \leqq \alpha, \beta < 0$ *and the sequence* $\{v(n)\}$ *such that* $\theta_{v(n)} \leqq \frac{\pi}{2}$ *and* $n\theta_{v(n)}^{\alpha+1}$ *is monotone, there exists a function* $\bar{f}(x)$ *such that* (11) *and* (23) *hold, but*

$$\limsup_{n\to\infty} \|\bar{f}(x) - H^*_{v(n),n}(\bar{f}, x)\|_{[-1,1]} > 0. \tag{25}$$

PROOF. If (6) holds or $\alpha = -1$ then (23) requires only the boundedness of $f(x)$, and this is clearly not enough to (12), even if (11) holds. Hence we may assume that $\theta_{v(n)}$ monoton decreases to 0, and $-1 < \alpha < 0$.

Let $\xi = \dfrac{\beta-\alpha}{\alpha+\beta+2}$, $-1 \leqq \xi < 1$ be the unique root of $w'_{\alpha+1,\beta+1}(x)$ and let l and m be such that

$$-1 \leqq \xi < x_{l+1} \leqq \frac{3\xi+1}{4} < x_l < x_{m+1} \leqq \frac{2\xi+1}{3} < x_m < 1.$$

Define

$$g_n(x) = \begin{cases} 0 & \text{if } -1 \leqq x \leqq x_{l+1} \text{ or } x_m \leqq x \leqq 1 \\ \dfrac{(x-x_{l+1})^2\{(x_l-x_{l+1})w'_{\alpha+1,\beta+1}(x_l) - [2w'_{\alpha+1,\beta+1}(x_l) - (x_l-x_{l+1})w''_{\alpha+1,\beta+1}(x_l)](x-x_l)\}}{(x_l-x_{l+1})^3 w'_{\alpha+1,\beta+1}(x)} + \\ + a_n \dfrac{(x-x_l)^2(x-x_{l+1})^2}{(x_l-x_{l+1})^3 w'_{\alpha+1,\beta+1}(x)} & \text{if } x_{l+1} \leqq x \leqq x_l \\ 1 - \dfrac{30[w_{\alpha+1,\beta+1}(x_k) - w_{\alpha+1,\beta+1}(x_{k+1})](x-x_k)^2(x-x_{k+1})^2}{(x_k-x_{k+1})^5 w'_{\alpha+1,\beta+1}(x)} & \text{if } x_{k+1} \leqq x \leqq x_k, \\ & k = m+1, m+2, \ldots, l-1, l \\ \dfrac{(x-x_m)^2(x_{m+1}-x_m)w'_{\alpha+1,\beta+1}(x_{m+1})}{(x_{m+1}-x_m)^3 w'_{\alpha+1,\beta+1}(x)} + \\ + \dfrac{(x-x_m)^2(x-x_{m+1})[b_n(x-x_{m+1}) - 2w'_{\alpha+1,\beta+1}(x_{m+1}) + (x_{m+1}-x_m)w''_{\alpha+1,\beta+1}(x_{m+1})]}{(x_{m+1}-x_m)^3 w'_{\alpha+1,\beta+1}(x)} \\ & \text{if } x_{m+1} \leqq x \leqq x_m \end{cases}$$

where a_n and b_n are such that

$$\int_{x_{l+1}}^{x_l} w'_{\alpha+1,\beta+1}(x) g_n(x)\,dx = \int_{x_{m+1}}^{x_m} w'_{\alpha+1,\beta+1}(x) g_n(x)\,dx = 0.$$

Easy to see that such a_n and b_n exist and $a_n, b_n = O(n)$.

Let us list some properties of $g_n(x)$:

$$\|g_n(x)\| \leqq c_7, \tag{26}$$

$$g_n(x_k) = 1, \quad g_n(x_k) = 0 \qquad (k = m+1, m+2, \ldots, l-1, l), \tag{27}$$

$$\max_{[-1, x_{l+1}) \cup \bigcup_{k=l}^{m} (x_{k+1}, x_k) \cup (x_m, 1]} |g_n''(x)| = O\left(\frac{1}{\min\limits_{m \leqq k \leqq l} (x_k - x_{k+1})^2}\right) = O(n^2), \tag{28}$$

and

$$\lim_{\delta \to \beta + 0} \int_{-1}^{1} w'_{\alpha+1, \delta+1}(x) g_n(x)\, dx = 0. \tag{29}$$

As we have seen in (5), (25) depends on the behaviour of

$$R_n(f, x) \overset{\text{def}}{=} \frac{P_n^{(\alpha,\beta)}(x)^2}{x - x_{\nu(n)}} \sum_{k=1}^{n} f(x_k) \frac{P_n^{(\alpha,\beta)''}(x_k)}{P_n^{(\alpha,\beta)'}(x_k)^3}. \tag{30}$$

Using the point y_n constructed in (23), and the relation

$$\lambda_\nu \sim \nu^{2\alpha+1} n^{-2\alpha-2}, \quad 0 < \theta_\nu < \frac{\pi}{2}$$

(see [2], (15.3.14)) we obtain by (13) and (27)

$$R_n(g_n, y_n) \geqq c_8 \theta_{\nu(n)}^{-2\alpha-2} \sum_{k=m+1}^{l} [(\alpha+\beta+2)x_k + \alpha - \beta] \lambda_k \tag{31}$$

$$\geqq c_9 \theta_{\nu(n)}^{-2\alpha-2} n^{-2\alpha-2} \sum_{k=m+1}^{l} \nu^{2\alpha+1} \geqq c_{10} \theta_{\nu(n)}^{-2\alpha-2} \quad (0 < c_{10} < 1).$$

Further we have by (13), (18) and (27)

$$\|R_m(g_n, x)\| = O(\theta_{\nu(m)}^{-2\alpha-2}) Q_m(g_n) \leqq c_{11} \theta_{\nu(m)}^{-2\alpha-2} \qquad (m = 1, 2, \ldots).$$

Now define a sequence of indices $\{n_i\}$ as follows:

$$\theta_{\nu(n_{i+1})}^{2\alpha+2} \leqq \frac{c_{10}}{4} \theta_{\nu(n_i)}^{2\alpha+2}; \quad \theta_{\nu(n_1)}^{2\alpha+2} \leqq \frac{c_{10}}{4} \tag{32}$$

(this is possible since $\lim\limits_{n \to \infty} \theta_{\nu(n)} = 0$);

$$n_{i+1} \theta_{\nu(n_{i+1})}^{\alpha+1} \geqq 2 n_i \theta_{\nu(n_i)}^{\alpha+1} \tag{33}$$

(this is possible since evidently $\lim\limits_{n \to \infty} n \theta_{\nu(n)}^{\alpha+1} = \infty$);

$$\|R_{n_k}(g_{n_i}, x)\| \leqq 1 \quad (k \geqq i) \tag{34}$$

MTA Könyvtára
Periodika 3827/19 74 sz.

(if this were not possible for a certain i, then by (29) and $E_n(g_{n_i})=O(n^{-2})=O(\theta_{\nu(n)}^{2\alpha+2})$, $g_{n_i}(x)$ would be a suitable counterexample). Let

$$\bar{f}(x) = \sum_{i=1}^{\infty} \theta_{\nu(n_i)}^{2\alpha+2} g_{n_i}(x). \tag{35}$$

Here by (32) and (27) the right hand side series uniformly converges in $[-1, 1]$. Let $0<h\leqq\frac{1}{n_1}$ be an arbitrary positive number, and j that index for which $n_j \leqq \frac{1}{h} < n_{j+1}$. Then we have by (35), (28), (27), (32) and (33)

$$|\bar{f}(x+h)-2\bar{f}(x)+\bar{f}(x-h)| \leqq \left(\sum_{i=1}^{j}+\sum_{i=j+1}^{\infty}\right)\theta_{\nu(n_i)}^{2\alpha+2}|g_{n_i}(x+h)-2g_{n_i}(x)+g_{n_i}(x-h)| \leqq$$

$$\leqq h^2 \sum_{i=1}^{j} \max|g_{n_i}''(x)|\cdot\theta_{\nu(n_i)}^{2\alpha+2} + O\left(\sum_{i=j+1}^{\infty}\theta_{\nu(n_i)}^{2\alpha+2}\right) = O(h^2)\sum_{i=1}^{j} n_i^2\theta_{\nu(n_i)}^{2\alpha+2} +$$

$$+O(\theta_{\nu(n_{j+1})}^{2\alpha+2}) = O(h^2 n_j^2\theta_{\nu(n_j)}^{2\alpha+2}+\theta_{\nu(n_{j+1})}^{2\alpha+2}).$$

If especially $|h| \leqq \frac{1}{n}$ then we get by the monotonicity of $n\theta_{\nu(n)}^{\alpha+1}$

$$E_n(\bar{f}) = O\left(\omega_2\left(\bar{f}, \frac{1}{n}\right)\right) = O(\theta_{\nu(n)}^{2\alpha+2})$$

i.e. (24) holds. (11) obviously fulfils by (29). Finally, we obtain by (35), (31), (32) and (34)

$$R_{n_k}(\bar{f}, y_{n_k}) \geqq \theta_{\nu(n_k)}^{2\alpha+2} R_{n_k}(g_{n_k}, y_{n_k}) - \left(\sum_{i=1}^{k-1}+\sum_{i=k+1}^{\infty}\right)\theta_{\nu(n_i)}^{2\alpha+2}|R_{n_k}(g_{n_i}, y_{n_k})| \geqq$$

$$\geqq c_{10} - \theta_{\nu(n_1)}^{2\alpha+2}\frac{1}{1-\frac{1}{4}c_{10}} - \frac{1}{4}c_{10}\frac{1}{1-\frac{1}{4}c_{10}} \geqq c_{10} - \frac{c_{10}}{4}\cdot\frac{4}{3} - \frac{c_{10}}{4}\cdot\frac{4}{3} = \frac{c_{10}}{3} > 0.$$

Q.e.d.

REMARK. By Theorem 3, we have (24) for the constructed $f(x)$. This together with (25) means that $H_{\nu(n),n}^{*}(f, x)$ oscillates between finite bounds but it does not convergence uniformly to $f(x)$. This phenomenon is similar to the one described in Theorem 2.

REFERENCES

[1] TURÁN, P.: A remark on Hermite—Fejér interpolation, *Annales Univ. Sci. Budapest, Sectio Math.*, **3—4** (1960—61), 369—377.

[2] SZEGŐ, G.: *Orthogonal Polynomials*, American Math. Soc. Coll. Publ., Vol. XXIII (New York, 1959).

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received November 1, 1973)

INDEX

Máté, L.: On a representation of module-homomorphisms (multipliers) 187
Mogyoródi, J.: On the rarefaction of renewal processes, V. 193
Mogyoródi, J.: On the rarefaction of renewal processes, VI. 199
Sujan Srivastava.: Joint distributions based on runs and the other number of intersections 211
Simonovits, A.: Direct comparison of different priority queuing disciplines 225
Bihari, I.: Perturbations of n dimensional linear systems of differential equations 245
Scherer, K.: On extensions of a theorem of S. N. Bernstein 269
Gerencsér, L.: On a theorem of Lyapunoff 273
Freud, G.: Investigations on weighted approximation by polynomials 285
Kapoor, S. F. and *Stewart, M. J.:* A note on the Hamiltonian index of graphs 307
Fritz, J.: Stochastic approximation for finding local maxima of probability measures 309
Невai, Г. П. (Névai, G. P.): Две теоремы об отклонении сумм Фурье 323
Komlós, J. and *Révész, P.:* A modification of the Robbins-Monro process 329
Nemetz, T.: Bounds on error probability 341
Major, P.: On a non-parametric estimation of the regression function 347
Brass, H. und *Günttner, R.:* Eine Fehlerabschätzung zur Interpolation stetiger Funktionen 363
Halász, G.: On a result of Salem and Zygmund concerning random polynomials 369
Katona, G.: A three part Sperner theorem 379
Révész, P.: Robbins-Monro procedure in a Hilbert space and its application in the theory of learning processes I. 391
Freud, G. and *Németh, G.:* On the L_p-norms of orthonormal Hermite functions 399
Körner, J.: An extension of the class of perfect graphs 405
Berkes, I.: Functional limit theorems for lacunary trigonometric and Walsh series 411
Berkes, I.: On Strassen's version of the loglog law for multiplicative systems 425
Cvetković, D. M.: Inequalities obtained on the basis of the spectrum of the graph 433
Saxena, R. B.: Approximation of continuous functions by polynomials involving Freud's interpolatory polynomials 437
Joshi, V. N.: On the class of matrices which are convergent for a given matrix 447
Joshi, V. N.: Remarks on iterative methods for computing the generalized inverse 457
Andréka, H. and Németi, I.: On the equivalence of sets definable by satisfaction and ultrafilters 463
Révész, P.: Robbins-Monro procedure in a Hilbert space, II. 469
Kallenberg, O.: Characterization of continuous random processes and signed measures 473
Günttner, R.: Asymptotisch optimale Abschätzungen zwischen Halbnormen 479
Szabados, J.: Generalization of a problem of P. Turán 485

Printed in Hungary

A kiadásért felel az Akadémiai Kiadó igazgatója — Műszaki szerkesztő: Botyánszky Pál
A kézirat a nyomdába érkezett: 1973. VI. 11. — Terjedelem: 27,3 (A/5) ív, 15 ábra

73-2758 Szegedi Nyomda

MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

Die *Studia Scientiarum Mathematicarum Hungarica* ist eine Halbjahrsschrift der Ungarischen Akademia der Wissenschaften. Sie veröffentlicht Originalbeiträge aus dem Bereiche der Mathematik in deutscher, englischer, französischer oder russischer Sprache. Es erscheint jährlich ein Band.

Adresse der Redaktion: 1053 Budapest V., Reáltanoda u. 13—15, Ungarn.
Technischer Redaktor: Gy. Petruska

Abonnementspreis pro Band (pro Jahr): $ 16.00. Bestellbar bei Buch- und Zeitungs-Aussenhandelsunternehmen *Kultúra* (Budapest 62, P. O. B. 149), oder bei den Vertretungen im Ausland.
Austauschabmachungen können mit der Bibliothek des Matematischen Instituts (1061 Budapest V., Reáltanoda u. 13—15) getroffen werden.
Die zur Veröffentlichung bestimmten Manuskripte sind in zwei Examplaren an die Redaktion zu schicken.

Studia Scientiarum Mathematicarum Hungarica est une revue biannuelle de l'Académie Hongroise des Sciences publiant des essais originaux, en francais, anglais, allemand ou russe, du domaine des mathématiques.

Rédaction: 1053 Budapest V., Reáltanoda u. 13—15, Hongrie.
Rédacteur téchnique: Gy. Petruska

Le prix de l'abonnement: $ 16.00 par an (volume). On s'abonne chez *Kultúra*, Société pour le Commerce de Livres et Journaux (Budapest 62, P. O. B. 149) ou chez ses représentants à l'étranger.
Pour établir des relations d'échange on est prié de s'adresser à la Bibliothèque de l'Institut de Matématique (1053 Budapest V., Reáltanoda u. 13—15).
On est prié 'envoyer les articles destinés à la publication en deux exemplaires à l'adresse de la rédaction.

Studia Scientiarum Mathematicarum Hungarica — выходит два раза в год в Издании Академии Наук Венгрии. Журнал публикует оригинальные исследования в области математики на немецком, английском, французском и русском языках. Отдельвные вышуски составляют ежегодно одни том.

Адрес редакции: 1053 Budapest V., Reáltanoda u. 13—15, Венгрия.
Технический редактор: Gy. Petruska.

Подписная цена на год (за одниии том): $ 16.00. Подписка на журнал принимается Внешнеторговым предприятием „Култура" (Budapest 62, P. O. B. 149) или его представительствами за границей.
По поводу отношения обмена просим обрашаться к Библиотеке Института Математики (1053 Budapest V., Reáltanoda u. 13—15).
Работы, предназначенные для опубликования в журнале следует направлять по адресу редакции в двух экземплярах.

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.

CHECHOSLOVAKIA

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 Knjga
Terazije 27.
Beograd

31. VII. 1974. **Index: 26 748**