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 VI.
FASC. 1–2.
1971

AKADÉMIAI KIADÓ, BUDAPEST

Studia Scientiarum Mathematicarum Hungarica

A Magyar Tudományos Akadémia matematikai folyóirata

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

Technikai szerkesztő: Petruska Gy.

Kiadja az Akadémiai Kiadó, 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 Hírlap Külkereskedelmi Vállalatnál (Budapest II., Fő u. 32).

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

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: 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 and 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 (Budapest V., Reáltanoda u. 13—15.)

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

Studia Scientiarum Mathematicarum Hungarica

Auxilio
Consilii Instituti Mathematici
Academiae Scientiarum Hungaricae

Redigit

L. Fejes Tóth

Tomus VI

Akadémiai Kiadó, Budapest

1971

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 3—22.

ALFRÉD RÉNYI'S WORKS*

1945

1. Egy Stieltjes féle integrálról. Doktori értekezés. (Doctoral dissertation.) Szeged, 1945. 40 p. (Kézirat.) (Manuscript.)

1946

1. On a Tauberian theorem of O. Szász. Acta Sci. Math. Szeged 11 (1946/48) 119—123.
2. Integral formulae in the theory of convex curves. Acta Sci. Math. Szeged 11 (1946/48) 158—166.

1947

1. On the minimal number of terms of the square of a polynomial. Hung. Acta Math. 1 (1946/49) No. 2., 30—34.
2. О представлении четных чисел в виде суммы одного простого и одного почти-простого числа. Доклады Акад. Наук СССР 56 (1947) 455—458.
3. О представлении четных чисел в виде суммы простого и одного почти-простого числа. Kandidátusi disszertáció. Leningrád, 1947. (Kézirat.)
4. Об одном новом применении метода академика И. М. Виноградова. Доклады Акад. Наук СССР 56 (1947), 675—678.
5. О некоторых гипотехаз теории характеров Дирихле (с Ю. В. Линником). Изв. Акад. Наук СССР 11 (1947) 539—546.

1948

1. О представлении четных чисел в виде суммы простого и почти-простого числа. Изв. Акад. Наук СССР 12 (1948) 57—78. (cf. 1947/3).
2. Játék a véletlennel. Középisk. Mat. Lapok 1 (1948) 101—111.
3. Játék a véletlennel II. Középisk. Mat. Lapok 1 (1948) 144—157.
4. Simple proof of a theorem of Borel and of the law of the iterated logarithm. Mat. Tidsskrift B, 1948, 41—48.
5. Remarque à la note précédente. (G. Alexits: Sur la convergence des séries lacunaires. Acta Sci. Math. Szeged 11 (1946/48) 251—253.) Acta Sci. Math. Szeged 11 (1946/48) 253.
6. Generalization of the „large sieve" of Ju. V. Linnik. Math. Centrum, Amsterdam, 1948. 5 p. (Mimeographed.)

* Compiled by P. Medgyessy. This list does not contain book reviews, prefaces written to the first numbers of new periodicals, abstracts of lectures delivered on Hungarian conferences and colloquiums, articles and comments without mathematical aspects, etc published in periodicals.

7. On the zeros of the L-function of Dirichlet. Math. Centrum, Amsterdam, 1948. 4 p. (Mimeographed.)
8. Proof of the theorem that every integer can be represented as the sum of a prime and an almost prime. Math. Centrum, Amsterdam, 1948. 3 p. (Mimeographed.)

1949

1. О представлении чисел 1, 2, . . ., *N* посредством разностей (с Ласло Редеи). Мат. Сборник 24 (1949) 385—389.
2. Some remarks on independent random variables. Hung. Acta Math. 1 (1946/49) No. 4, 17—20.
3. On the measure of equidistribution of point sets. Acta Sci. Math. Szeged 13 (1949) 77—92.
4. Un nouveau théorème concernant les fonctions indépendantes et ses applications à la théorie des nombres. Jour. Math. Pures Appl. 28 (1949) 137—149.
5. A szovjet matematika 30 éve. Természet és Technika 108 (1949) 220—226.
6. Probability methods in number theory. Publ. Math. Coll. Budapest 1 (1949) No. 21, 1—9.
7. Sur un théorème général de probabilité. Annales Inst. Fourier 1 (1949) 43—52.
8. On the coefficients of schlicht functions. Publ. Math. Debrecen 1 (1949) 18—23.
9. A szovjet matematika 30 éve. I. A valószínűségszámítás megalapozásáról. Mat. Lapok 1 (1949/50) 27—64.

1950

1. On a theorem of Erdős and Turán. Proc. Amer. Math. Soc. 1 (1950) 7—10.
2. Some problems and results on consecutive primes (with P. Erdős). „Simon Stevin" 27 (1949/50) 115—125.
3. A szovjet matematika 30 éve. II. A valószínűségszámítás új irányai. Mat. Lapok 1 (1949/50) 91—137.
4. On the large sieve of Ju. V. Linnik. Comp. Math. 8 (1950) 68—75.
5. On the geometry of conformal mapping. Acta Sci. Math. Szeged 12 (1950) Pars B, 215—222.
6. On the algebra of distributions. Publ. Math. Debrecen 1 (1950) 135—149.
7. Az aprítás matematikai elméletéről. Építőanyag 2 (1950) 9—10. szám. 7 p.
8. A Newton-féle gyökközelítő eljárásról. Mat. Lapok 1 (1949/50) 278—293.
9. On the summability of Cauchy—Fourier series. Publ. Math. Debrecen 1 (1950) 162—164. (cf. 1945/1.)
10. Об одной общей теореме теории вероятностей и о ее применении в теории чисел. Zprávy o společnem 3. sjezdu matematiků Československých a 7. sjezdu matematiku Polských, Praha, 1950. Časopis Pěst. Mat. Fys. 74 (1949) 167—175.
11. К теории предельных теорем для сумм независимых случайных величин. Acta Math. Acad. Sci. Hung. 1 (1950) 99—108.
12. Harc a formalizmus ellen a matematika tanításában. A középiskolai matematikatanítás kérdései. Szocialista Nevelés Kiskönyvtára 4. sz. Közoktatásügyi Kiadó Vállalat, Budapest, 1950; pp. 24—28.
13. Remarks concerning the zeros of certain integral functions. C. R. Acad. Bulg. Sci. 3 (1950) No. 2—3, 9—10.
14. On composed Poisson distributions, I. (with L. Jánossy and J. Aczél). Acta Math. Acad. Sci. Hung. 1 (1950) 209—224.

15. Valószínűségszámítás. 1949—50. I. f. é.—II. f. é. Egyetemi jegyzet. A Debreceni Tudományegyetem Matematikai Intézete, Debrecen, 1950.
16. Valószínűségszámítás. Egyetemi jegyzet. Az Eötvös Loránd Tudományegyetem Természettudományi Kara, Budapest, 1950.

1951

1. A valószínűségszámítás központi határértéktételének egy új általánosításáról. MTA III. Oszt. Közl. 1 (1951) 351—355. (cf.: 1950/11).
2. A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének feladatairól. Akad. Ért. 58 (1951) 483. füzet, 1951 január—február, 20—26.
3. A Poisson-eloszlás problémaköréről. MTA III. Oszt. Közl. 1 (1951) 202—212.
4. On some problems concerning Poisson processes. Publ. Math. Debrecen 2 (1951) 66—73.
5. Sur l'indépendence des domaines simples dans l'espace euclidien à n dimensions (with C. Rényi and J. Surányi). Colloqu. Math. 2 (1951) 130—135.
6. Összetett Poisson-eloszlásokról, I. (Jánossy Lajossal és Aczél Jánossal). MTA III. Oszt. Közl. 1 (1951) 315—238. (cf.: 1950/14).
7. Об основах теории вероятностей. Годишник Физ.-мат. Факултет София, 47 (1951) Книга 1, 227—236.
8. Základy theorie pravděpodobnosti. Mat. Ustav Československe Akad. Věd., Praha, 1951. 10 p. (Mimeographed.) (cf.: 1951/7.)
9. On composed Poisson distributions. II. Acta Math. Acad. Sci. Hung. 2 (1951) 83—98.
10. Összetett Poisson eloszlásokról II. MTA III. Oszt. Közl. 1 (1951) 329—341. (cf.: 1951/9).
11. Két bizonyítás Jánossy Lajos egy tételére. (Turán Pállal.) MTA III. Oszt. Közl. 1 (1951) 369—370.
12. Levél a szerkesztőhöz. Hőmunkások víz- és sóanyagcseréje. — A matematikai statisztika módszereinek alkalmazása az orvostudományban. (A következőhöz: Somfai Jenő és Nógrády György: A munkaklíma hatásának vizsgálata bányászokon. Orvosi Hetilap 91 (1950) 871—875.) Orvosi Hetilap és Szovjet Orvostudományi Beszámoló 92 (1951) 945—947.
13. On the approximation of measurable functions (with L. Pukánszky). Publ. Math. Debrecen 2 (1951) 146—149.
14. Komplex függvénytan. Egyetemi jegyzet. Tankönyvkiadó, 1. Jegyzetsokszorosító, Budapest, 1951. 34 p.
15. Komplex függvénytan. Egyetemi jegyzet. VKM 2. Jegyzetsokszorosító, Budapest, 1951. 59 p.

1952

1. Sztochasztikus függetlenség és teljes függvényrendszerek. Az Első Magyar Matematikai Kongresszus Közleményei. 1950 augusztus 27.—szeptember 2. Akadémiai Kiadó, Budapest, 1952; pp. 299—308.
2. Стохастическая независимость и полные системы функций. Az Első Magyar Matematikai Kongresszus Közleményei. 1950. augusztus 27.—szeptember 2. Akadémiai Kiadó, Budapest, 1952; pp. 309—316. (cf.: 1952/1).
3. On a conjecture of H. Steinhaus. Annales Soc. Polon. Math. 25 (1952) 279—287.

4. Hozzászólás. (A következőhöz: Kalmár László: A matematika alapjaival kapcsolatos újabb eredmények. MTA III. Oszt. Közl. 2 (1952) 89—103.) MTA III. Oszt. Közl. 2 (1952) 104—107.
5. Új eredmények a valószínűségszámítás terén. MTA III. Oszt. Közl. 2 (1952) 125—139.
6. A. Ja. Hincsin „A statisztikai mechanika analitikus módszerei" c. könyvéről. (Fényes Imrével.) MTA III. Oszt. Közl. 2 (1952) 275—280.
7. A valószínűségszámítás elvi kérdései a dialektikus materializmus megvilágításában. Filozófiai Évkönyv. 1952. Akadémiai Kiadó, Budapest, 1952; pp. 63—97.
8. On projections of probability distributions. Acta Math. Acad. Sci. Hung. 3 (1952) 131—142.
9. Jordan Károly matematikai munkásságáról. Mat. Lapok 3 (1952) 111—121.
10. Gépalkatrészek és felszerelési tárgyak törzskészletének valószínűségszámítási meghatározása. (Szentmártony Tiborral.) Mat. Lapok. 3 (1952) 129—139.
11. Gépipari üzemek elektromos energiaszükségletének és egyidejűségi, illetőleg szükségleti tényezőjének valószínűségszámítási meghatározása. (Szentmártony Tiborral.) MTA Alk. Mat. Int. Közl. 1 (1952) 85—104.
12. Kompresszorok és légtartályok racionális méretezése üzemek sűrített levegővel való ellátására. MTA Alk. Mat. Int. Közl. 1 (1952) 105—138.
13. Poisson-folyamatok által származtatott történés-folyamatokról és azok technikai és fizikai alkalmazásairól. (Takács Lajossal.) MTA Alk. Mat. Int. Közl. 1 (1952) 139—146.
14. Megjegyzések Gombás Pál és Gáspár Rezső egy dolgozatához. MTA Alk. Mat. Int. Közl. 1 (1952) 393—397.
15. On the zeros of polynomials (with P. Turán). Acta Math. Acad. Sci. Hung. 3 (1952) 275—284.
16. Bolyai János, a tudomány nagy forradalmára. Mat. Lapok 3 (1952) 173—178.
17. Valószínűségszámítás. Egyetemi jegyzet. Felsőoktatási Jegyzetellátó Vállalat, Budapest, 1952.

1953

1. H. Steinhaus egy sejtéséről. MTA III. Oszt. Közl. 3 (1953) 37—44. (cf.: 1952/3).
2. Укрепление связи математики с практикой. Природа, 1953, 69—73.
3. Poznámka u uhlech mnohoúhelnika. Časopis Pěst. Mat. 78 (1953) 305—306.
4. Valószínűség-eloszlások vetületeiről. MTA III. Oszt. Közl. 3 (1953) 59—69. (cf.: 1952/8).
5. Bolyai János felfedezésének tudományos és világnézeti jelentősége. Természet és Technika 112 (1953) 1—4.
6. A Bolyai—Lobacsevszkij geometria világnézeti jelentősége. MTA III. Oszt. Közl. 3 (1953) 253—273.
7. Ideologický význam geometrie Bolyai—Lobačevského. Časopis Pěst. Mat. 78 (1953) 149—168. (cf.: 1953/6.)
8. Hozzászólás. (A következőhöz: Jánossy Lajos: Beszámoló a berlini fizikus kongresszus egyes problémáiról. MTA III. Oszt. Közl. 3 (1953) 323—325.) MTA III. Oszt. Közl. 3 (1953) 326—327.
9. Hozzászólás. (A következőhöz: Gombás Pál: Elméleti fizikai kutatásokban alkalmazott matematikai módszerek különös tekintettel a kvantummechanikai

közelítő módszerekre. MTA III. Oszt. Közl. 3 (1953) 329—340.) MTA III. Oszt. Közl. 3 (1953) 344—347.

10. Az Alkalmazott Matematikai Intézet munkája a valószínűségszámítás ipari alkalmazásai terén. MTA III. Oszt. Közl. 3 (1953) 363—372.
11. On the theory of order statistics. Acta Math. Acad. Sci. Hung. 4 (1953) 191—231.
12. A rendezett minták elméletéről. MTA III. Oszt. Közl. 3 (1953) 467—503. (cf.: 1953/11).
13. Eine neue Methode in der Theorie der geordneten Stichproben. Bericht über die Mathematiker-Tagung in Berlin, Januar 1953. Deutscher Verlag der Wissenschaften, Berlin, 1953; pp. 203—212.
14. Hozzászólás. (A következőhöz: Ankét O. J. Smidt „Négy előadás a Föld keletkezésének elméletéről" című könyvéről. MTA III. Oszt. Közl. 3 (1953) 579—601.) MTA III. Oszt. Közl. 3 (1953) 595—600.
15. Kémiai reakciók tárgyalása a sztochasztikus folyamatok elmélete segítségével. MTA Alk. Mat. Int. Közl. 2 (1953) 83—101.
16. Újabb kritériumok két minta összehasonlítására. MTA Alk. Mat. Int. Közl. 2 (1953) 243—265.
17. A valószínűségszámítás alapfogalmairól. Mérnöki Továbbképző Intézet előadássorozatából. Felsőoktatási Jegyzetellátó Vállalat, Budapest, 1953. 51 p.
18. A raktárkészlet pótlásáról I. (Palásti Ilonával, Szentmártony Tiborral és Takács Lajossal.) MTA Alk. Mat. Int. Közl. 2 (1953) 187—201.
19. Valószínűségszámítás. Egyetemi jegyzet. Felsőoktatási Jegyzetellátó Vállalat, Budapest, 1953.
20. Játék a véletlennel. Pedagógiai Főiskolai jegyzet. 1. változatlan utánnyomás. Felsőoktatási Jegyzetellátó Vállalat, Budapest, 1953. 9 p.

1954

1. Základni problémy počtu pravděpodobnosti ve světle dialektického materialismu. Časopis Pěst. Mat. 79 (1954) 189—218. (cf.: 1952/7.)
2. Elementary proofs of some basic facts concerning order statistics (with Gy. Hajós). Acta Math. Acad. Sci. Hung. 5 (1954) 1—6.
3. Идеологическое значение геометрии Бояи—Лобачевского. Acta Math. Acad. Sci. Hung. 5 (1954) Supplementum, 21—42. (cf.: 1953/6).
4. A valószínűségszámítás új axiomatikus felépítése. MTA III. Oszt. Közl. 4 (1954) 369—427.
5. A valószínűségszámítás történetének rövid áttekintése. MTA III. Oszt. Közl. 4 (1954) 447—466.
6. Valószínűségszámítás. Tankönyvkiadó, Budapest, 1954.
7. Elemi bizonyítások a rendezett minták elméletének néhány alapvető összefüggésére. (Hajós Györggyel.) MTA III. Oszt. Közl. 4 (1954) 467—472. (cf.: 1953/2).
8. A kémiai frakcionáló megosztás matematikai tárgyalása nem-teljes diffúzió esetében. (Medgyessy Pállal, Tettamanti Károllyal és Vincze Istvánnal.) MTA Alk. Mat. Int. Közl. 3 (1954) 81—97.
9. A komplex potenciál egyrétűségéről, I. (Rényi Katóval.) MTA Alk. Mat. Int. Közl. 3 (1954) 353—367.

10. Die prinzipiellen Fragen der Wahrscheinlichkeitsrechnung im Lichte des dialektischen Materialismus. Philosophisches Jahrbuch, 1952. Zusammenfassung. Akadémiai Kiadó, Budapest, 1954; pp. 7—8.
11. Hozzászólás. (A következőhöz: Sedlmayer Kurt: Nagyobb termések elérésének tudományos alapjai. MTA IV. Oszt. Közl. 5 (1954) 187—197.) MTA IV. Oszt. Közl. 5 (1954) 198—200.
12. Egy lucerna nemesítésével kapcsolatos kombinatorikai problémáról. (Előadás: Matematikai Statisztikai Kollokvium. 1954. szeptember hó 27.—29., Jósvafő.) Kivonat: Az 1954. szeptember hó 27.-étől 29.-ig Jósvafőn, a Bolyai János Matematikai Társulat által rendezett Matematikai Statisztikai Kollokviumon elhangzott előadások kivonatai. Bolyai János Matematikai Társulat, Budapest, 1954; pp. 13—15.
13. Megoldatlan problémák a rendezett minták elméletében. — Referátum. (Előadás: Matematikai Statisztikai Kollokvium. 1954. szeptember hó 27.—29., Jósvafő.) Kivonat: Az 1954. szeptember hó 27.-étől 29.-ig Jósvafőn, a Bolyai János Matematikai Társulat által rendezett Matematikai Statisztikai Kollokviumon elhangzott előadások kivonatai. Bolyai János Matematikai Társulat, Budapest, 1954; pp. 18—20.

1955

1. Egy kombinatorikai probléma, amely a lucerna nemesítésével kapcsolatban merült fel. Mat. Lapok 6 (1955) 151—164.
2. Bizonyos trigonometrikus rendszerek teljességéről. (Czipszer Jánossal.) MTA III. Oszt. Közl. 5 (1955) 391—410.
3. A matematika fejlődése hazánkban a felszabadulás óta. (Alexits Györggyel és Hajós Györggyel.) A magyar tudomány 10 éve. 1945—1955. Akadémiai Kiadó, Budapest, 1955; pp. 87—106.
4. Generalization of an inequality of Kolmogorov (with J. Hájek). Acta Math. Acad. Sci. Hung. 6 (1955) 281—283.
5. On a new axiomatic theory of probability. Acta Math. Acad. Sci. Hung. 6 (1955) 285—335.
6. A sztochasztikus folyamatok elméletéről és annak néhány műszaki alkalmazásáról. (Mérnöki Továbbképző Intézet előadássorozatából.) Felsőoktatási Jegyzetellátó Vállalat, Budapest, 1955. 78 p.
7. On the density of certain sequences of integers. Publ. Inst. Math. Acad. Serbe Sci., Beograd 8 (1955) 157—162.
8. Matematikai statisztika. Egyetemi jegyzet. Jegyzetsokszorosító, Budapest, 1955.
9. A világ tudósainak tapasztalatcseréje egyaránt hasznos a béke és a tudomány számára. Beszámoló egy külföldi tanulmányút élményeiről és tanulságairól. Szabad Nép, 1955, december 18, 6.
10. Pour la coopération entre savants du monde. Journal article, 1955. 2 p. (cf.: 1955/9.).
11. A matematikai módszerek alkalmazásának eredményei és lehetőségei. Szabad Nép, 1955, szeptember 8, 4.

1956

1. Szakkörökben elvégezhető valószínűségszámítási kísérletekről. — Előadások az iskolai matematika köréből. A Bolyai János Matematikai Társulat kiadványa. Tankönyvkiadó, Budapest, 1956; pp. 135—150.

2. Axiomatischer Aufbau der Wahrscheinlichkeitsrechnung. Bericht über die Tagung Wahrscheinlicheitsrechnung und mathematische Statistik in Berlin, Oktober, 1954. Deutscher Verlag der Wissenschaften, Berlin 1956; pp. 7—15.
3. An inequality for uncorrelated random variables (with E. Zergényi). Czech. Math. Jour. 6 (81) (1956) 415—419.
4. A számjegyek eloszlása valós számok Cantor-féle előállításaiban. Mat. Lapok 7 (1956) 77—100.
5. On some combinatorical problems (with P. Erdős). Publ. Math. Debrecen 4 (1955/56) 398—405.
6. О предельном распределении для сумм независимых случайных величин на бикомпактных коммутативных топологических группах (с А. Прекопа и К. Урбаником). Acta Math. Acad. Sci. Hung. 7 (1956) 11—16.
7. On conditional probability spaces generated by a dimensionally ordered set of measures. Теория Вероятностей 1 (1956) 61—71.
8. Az entrópia fogalmáról. (Balatoni Jánossal.) MTA Mat. Kut. Int. Közl. 1 (1956) 9—40.
9. Az ingerületátvitel valószínűsége egy egyszerű konvergens kapcsolású interneuronális synapsis-modellben. (Szentágothay Jánossal.) MTA Mat. Kut. Int. Közl. 1 (1956) 83—91.
10. On the number of zeros of successive derivatives of analytic functions. (with P. Erdős.) Acta Math. Acad. Sci. Hung. 7 (1956) 125—144.
11. Az árrendezés problémájáról. (Bródy Andrással.) MTA Mat. Kut. Int. Közl. 1 (1956) 325—335.
12. A Monte-Carlo módszer mint minimax stratégia. (Palásti Ilonával.) MTA Mat. Kut. Int. Közl. 1 (1956) 529—545.
13. Discussion on Dr. David's and Dr. Johnson's paper. (F. N. David and N. L. Johnson: Some tests of significance with ordered variables. Jour. Roy. Stat. Soc. Ser. B, 18 (1956) 1—20.) Jour. Roy. Soc. Ser. B, 18 (1956) 29.
14. A Poisson-folyamat egy jellemzése. MTA Mat. Kut. Int. Közl. 1 (1956) 519—527.
15. On the independence in the limit of sums depending on the same sequence of independent random variables (with A. Prékopa). Acta Math. Acad. Sci. Hung. 7 (1956) 319—326.
16. Internationaler Erfahrungsaustausch der Wissenschaftler erhält den Frieden und fördert die Wissenschaft. Journal article, 1956. 2 p. (cf.: 1955/9.).
17. "Scientific exchange is beneficial to world peace and science". Journal article, 1956. 2 p. (cf.: 1955/9.).

1957

1. On a new axiomatic foundation of the theory of probability. Proceedings of the International Congress of Mathematicians 1954. Amsterdam September 2—September 9. Vol. I. Noordhoff N. V., Groningen — North-Holland Publishing Co., Amsterdam, 1957; pp. 506—507.
2. On the theory of order statistics. Proceedings of the International Congress of Mathematicians 1954. Amsterdam September 2—September 9. Vol. I. Noordhoff N. V., Groningen—North-Holland Publishing Co., Amsterdam, 1957; pp. 508—509.

3. A new deduction of Maxwell's law of velocity distribution. Изв. Мат. Инст. София 2 (1957) Книга 2, 45—55.
4. Probabilistic proof of a theorem on the approximation of continuous functions by means of generalized Bernstein polynomials (with M. Arató). Acta Math. Acad. Sci. Hung. 8 (1957) 91—98.
5. On the asymptotic distribution of the sum of a random number of independent random variables. Acta Math. Acad. Sci. Hung. 8 (1957) 193—199.
6. On the number of zeros of successive derivatives of entire functions of finite order (with P. Erdős). Acta Math. Acad. Sci. Hung. 8 (1957) 223—225.
7. A probabilistic approach to problems of diophantine approximation (with P. Erdős). Illinois Jour. Math. 1 (1957) 303—315.
8. Mathematical Notes. II. On the sequence of generalized partial sums of a series. Publ. Math. Debrecen 5 (1957/58) 129—141.
9. A remark on the theorem of Simmons. Acta Sci. Math. Szeged 18 (1957) 21—22.
10. Valós számok előállítására szolgáló algoritmusokról. MTA III. Oszt. Közl. 7 (1957) 265—293.
11. Representations for real numbers and their ergodic properties. Acta Math. Acad. Sci. Hung. 8 (1957) 477—493.
12. Az $L(z)$ valószínűség-eloszlásfüggvényről. MTA Mat. Kut. Int. Közl. 2 (1957) 43—50.
13. Szénszemcsés ellenállások vizsgálata valószínűségszámítási módszerrel. MTA Mat. Kut. Int. Közl. 2 (1957) 247—256.
14. Über den Begriff der Entropie in der Wahrscheinlichkeitsrechnung. (Előadás: IV. Österreichischer Mathematikerkongress Wien, 17.—22. IX. 1956.) Kivonat: Nachr. Österr. Math. Ges. Beilage zu "Internat. Math. Nachr." 11 (1957) April, Nr. 47/48., Sondernummer, 83.
15. Über den Begriff der Entropie. (Balatoni Jánossal.) Arbeiten zur Informationstheorie. I. Deutscher Verlag der Wissenschaften, Berlin, 1957; pp. 117—134. (cf.: 1956/8.)

1958

1. On a theorem of Erdős—Kac (with P. Turán). Acta Arith. 4 (1958) 71—84.
2. Some remarks on univalent functions. Изв. Мат. Инст. София 3 (1959) Книга 2, 111—121.
3. Some remarks on univalent functions II. Ann. Acad. Sci. Fennicae Series A. I. Mathematica. 250/29. Suomalainen Tiedeakatemia, Helsinki, 1958, 7 p.
4. Quelques remarques sur les probabilités d'événements dépendants. Jour. Math. Pures Appl. (9) 37 (1958) 393—398.
5. On mixing sequences of sets. Acta Math. Acad. Sci. Hung. 9 (1958) 215—228.
6. Egy egydimenziós véletlen térkitöltési problémáról. MTA Mat. Kut. Int. Közl. 3 (1958) 109—127.
7. On Engel's and Sylvester's series (with P. Erdős and P. Szüsz). Annales Univ. Sci. Budapest. Sect. Math. 1 (1958) 7—32.
8. Probability methods in number theory (in Chinese). Shuxue Jinzhan 4 (1958) 465—510.
9. On Cantor's products. Colloqu. Math. 6 (1958) 135—139.
10. On mixing sequences of random variables (with P. Révész). Acta Math. Acad. Sci. Hung. 9 (1958) 389—393.

11. On singular radii of power series (with P. Erdős). MTA Mat. Kut. Int. Közl. 3 (1958) 159—169.
12. On the probabilistic generalization of the large sieve of Linnik. MTA Mat. Kut. Int. Közl. 3 (1958) 199—206.
13. Matematikai statisztika IV. éves alkalmazott matematika szakos hallgatók számára. Felsőoktatási Jegyzetellátó Vállalat, Budapest, 1958. 211. p.
14. Levél a szerkesztőhöz. (A következőhöz: Pólya György, A gondolkodás iskolája c. könyvéről (Vámosi Pál) (Könyvbarát) VII. évf. 5. szám 46. o.) Könyvbarát, 8 (1958) 1. szám, 30.

1959

1. Some further statistical properties of the digits in Cantor's series (with P. Erdős). Acta Math. Acad. Sci. Hung. 10 (1959) 21—29.
2. On random graphs I. (with P. Erdős). Publ. Math. Debrecen 6 (1959) 290—297.
3. On a theorem of P. Erdős and its application in information theory. Mathematica, Cluj 1 (24) (1959) 341—344.
4. On the dimension and entropy of probability distributions. Acta Math. Acad. Sci. Hung. 10 (1959) 193—215.
5. New version of the probabilistic generalization of the large sieve. Acta Math. Acad. Sci. Hung. 10 (1959) 217—226.
6. On the central limit theorem for samples from a finite population. (With P. Erdős.) MTA Mat. Kut. Int. Közl. 4 (1959) 49—61.
7. Some remarks on the theory of trees. MTA Mat. Kut. Int. Közl. 4 (1959) 73—85.
8. On Cantor's series with convergent $\sum \frac{1}{q^n}$ (with P. Erdős). Annales Univ. Sci. Budapest, Sect. Math. 2 (1959) 93—109.
9. Autoklávok soros és párhuzamos kapcsolásáról és a keverés elméletéről. MTA Mat. Kut. Int. Közl. 4 (1959) 155—165.
10. On measures of dependence. Acta Math. Acad. Sci. Hung. 10 (1959) 441—451.
11. On connected graphs, I. MTA Mat. Kut. Int. Közl. 4 (1959) 385—388.
12. Summation methods and probability theory. MTA Mat. Kut. Int. Közl. 4 (1959) 389—399.
13. Dialógus a matematika tanításáról. (Előadás: Középiskolai Szakfelügyelői Tanácskozás a Központi Pedagógus Továbbképző Intézetben. Budapest, 1959 január.) Az MTA matematikai Kutatóintézete, Budapest, 1959. 17 p. (Mimeographed.)
14. Sztochasztikus kapcsolatok mérőszámairól. (On measures of correlation.) (Lecture: Biometriai Symposion. Budapest, 1959. szeptember 7—9.) Abstract: Biometriai Symposion. Budapest, 1959. szeptember 7—9. Előadáskivonatok. Budapest, 1959. 1 p.

1960

1. On the central limit theorem for the sum of a random number of independent random variables. Acta Math. Acad. Sci. Hung. 11 (1960) 97—102.
2. Additive properties of random sequences of positive integers (with P. Erdős). Acta Arithm. 6 (1960) 83—110.
3. Bolyongási problémákra vonatkozó határeloszlástételek. MTA III. Oszt. Közl. 10 (1960) 149—169.

4. On the evolution of random graphs (with P. Erdős). MTA Mat. Kut. Int. Közl. 5 (1960) 17—61.
5. Probabilistic methods in number theory. Proceedings of the International Congress of Mathematicians 14—21 August 1958. (Edinburgh.) Cambridge U. P., London, 1960; pp. 529—539.
6. Az információelmélet néhány alapvető kérdése. MTA III. Oszt. Közl. 10 (1960) 251—282.
7. Dimension, entropy and information. Transactions of the Prague Conference on information theory, statistical decision functions, random processes held at Liblice near Prague, from June 1 to 6, 1959. Publ. House of the Czech. Acad. Sci., Prague, 1960; pp. 545—556.
8. Bemerkungen zur Arbeit „Über gewisse Elementenfolgen des Hilbertschen Raumes" von K. Koncz. MTA Mat. Kut. Int. Közl. 5 (1960) 265—267.
9. Üzletek áruellátásával kapcsolatos szélsőértékfeladatok. (Ziermann Margittal.) MTA Mat. Kut. Int. Közl. 5. B (1960) 495—506.
10. Turán Pál matematikai munkásságáról. Mat. Lapok 11 (1960) 229—263.
11. On measures of entropy and information. Fourth Berkeley Symposium on Mathematical Statistics and Probability. Held at the Statistical Laboratory, University of California, June 20—July 30. 1960. 28 p. (Mimeographed.)
12. On the evolution of random graph (with P. Erdős). Random Graphs, 3 seminarer holdt af A. Rényi. Februar—marts 1960. Matematisk Institut, Aarhus Universitet, Aarhus, 1960. 57 p. (Mimeographed.)
13. On the evolution of random graphs (with P. Erdős). 32 nd Session of the International Statistical Institute, Tokyo, 1960. Tokyo, 1960. 5 p. (Mimeographed.)
15. Jordan Károly 1871—1959. Magyar Tudomány 5 (1960) 233—235.
15. Boszorkányság-e a matematika? Népszabadság, 1960, augusztus 25, 8.

1961

1. On the evolution of random graphs (with P. Erdős). Bull. Inst. Internat. Stat. 38 (1961) 4 e Livraison, 343—347. (cf.: 1960/13.)
2. On measures of entropy and information. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability. Held at the Statistical Laboratory, University of California, June 20—July 30. 1960. Vol. 1. University of California Press, Berkeley—Los Angeles, 1961; pp. 547—561. (cf. 1960/11.)
3. Egy általános módszer valószínűségszámítási tételek bizonyítására és annak néhány alkalmazása. MTA III. Oszt. Közl. 11 (1961) 79—105.
4. On random generating elements of a finite Boolean algebra. Acta Sci. Math. Szeged 22 (1961), 75—81.
5. On the strength of connectedness of a random graph (with P. Erdős). Acta. Math Acad. Sci. Hung. 12 (1961) 261—267.
6. On a classical problem of probability theory (with P. Erdős). MTA Mat. Kut. Int. Közl. 6. A (1961) 215—220.
7. On Kolmogorov's inequality. MTA Mat. Kut. Int. Közl. 6. A (1961) 411—415.
8. Legendre polynomials and probability theory. Annales Univ. Sci. Budapest, Sect. Math. 3—4 (1960/61) 247—251.
9. Egy információelméleti problémáról. MTA Mat. Kut. Int. Közl. 6. B. (1961) 505—516.

10. On random subsets of a finite set. Mathematica, Cluj 3 (26) (1961) 355—362.
11. Über verschiedene Masszahlen von Entropie und Informationsgewinn. (Lecture: V. Österreichische Mathematikerkongress, Innsbruck, 12.—17. IX. 1960.) Abstract: Nachr. Österr. Math. Ges. Beilage zu „Internat. Math. Nachr." 15 (1961) Jänner, Nr. 66, Sondernummer, 79—80.
12. On different measures of information. (Lecture: Second Hungarian Mathematical Congress, Budapest, August 24.—31. 1960.) Abstract: Deuxième Congrès Mathématique Hongrois Budapest, 24.—31., August 1960. II. Akadémiai Kiadó, Budapest, 1961; Section IV, pp. 26—28.
13. Gondolatak a matematikusképzés továbbfejlesztéséről. Magyar Tudomány 6 (1961) 593—600.
14. Statistical laws of accumulation of information. 33rd Session of the International Statistical Institute, Paris, 1961. Paris, 1961. (Mimeographed.)
15. Matematikai kongresszusok és a II. Magyar Matematikai Kongresszus. Magyar Tudomány 6 (1961) 13—23.
16. Véletlen információ akkumulációja. (Német nyelven.) (In German.) (Lecture: Kollokvium über Wahrscheinlichkeitsrechnung und Statistik, Eisenstadt, 15.—18. XI. 1961. (Manuscript.)
17. Statistical laws of accumulation of information. Michigan State University, Department of Statistics, East Lansing, 1961. 11 p. (Mimeographed.) (cf.: 1961/14).

1962

1. Statistical laws of accumulation of information. Bull. Inst. Internat. Stat. 39 (1962) 2 e Livraison, 311—316. (cf.: 1961/14).
2. Az információ-akkumuláció statisztikus törvényszerűségeiről. MTA Mat. Kut. Int. Közl. 12 (1962) 15—33.
3. Egy megfigyeléssorozat kiemelkedő elemeiről. MTA III. Oszt. Közl. 12 (1962) 105—121.
4. Three new proofs and a generalization of a theorem of Irving Weiss. MTA Mat. Kut. Int. Közl. 7. A (1962) 203—214.
5. Théorie des éléments saillants d'une suite d'observations. Annales Fac. Sci. Univ. Clermont-Ferrand 2 (1962) No. 8, 7—12.
6. Dialógus a matematikáról. Az MTA Matematikai Kutatóintézete, Budapest 1962. 25 p. (Mimeographed.)
7. On a problem of A. Zygmund (with P. Erdős). Studies in mathematical analysis and related topics. Essays in honor of George Pólya. Stanford Univ. Press, Stanford, Cal., 1962; pp. 110—116.
8. A new approach to the theory of Engel's series. Annales Univ. Budapest. Sect. Math. 5 (1962) 25—32.
9. On the representation of an even number as the sum of a prime and of an almost prime. American Mathematical Society. Translations. Series 2, Vol. 19. American Mathematical Society, Providence, 1962; pp. 299—321. (cf.: 1948/1).
10. Egy gráfelméleti problémáról. (Erdős Pállal.) MTA Mat. Kut. Int. Közl. 7. B (1962) 623—641.
11. A matematika alkalmazásairól tartandó vita tézisei. Magyar Tudomány 7 (69) 553—559.

12. On the theory of outstanding observations. (Lecture: International Congress of Mathematicians, Stockholm 1962.) Abstract: International Congress of Mathematicians, Abstracts of short communications. Stockholm, 1962. Almqvist and Wiksells, Uppsala, 1962; pp. 165—166.
13. Théorie des éléments saillants d'une suite d'observations. Colloquium on Combinatorial Methods in Probability Theory. August 1—10, 1962. Matematisk Institut, Aarhus Universitet, Danmark, Aarhus, 1962; pp. 104—117.
14. Wahrscheinlichkeitsrechnung, mit einem Anhang über Informationstheorie. VEB Deutscher Verlag der Wissenschaften, Berlin, 1962.
15. Sur les graphes aléatoires (I). L'évolution des graphes aléatoires. — Sur les graphes aléatoires II. Symetrie et asymetrie des graphes aléatoires. Institut H. Poincaré, Paris, 1962. 20 p. (Mimeographed.)
16. Dialógus a matematika tanításáról. — Előadások a középiskolai matematika köréből. A Központi Pedagógus Továbbképző Intézet és a Bolyai János Matematikai Társulat kiadványa. Tankönyvkiadó, Budapest, 1962; pp. 5—19. (cf.: 1959/13.)
17. Dialógus a matematikáról. Valóság 5 (1962) 3. szám, 40—56.
18. A matematika és a társadalom. Népszabadság, 1962, október 21, 7.

1963

1. Remarks on a problem of Obreanu (with P. Erdős). Canadian Math. Bull. 6 (1963) 267—273.
2. Über die konvexe Hülle von n zufällig gewählten Punkten (mit R. Sulanke). Zeitschr. Wahrscheinlichkeitstheorie 2 (1963/64) 75—84.
3. On stable sequences of events. Sankhyā, Ser. A, 25 (1963) 293—302.
4. On the distribution of values of additive number-theoretical functions. Publ. Math. Debrecen 10 (1963) 264—273.
5. A study of sequences of equivalent events as special stable sequences (with P. Révész). Publ. Math. Debrecen 10 (1963) 319—325.
6. On "small" coefficients of the power series of an entire function (with C. Rényi). Annales Univ. Budapest, Sect. Math. 6 (1963) 27—38.
7. On two problems of information theory (with P. Erdős). MTA Mat. Kut. Int. Közl. 8. A (1963) 229—243.
8. On random matrices (with P. Erdős). MTA Mat. Kut. Int. Közl. 8. A (1963) 455—461.
9. An elementary inequality between the probability of events (with P. Erdős and J. Neveu). Math. Scand. 13 (1963) 99—104.
10. Un dialogue. Les cahiers rationalistes, 33 (1963) janvier—février, Nos. 208—209, 4—32. (cf.: 1962/17.)
12. Blaise Pascal. 1623—1662. Magyar Tudomány 8 (70) (1963) 102—108.
13. Megjegyzések egyes „megjegyzések"-hez. (A következőhöz: Tekse Kálmán: Néhány megjegyzés Rényi A. „A matematika alkalmazásairól tartandó vita téziseihez" című cikkéhez. Magyar Tudomány 8 (70) (1963) 46—50.) Magyar Tudomány 8 (70) (1963) 419—429.
14. Über stabile Folgen von Ereignissen und Zufallsveränderlichen. (Vortrag: Tagung über Mathematische Statistik und Wahsrscheinlichkeitstheorie. Oberwolfach, 4—8. März 1963.) Auszug: Tagungsbericht. Mathematische Statistik und Wahr-

scheinlichkeitstheorie. Oberwolfach, 4—8. März 1963. Mathematisches Forschungsinstitut, Oberwohlfach, 1963. (Mimeographed.)

15. Dialógus a matematikáról. (Részletek.) A matematika tanítása. Szemelvénygyűjtemény. (Szerkesztette Varga Tamás.) Kézirat (255—504. o.). Tankönyvkiadó, Budapest, 1963; pp. 381—415. (cf.: 1962/17.)
16. Dialógus a matematika tanításáról. A matematika tanítása. Szemelvénygyűjtemény. (Szerkesztette Varga Tamás.) Kézirat (255—504. o.). Tankönyvkiadó, Budapest, 1963; pp. 365—380. (cf.: 1959/13.)
17. A kultúra egységéről, matematikus szemmel. Valóság 6 (1963) 3. szám 51—53.
18. A Socratic dialogue on mathematics. Az MTA Matematikai Kutatóintézete, Budapest, 1963. 24 p. (Mimeographed.)
19. On the foundations of information theory. 34th Session of the International Statistical Institute, Ottawa, 1963. (Mimeographed.)

1964

1. Über die konvexe Hülle von n zufällig gewählten Punkten. II (mit R. Sulanke) Zeitschr. Wahrscheinlichkeitstheorie 3 (1964/65) 138—147.
2. Információelmélet és nyelvészet. Általános nyelvészeti tanulmányok. II. A matematikai nyelvészet és a gépi fordítás kérdései. Akadémiai Kiadó, Budapest, 1964; pp. 245—251.
3. Additive and multiplicative number-theoretical functions. University of Michigan, Ann Arbor, 1964. 23 p. (Mimeographed.)
4. Dialógus a matematika alkalmazásairól. Az MTA Matematikai Kutató Intézete, Budapest, 1964, 20 p. (Mimeographed.)
5. A természet könyvének nyelve. Dialógus. 1964. Az MTA Matematikai Kutató Intézete, Budapest, 1964. (Mimeographed.)
6. On an extremal property of the Poisson process. Annals Inst. Stat. Math. Tokyo 16 (1964) 129—133.
7. A generalization of a theorem of E. Vincze (with R. G. Laha and E. Lukács). MTA Mat. Kut. Int. Közl. 9. A (1964) 237—239.
8. On two mathematical models of the traffic on a divided highway. Jour. Appl. Prob. 1 (1964) 311—320.
9. On the amount of information concerning an unknown parameter in a sequence of observations. MTA Mat. Kut. Int. Közl. 9. A (1964) 617—625.
10. Hervorragende Elemente von Beobachtungsreihen. (Lecture: Internationale Tagung über Mathematische Statistik und ihre Anwendungen, Berlin, von 4. bis 8. September 1962.) Abstract: Abh. Deutsch. Akad. Wissensch. Berlin, Klasse Math., Phys. Technik, 1964, No. 4, 101.
11. A Socratic dialogue on mathematics. Canadian Math. Bull. 7 (1964) 441—462. (cf.: 1962/17).
12. Mathematics. A Socratic dialogue. Physics Today 17 (1964) December, 24—36. (cf.: 1962/17).
13. A Socratic dialogue on mathematics. "Simon Stevin" 38 (1963/64) 125—144. (cf.: 1962/17).
14. On the amount of information in a frequency count. Bull. Inst. Internat. Stat. 41 (1964) 623—626.

15. Sur les espaces simples des Probabilités conditionnelles. Annales Inst. H. Poincaré, Nouvelle Série, Sect. B, 1 (1964) no. 1., 3—21.
16. A dialogue on the applications of mathematics. Ontario Mathematics Gazette 31 (1964) Number 2, 28—40. (cf.: 1964/4).
17. A dialogue on the applications of mathematics. University of Michigan, Ann Arbor, 1964. 15 p. (Mimeographed.) (cf.: 1964/4).
18. Discussion on Mr. Lewis's paper. (P. A. W. Lewis: A branching Poisson process model for the analysis of computer failure patterns. Jour. Roy. Stat. Soc., Series B, 26 (1964) 398—456.) Jour. Roy. Stat. Soc., Series B, 26 (1964) 445—446.
19. Eine Extremaleigenschaft des Poissonschen Prozesses. (Lecture: Tagung über Mathematische Statistik and Wahrscheinlichkeitstheorie, Oberwolfach, 2.—7. August 1964.) Abstract: Tagungsbericht. Mathematische Statistik und Wahrscheinlichkeitstheorie. Oberwolfach, 2.—7. August 1964. Mathematisches Forschungsinstitut, Oberwohlfach, 1964. (Mimeographed.)
20. Mellékvágány? (A következőhöz: Farkas Miklós: A magyar matematikusok közötti vitáról. Magyar Tudomány 8 (1963) 824—829.) Magyar Tudomány 9 (1964) 228—230.
21. Information and statistics. Symposium on mathematical statistics, Budapest, 1964. 5 p. (Mimeographed.)

1965

1. On the foundations of information theory. Rev. Inst. Internat. Stat. 33 (1965) 1—14. (cf.: 1963/19).
2. Probabilistic methods in group theory (with P. Erdős). Jour. Analyse Mathématique 14 (1965) 127—138.
3. Some remarks on periodic entire functions (with C. Rényi). Jour. Analyse Mathématique 14 (1965) 303—310.
4. Levelek a valószínűségről. 1.—4. Az MTA Matematikai Kutatóintézete, Budapest, 1965. 46 p. (Mimeographed.)
5. On some basic problems of statistics from the point of view of information theory. École d' Été de l' OTAN 1965 sur les Méthodes Combinatoires en Théorie de l' Information et du Codage (Royan: 26 Août — 8 Septembre 1965). Royan, 1965. 26 p. (Mimeographed.)
6. On certain representations of real numbers and on sequences of equivalent events. Acta Sci. Math. Szeged 26 (1965) 63—74.
7. Dialógusok a matematikáról. Akadémiai Kiadó, Budapest, 1965.
8. On the theory of random search. Bull. Amer. Math. Soc. 71 (1965) 809—828.
9. On the mean value of nonnegative multiplicative number-theoretical functions (with P. Erdős). Michigan Math. Jour. 12 (1965) 321—338.
10. A new proof of a theorem of Delange. Publ. Math. Debrecen 12 (1965) 323—329.
11. A természet könyvének nyelve. Fizikai Szemle 15 (1965) 129—138. (Cf.: 1964/5.)
12. The language of the Great Book of Nature. Az MTA Matematikai Kutatóintézete, Budapest, 1965. 39 p. (Mimeographed.) (Cf:. 1964/5.)
13. A Matemática — Um Diálogo Socrático. Gazeta de Matemática, 1965, No. 100, Julho—Dezembro, 59—71 (cf.: 1962/17).
14. On the amount of information concerning an unknown parameter in a sequence of observations. Lectures to the Fifth Summer Research Institute of the Austra-

lian Mathematical Society. Preprint. Australian Mathematical Society, 1965; pp. 1—13. (cf.: 1964/9.)

15. On the theory of random search. Lectures to the Fifth Summer Research Institute of the Australian Mathematical Society. Preprint. Australian Mathematical Society, 1965; pp. 14—16.
16. On an extremal property of the Poisson process. Lectures to the Fifth Summer Research Institute of the Australian Mathematical Society. Preprint. Australian Mathematical Scciety, 1965; 17—20.
17. Véges geometriák kombinatorikai alkalmazásai I. Az MTA Matematikai Kutatóintézete, Budapest, 1965. 67 p. (Mimeographed.)
18. Új módszerek és eredmények a kombinatorikus analízisben I. Az MTA Matematikai Kutatóintézete, Budapest, 1965. 89 p. (Mimeographed.)
19. On the amount of information in a frequency count. 35th Session of the International Statistical Institute, Beograd, 1965. Beograd, 1965, 8 p. (Mimeographed.)
20. On the aount of information in a frequency-count. Bull. Inst. Internat. Stat. 41 (1965) 623—626, (cf.: 1965/19).

1966

1. On a problem of graph theory (with P. Erdős and V. T. Sós). Studia Sci. Math. Hung. 1 (1966) 215—235.
2. Véges geometriák kombinatorikai alkalmazásai I. Mat. Lapok 17 (1966) 33—76. (cf.: 1965/17).
3. Probability theory. Lecture notes. Stanford University, Stanford, 1966, 110 p.
4. On the existence of a factor of degree one of a connected random graph (with P. Erdős). Acta Math. Acad. Sci. Hung. 17 (1966) 359—368.
5. Új módszerek és eredmények a kombinatorikus analízisben, I. MTA III. Oszt. Közl. 16 (1966) 77—105. (cf.: 1965/18).
6. On the amount of missing information and the Neyman—Pearson-lemma. Rsearch papers in statistics. Festschrift for J. Neyman. Wiley, London, 1966; pp. 281—288.
7. On the amount of information in a random variable concerning an event. Jour. Math. Sci. Delhi 1 (1966) 30—33.
8. Statistics based on information theory. European Meeting of Statisticians, London, 1966. 17 p. (Mimeographed.)
9. Mathematics. A Socratic dialogue. (Mimeographed form of: Alfréd Rényi: Mathematics. A Socratic dialogue. Physics Today 17 (1964) December, 24—36.) Mathematics Department of Ohio University, 1966, 16 p. (cf. 1962/17).
10. Sokratischer Dialog. Neue Sammlung 6 (1966) 284—304. (cf.: 1962/17).
11. A dialogue on the applications of mathematics. "Simon Stevin", 39 (1965/66) 3—17. (cf.: 1964/4).
12. Új módszerek és eredmények a kombinatorikus analízisben, II. MTA III. Oszt. Közl. 16 (1966) 159—177. (cf.: 1965/18).
13. Levelek a valószínűségről. Fizikai Szemle 16 (1966) 278—288. (cf.: 1965/4)
14. A Matematikai Kutató Intézet 10 éve. Magyar Tudomány 11 (1966) 81—91.
15. Valószínűségszámítás. Tankönyvkiadó, Budapest, 1966.
16. Wahrscheinlichkeitsrechnung, mit einem Anhang über Informationstheorie. 2., berichtigte Auflage. VEB Deutscher Verlag der Wissenschaften, Berlin, 1966. (cf.: 1962/14.)

17. Calcul des probabilités. Avec un appendice sur la théorie de l'information. Dunod, Paris, 1966. (cf.: 1962/14).
18. Dialógusok a matematikáról. 2. kiadás. Akadémiai Kiadó, Budapest, 1966. (cf.: 1965/7.)
19. Valószínűségszámítási módszerek az analízisben. Az MTA Matematikai Kutatóintézete, Budapest, 1966. 65 p. (Mimeographed.)
20. Eine Ungleichung zwischen die Irrtumswahrscheinlichkeit und die fehlende Information. (Lecture: Tagung über Mathematische Statistik und Wahrscheinlichkeitstheorie. Oberwohlfach, 17.—24. April 1966.) Abstract: Tagungsbericht. Mathematische Statistik und Wahrscheinlichkeitstheorie. Oberwolfach, 17.—24. April 1966. Mathematisches Forschungsinstitut, Oberwolfach, 1966. (Mimeographed.)
21. On the mathematical theory of trees. "Rouse Ball" lectures, Cambridge, 1966. (Manuscript.)
22. A Poisson-folyamatról. (Lecture: Matematikai Statisztikai Kollokvium, Debrecen, 1966 október 13—15.) Abstract: Matematikai Statisztikai Kollokvium. Debrecen, 1966. október 13—15. Előadás kivonatok. Debrecen, 1966; p. 4.

1967

1. Levelek a valószínűségről. Akadémiai Kiadó, Budapest, 1967.
2. Valószínűségszámítási módszerek az analízisben. I. Mat. Lapok 18 (1967) 5—35. (cf.: 1966/19).
3. Valószínűségszámítási módszerek az analízisben II. Mat. Lapok 18 (1967) 175—194. (cf.: 1966/19).
4. Remarks on the Poisson process. Studia Sci. Math. Hung. 2 (1967) 119—123.
5. Statistics and information theory. Studia Sci. Math. Hung. 2 (1967) 249—256.
6. On the height of trees (with G. Szekeres). Jour. Australian Math. Soc. 7 (1967) 497—507.
7. Remarks on the Poisson process. Symposium on probability methods in analysis. Lectures delivered at a symposium at Loutraki, Greece, 22. 5.—4.6. 1966. Lecture Notes in Mathematics, 31. Springer, Berlin, 1967; pp. 280—286.
8. On some basic problems of statistics from the point of view of information theory. Proceedings of the Fifth Berkeley Symposium on mathematical statistics and probability. Held at the Statistical Laboratory, University of California, June 21—July 18, 1965 and December 27, 1965—January 7, 1966. Vol. I. University of California Press, Berkeley and Los Angeles, 1967; pp. 531—543 (cf.: 1965/19).
9. Játék és matematika (I). Természettudományi Közlöny 11 (98) (1967) 61—63.
10. Játék és matematika (II). Természettudományi Közlöny 11 (98) (1967) 116—119.
11. Játék és matematika (III). Természettudományi Közlöny 11 (98) (1967) 211—213.
12. „Az ember gúnnyal — tudjuk — arra támad, amit meg nem ért!" (A televízióban elhangzott „Játék és matematika" c. sorozat befejező része.) Természettudományi Közlöny 11 (98) (1967) 296—298.
13. Dialogues on mathematics. Holden—Day, Inc., San Francisco, 1967. (cf.: 1965/7.)
14. Dialoge über Mathematik. VEB Deutscher Verlag der Wissenschaften, Berlin, 1967. (cf.: 1965/7.)

15. Dialoge über Mathematik. Birkhäuser, Basel, 1967. (cf.: 1965/7.)
16. Probabilistic methods in combinatorial mathematics. (Lecture: Symposium on combinatorial mathematics, Chapel Hill, 1967.) Mimeographed text: Symposium on combinatorial mathematics, Chapel Hill. University of North Carolina. Monographs Series, 1967. 13 p.
17. Probabilistic methods in combinatorial mathematics. Combinatorial mathematics and its applications. Proceedings of the Conference held at the University of North Carolina at Chapel Hill, April 10—14, 1967. Preprint. 13 p. (cf.: 1967/16).
18. Letters on probability. 4th letter. Az MTA Matematikai Kutatóintézete, Budapest, 1967. 18 p. (Mimeographed.)
19. Dialoguri despre matematică. Editura Ştiintifică, Bucureşti, 1967. (cf.: 1965/7).
20. On some problems in the thory of order statistics. Az MTA Matematikai Kutatóintézete, Budapest, 1967. 17 p. (Mimeographed.)
21. On some problems in the theory of order statistics. 36th Session of the International Statistical Institute, Sydney, Australia, 28 August to 7 September 1967. Sydney, 1967. 17 p. (Mimeographed.)

1968

1. Kerekasztal-konferencia szovjet matematikusokkal a matematika elvi kérdéseiről. Mat. Lapok 19 (1968) 3—8.
2. A rendezett minták elméletének egy problémaköréről. MTA III. Oszt. Közl. 18 (1968) 23—30. (cf.: 1967/20).
3. Zufällige konvexe Polygone in einem Ringgebiet (mit R. Sulanke). Zeitschr. Wahrscheinlichkeitstheorie 9 (1968) 146—157.
4. On quadratic inequalities in the theory of probability (with J. Galambos). Studia Sci. Math. Hung. 3 (1968) 351—358.
5. On random matrices II. (with P. Erdős). Studia Sci. Math. Hung. 3 (1968) 459—464.
6. Information and statistics. Studies in mathematical statistics. Theory and applications. Akadémiai Kiadó, Budapest, 1968; pp. 129—131. (cf.: 1964/21).
7. Sur la théorie de la recherche aléatoire. Colloques internationaux du Centre National de la Recherche Scientifique. No. 165. Programmation en mathématiques numériques. Besançon, 7—14 Septembre 1966. Éditions du Centre National de la Recherche Scientifique, Paris, 1968; pp. 281—287.
8. On the distribution of numbers prime to *n*. Abhandlungen aus Zahlentheorie und Analysis. Zur Erinnerung an Edmund Landau (1877—1938). VEB Deutscher Verlag der Wissenschaften, Berlin 1968; pp. 269—278.
9. Valószínűségszámítás. Második kiadás. Tankönyvkiadó, Budapest, 1968. (cf.: 1966/15).
10. Változatok egy Fibonacci-témára. Természet Világa, 1968, 22—27.
11. Változatok egy Fibonacci-témára (II). Természet Világa, 1968, 87—90.
12. Ars Mathematica. Fizikai Szemle 18 (1968) 60—61.
13. Ars Mathematica. Élet és Tudomány 23 (1968) 654—655. (cf.: 1968/12).
14. Die Sprache des Buches der Natur. Neue Sammlung 8 (1968) 117—123. (cf.: 1964/5).

15. Some remarks on the large sieve of Yu. V. Linnik (with P. Erdős).[1] Annales Univ. Sci. Budapest. Sect. Math. 11 (1968) 3—13.
16. On some problems of statistics from the point of view of information theory. (Lecture: Információelméleti Kollokvium. Kossuth Lajos Tudományegyetem, Debrecen 1967, szeptember 19—24.) Abstract: Információelméleti Kollokvium. Kossuth Lajos Tudományegyetem, Debrecen 1967. szeptember 19—24. Előadáskivonatok. Bolyai János Matematikai Társulat, Budapest, 1968; pp. 87—90.
17. Stochastische Prozesse in der Biologie. (Lecture: A Nemzetközi Biometriai Társaság Magyar Csoportjának Biometriai Konferenciája. Magyar Tudományos Akadémia, Budapest 1968. március 19—22.) Abstract: A Nemzetközi Biometriai Társaság Magyar Csoportjának Biometriai Konferenciája. Magyar Tudományos Akadémia, Budapest 1968. március 19—22. Előadáskivonatok. Az MTA Biológiai Tudományok Osztálya, Budapest, 1968; p. 50.
18. Über die Potenzreihen ganzer Funktionen (mit P. Erdős). (Lecture: Tagung über komplexe Analysis. Oberwolfach, 8.—14. September 1968.) (Manuscript.)

1969

1. On random entire functions (with P. Erdős). Zastosowania Matematyki 10 (1969) 47—55. (cf.: 1968/19).
2. Measures in denumerable spaces (with A. Hanisch and W. M. Hirsch). American Math. Monthly 76 (1969) 494—502.
3. Lectures on the theory of search. Department of Statistics. University of North Carolina at Chapel Hill. Institute of Statistics Mimeo Series No. 600. 7. May 1969. 78 p.
4. A szerencsejátékok és a valószínűségszámítás. Matematikai érdekességek. Gondolat, Budapest, 1969; pp. 197—220.
5. A Barkochba játék és az információelmélet. Matematikai érdekességek. Gondolat, Budapest, 1969; pp. 269—286.
6. Dialógusok a matematikáról. 3. kiadás. Akadémiai Kiadó, Budapest, 1969. (cf.: 1965/7).
7. Диалоги о математике. Мир, Москва, 1969. (cf.: 1965/7).
8. Briefe über die Wahrscheinlichkeit. Akadémiai Kiadó, Budapest, 1969. (cf.: 1967/1).
9. Briefe über die Wahrscheinlichkeit. VEB Deutscher Verlag der Wissenschaften, Berlin, 1969. (cf.: 1967/1).
10. Briefe über die Wahrscheinlichkeit. Birkhäuser, Basel, 1969. (cf.: 1967/1).
11. On some problems of statistics from the point of view of information theory. Proceedings of the Colloquium on Information Theory organized by the Bolyai Mathematical Society, Debrecen (Hungary), 1967. Bolyai János Matematikai Társulat, Budapest, 1969; pp. 343—357.
12. Lezioni sulla probabilità e l'informazione. Lezioni e conferenze. Università di Trieste, Istituto di Mecanica. Trieste, 1969 (to appear).
13. Remarks on the teaching of probability. Lecture to the First CSMP International Conference (March 18—27, 1969). The International Conference on the Teaching of Probability and Statistics at the Pre-College Level. Co-Sponsored by Southern Illinois University at Carbondale and Central Midwestern Regional Educational Laboratory. Carbondale, 1969. 16 p. (Mimeographed.)

14. On the enumeration of trees. Lecture to the International Conference on Combinatorial Structures and their Applications, Calgary, 1969. (Mimeographed.)
15. On Cayley's polynomials for counting trees. Proceedings of the Calgary International Conference on Combinatorial Structures and their Applications. Calgary, 1969. (To appear.)
16. Gondolatok a valószínűségszámítás tanításáról. Az MTA Matematikai Kutatóintézete, Budapest, 1969. (Mimeographed.) (cf.: 1969/13).
17. Mathematical models of biological processes. Lecture. Kingston, 1969. (Manuscript.)
18. Applications of probability theory to other areas of mathematics. Lectures held at the 12th Biennial International Seminar of the Canadian Mathematical Congress at Vancouver, August 1969. Preprint. Canadian Mathematical Society, Vancouver, 1969.
19. Probabilistic methods in combinatorial mathematics. Combinatorial mathematics and its applications. Proceedings of the Conference held at the University of North Carolina at Chapel Hill, April 10—14, 1967. (Edited by R. C. Bose and T. A. Dowling.) The University of North Carolina Press, Chapel Hill, 1969; pp. 1—13 (cf.: 1967/16).
20. On some problems in the theory of order statistics. Bull. Inst. Internat. Stat. 42 (1969) 165—176 (cf.: 1967/21).
21. My fourth letter to Pierre Fermat. By Blaise Pascal, Paris, France. First CSMP International Conference (March 18—27, 1969). The International Conference on the Teaching of Probability and Statistics at the Pre-College level. Co-sponsored by Southern Illinois University at Carbondale and Centre Midwestern Regional Educational Laboratory. Carbondale, 1969. 18 p. (Mimeographed.) (cf.: 1967/18).
22. On the distribution of numbers prime to *n*. Number theory and Analysis. A Collection of Papers in Honor of Edmund Landau (1877—1938). Plenum Press, New York, 1969; pp. 269—278. (cf.: 1968/8).
23. Ars Mathematica. (In English.) 1969. (To appear in the Festschrift in honor of Professor Herman Wold.) 5 p. (Manuscript.) (cf. 1968/12).

1970

1. Gondolatok a valószínűségszámítás tanításáról. Mat. Lapok 21 (1970) 31—37. (cf.: 1969/13).
2. Valószínűségszámítási feladatgyűjtemény (with K. Bognár, J. Mogyoródi, A. Prékopa, D. Szász). Tankönyvkiadó, Budapest, 1971.
3. Probability theory. Akadémiai Kiadó, Budapest — North-Holland Publishing Co., Amsterdam, 1970.
4. Foundations of probability theory. Holden-Day, Inc., San Francisco, 1970.
5. On the mathematical theory of trees. North-Holland Publishing Co., Amsterdam, 1970 (to appear).
6. Stochastische Prozesse in der Biologie. Vorträge der II. Ungarischen Biometrischen Konferenz (Budapest, vom 19. bis 22. März 1968). Akadémiai Kiadó, Budapest, 1970; pp. 27—33.
7. On a new law of large numbers (with P. Erdős). Journal d'Analyse Mathématique 23 (1970) 103—111.

8. The Prüfer code for k-trees (with C. Rényi). Combinatorial theory and its applications III. Bolyai János Matematikai Társulat, Budapest—North-Holland Publishing Co., Amsterdam—London, 1970; pp. 945—971.
9. On the enumeration of search codes. Acta Math. Acad. Sci. Hung. 21 (1970) 27—33.
10. Valószínűségszámítás. (In Czech.) Praha, 1970. (To appear.)
11. Uniform flow in cascade graphs. In a publication edited by M. Behara (Springer, Berlin). (To appear.) 19 p. (Manuscript.)
12. On the number of endpoints of a k-tree. Studia Sci. Math. Hung. 5 (1970) 5—10.
13. Naplójegyzetek az információelméletről. Part of a book in preparation for Magvető Kiadó, entitled „Ars Mathematica". 1970. (Manuscript.)
14. Dialógusok a matematikáról. 3. rész. (In Italian.) Sapere, 1970. (To appear.)
15. Applications of probability theory to other areas of mathematics. Proceedings of the 12th Biennial International Seminar of the Canadian Mathematical Congress (Vancouver, 1969). Vancouver, 1970. (To appear.) (cf.: 1969/18).
16. On the enumeration of trees. Proceedings of the Calgary International Conference on Combinatorial Structures and their Applications, Calgary, 1969. Gordon and Breach, New York 1970; pp. 355—360. (cf.: 1969/14).
17. Napló az információelméletről. Fizikai Szemle 20 (1970) 161—172. (cf.: 1970./13).
18. Remarks on the teaching of probability. The teaching of probability and statistics. Proceedings of First CSMP International Conference on Teaching of Mathematics at the Pre-college level. Jointly sponsored by Southern Illinois University and CEMREL. March, 1969. (Edited by Lennart Råde.) Almqvist and Wiksells, Stockholm, 1970; pp. 273—281. (cf.: 1969/13).
19. Письма о вероятности. Мир, Москва, 1971. (cf.: 1967/1).
20. Letters on probability. Akadémiai Kiadó, Budapest, 1970. (To appear.) (cf.: 1967/1).

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 23—25.

POLYNOMIAL APPROXIMATION ON THE REAL LINE

by

G. FREUD and J. T. SCHEICK

The aim of this work is to define a linear method of approximation by polynomials which, for functions f of restricted growth, will converge uniformly to f on each compact interval of the real line upon which f is continuous. The following theorem of CHLODOUSKI ([1], p. 36) is a prototype:

THEOREM. *Suppose* $\varrho_n \to \infty$ *and* $\varrho_n/n \to 0$ *where* $\varrho_n > 0$ *for all* n. *Let* f *be given and define* p_n *by*

$$p_n(x) = \sum_0^n f\left(\varrho_n \frac{k}{n}\right)\binom{n}{k}\left(\frac{x}{\varrho_n}\right)^k \left(1 - \frac{x}{\varrho_n}\right)^{n-k} = B_n\left(f(\varrho_n u), \frac{x}{\varrho_n}\right)^1.$$

If $\sup\{|f(t)|: 0 \leqq t \leqq \varrho_n\} = O(e^{(\alpha n/\varrho_n)})$ *for all* $\alpha > 0$ *then* $p_n(x_0) \to f(x_0)$ *at each point* x_0 *where* f *is continuous.*

We will construct a more flexible device of approximation. Let (μ_n), (ν_n) be sequences of natural numbers and put $K_n(t) = \left(\frac{\sin(\nu_n t/2)}{\sin(t/2)}\right)^{2\mu_n}$ where $\int_{-\pi}^{\pi} K_n(t)\,dt = \lambda_n$. Then K_n is an even trigonometric polynomial of degree $\mu_n(\nu_n - 1)$. Further, if g is an even integrable 2π periodic function, $\int_{-\pi}^{\pi} g(\Theta + t)K_n(t)\,dt$ is a polynomial in $\cos\Theta$ of degree not exceeding $\mu_n(\nu_n - 1)$. We wish to consider functions measurable on the real line and bounded on each bounded interval: such functions we shall call *locally bounded.* Let (ϱ_n) be a sequence of positive numbers. For f locally bounded, put $\|f\|_{\varrho_n} = \sup\{|f(t)|: |t| \leqq \varrho_n\}$. The expression

$$p_n(x) = p_n(f; x) = \lambda_n^{-1} \int_{-\pi}^{\pi} f\big(\varrho_n(\Theta + t)\big)K_n(t)\,dt, \qquad x = \varrho_n \cos\Theta \tag{1}$$

defines a sequence of polynomials with degree $p_n \leqq \mu_n \nu_n$. The basic theorem is

THEOREM 1. *Let* f *be locally bounded. Suppose*

$$\varrho_n \to \infty \quad \textit{and} \quad \varrho_n = o(\nu_n) \tag{2}$$

and

$$\|f\|_{\varrho_n} = o(e^{\mu_n}). \tag{3}$$

---

[1] $B_n(h; t)$ denotes the Bernstein polynomial of h evaluated at t.

Then, for the polynomials (1) *we have*

(4) *f continuous at x_0 implies $p_n(f; x_0) \to f(x_0)$*

(5) *f continuous at each point of $[a, b]$ implies $p_n(f) \to f$ uniformly on $[a, b]$.*

PROOF. It will be sufficient to prove (5) since we do not exclude $a=b$. First let us estimate the mass of the kernel on (s, π). From $x \geqq \sin x \geqq \left(1-\frac{x}{\pi}\right) x$ for $0<x<\pi$ we have

$$\lambda_n \geqq 2 \int_0^{2\pi/\nu_n} \left[\nu_n\left(1-\frac{\nu_n t}{2\pi}\right)\right]^{2\mu_n} dt = \frac{4\pi\nu_n^{2\mu_n-1}}{2\mu_n+1}$$

and $1 \geqq \sin\frac{x}{2} \geqq \frac{x}{\pi}$ for $0<x<\pi$ yields

$$\lambda_n^{-1} \int_s^\pi K_n(t)\,dt \leqq \frac{1}{\lambda_n} \int_s^\pi \left(\frac{\pi}{t}\right)^{2\mu_n} dt \leqq \frac{\pi}{\lambda_n(2\mu_n-1)} \left(\frac{\pi}{s}\right)^{2\mu_n-1}$$

whence

$$\lambda_n^{-1} \int_s^\pi K_n(t)\,dt \leqq \left(\frac{\pi}{\nu_n s}\right)^{2\mu_n-1} \quad \text{if} \quad 0<s<\pi.$$

Now let $\varepsilon>0$ be given. There exists $\delta>0$ such that $\{x\in[a, b]$ and $|y-x| \leqq \delta\} \Rightarrow |f(x)-f(y)| < \varepsilon$. But $|\varrho_n \cos(\Theta+t)-\varrho_n \cos\Theta| \leqq \delta$ if $|t| \leqq \delta/\varrho_n$. Thus if $x\in[a, b]$ and $x=\varrho_n \cos\Theta$, one has

$$|f(x)-p_n(x)| \leqq \frac{1}{\lambda_n} \int_{-\pi}^{\pi} |f(\varrho_n \cos(\Theta+t))-f(\varrho_n \cos\Theta)|\, K_n(t)\,dt \leqq$$

$$\leqq \frac{1}{\lambda_n} 2\varepsilon \int_0^{\delta/\varrho_n} K_n(t)\,dt + 4\frac{1}{\lambda_n} \|f\|_{\varrho_n} \int_{\delta/\varrho_n}^{\pi} K_n(t)\,dt \leqq 2\varepsilon + 4\|f\|_{\varrho_n} \left(\frac{\pi\varrho_n}{\delta\nu_n}\right)^{2\mu_n-1}.$$

Choose N so that for $n \geqq N$ one has

$$[a, b] \subset [-\varrho_n, \varrho_n]$$

and

$$\frac{\pi}{\delta}\frac{\varrho_n}{\nu_n} < e^{-1}.$$

Hence for all $x\in[a, b]$ and all $n \geqq N$

$$|f(x)-p_n(x)| \leqq 2\varepsilon + 4\|f\|_{\varrho_n} e^{-\mu_n}$$

from which the assertion follows.

The maps $f \to p_n(f)$ defined by (1) are linear, so (1) defines a linear method of approximation. We will now investigate the possibility of adjusting the parameters to provide a fixed linear method of approximation for a class of functions whose growth is known. We desire the additional condition that $\mu_n \nu_n \leqq n$ so degree $p_n \leqq n$, for all n. Let φ be even and continuous on the reals, increasing in $[0, \infty)$ and $\varphi(x) \to \infty$ as $x \to \infty$. Denote by $M(\varphi)$ the set of all locally bounded f for which $|f(t)| \leqq e^{\varphi(|t|)}$ for all t. It turns out that (μ_n) may be fairly arbitrary.

THEOREM 2. *Let (μ_n) be a sequence of natural numbers such as*

$$\mu_n \to \infty, \frac{\mu_n}{n} \to 0 \quad as \quad n \to \infty.$$

Define[2] $\nu_n = \left[\frac{n}{\mu_n}\right]$. *Then for each class $M(\varphi)$, there exists $\{\varrho_n\}$ such that for all $f \in M(\varphi)$, (4) and (5) of Theorem 1 hold.*

PROOF. Since $\|f\|_{\varrho_n} e^{-\mu_n} \leqq \exp(\varphi(\varrho_n) - \mu_n)$, it is sufficient to define $\{\varrho_n\}$ so that $\varphi(\varrho_n) \leqq \frac{1}{2}\mu_n$ for all n and (2) holds. To this end, let

$$x_n = \sup\left\{x \geqq 0 : \varphi(x) \leqq \frac{\mu_n}{2}\right\}.$$

Then $\varphi(x_n) = \frac{1}{2}\mu_n$ and $x_n \to \infty$. Put $n\varepsilon_n = \min\{\mu_n x_n, \sqrt{\mu_n n}\}$ and $\varrho_n = \nu_n \varepsilon_n$. Then $\varrho_n = \left(\frac{\mu_n}{n}\left[\frac{n}{\mu_n}\right]\right) \min\left\{x_n, \sqrt{\frac{n}{\mu_n}}\right\} \to \infty$ and $0 \leqq \frac{\varrho_n}{\nu_n} = \varepsilon_n \leqq \sqrt{\frac{\mu_n}{n}} \to 0$. Clearly $\varrho_n \leqq x_n$ so that $\varphi(\varrho_n) \leqq \varphi(x_n) = \frac{1}{2}\mu_n$; Q.e.d.

REFERENCE

[1] LORENTZ, G. G.: *Bernstein Polynomials*. University of Toronto Press, 1953.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

Ohio State University, Columbus, Ohio, U.S.A.

(Received January 5, 1970.)

[2] $[x]$ denotes the greatest integer not exceeding x.

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 27—40.

ON THE WEAK* CONTINUITY OF CONVOLUTION IN A CONVOLUTION ALGEBRA OVER AN ARBITRARY TOPOLOGICAL GROUP

by

I. CSISZÁR

To the memory of C. Rényi

Summary

It is proved that the convolution of τ-regular Borel measures on an arbitrary topological group X is joint continuous, i.e. $\mu_\alpha \to \mu$ and $\nu_\alpha \to \nu$ imply $\mu_\alpha \nu_\alpha \to \mu\nu$; this has been known so far only under the condition that the measures considered are uniformly tight. This result is obtained as a corollary of a general theorem on the weak* continuity of convolution of linear functionals on the space of uniformly continuous functions on X (theorem 1). In § 3 some possible applications and a generalization of theorem 1 are pointed out and also several open problems are mentioned.

§ 1. Preliminaries

Let X be an arbitrary topological group; let e denote the unit element of X and $\mathscr{V}$ the class of all symmetric (open) neighbourhoods of e.

Let $U_r(X)$ be the Banach space of all bounded and right uniformly continuous (real-valued) functions on X. I.e., $f \in U_r(X)$ means that $\|f\| = \sup_{x \in X} |f(x)| < \infty$ and for every $\varepsilon > 0$ there exists $V \in \mathscr{V}$ such that $x_1 x_2^{-1} \in V$ implies $|f(x_1) - f(x_2)| < \varepsilon$. The advantages of considering $U_r(X)$ rather than $C(X)$ (the Banach space of all bounded continuous functions on X) will be apparent soon.

Let $\mathscr{L}_r(X)$ denote the dual of $U_r(X)$ endowed with the weak* topology. I.e., $\mathscr{L}_r(X)$ is the set of all bounded linear functionals on $U_r(X)$ and $L_\alpha \to L$ means that

$$(1) \qquad L_\alpha f \to Lf \quad \text{for every} \quad f \in U_r(X);$$

here we have convergence of nets in the sense of Moore—Smith in mind, cf. e.g. [6], Chapter 2.

If $L \in \mathscr{L}_r(X)$, consider the linear operator $\bar{L}$ defined by

$$(2) \qquad (\bar{L}f)(x) = Lf_x, \quad f_x(y) = f(xy) \qquad (f \in U_r(X))$$

$\bar{L}$ maps $U_r(X)$ into itself; moreover, if for a given $f \in U_r(X)$ and $V \in \mathscr{V}$ we have $|f(x_1) - f(x_2)| < \varepsilon$ whenever $x_1 x_2^{-1} \in V$, then $\|f_{x_1} - f_{x_2}\| = \sup_{y \in X} |f(x_1 y) - f(x_2 y)| \leqq \varepsilon$ if $x_1 x_2^{-1} \in V$ (using the fact that $(x_1 y)(x_2 y)^{-1} = x_1 x_2^{-1}$), implying

$$(3) \qquad |(\bar{L}f)(x_1) - (\bar{L}f)(x_2)| = |L(f_{x_1} - f_{x_2})| \leqq \|L\| \varepsilon \quad \text{if} \quad x_1 x_2^{-1} \in V.$$

Here $\|L\| = \sup_{\|f\| \leq 1} |Lf|$ denotes the norm of the functional L; actually the operator $\bar{L}$ has the same norm:

$$\|\bar{L}\| = \sup_{\|f\| \leq 1} \|\bar{L}f\| = \sup_{\|f\| \leq 1} \sup_{x \in X} |Lf_x| = \sup_{\|f\| \leq 1} |Lf| = \|L\|.$$

We shall consider $\mathscr{L}_r(X)$ as a convolution algebra, cf. [4], § 19. The convolution LM of L and M in $\mathscr{L}_r(X)$ is defined as the functional $L(\bar{M}f)$ having the associated operator $\bar{L}\bar{M}$:

$$LMf = L(\bar{M}f); \quad \overline{LM}f = \bar{L}\bar{M}f \qquad (f \in U_r(X)). \tag{4}$$

Here the second equation follows from the first one using the relation

$$(\bar{M}f)_x(y) = (\bar{M}f)(xy) = Mf_{xy} = (\bar{M}f_x)(y); \tag{5}$$

(5) can be interpreted by saying that the left translations $f \to f_x$ commute with the operators $\bar{M}$ $(x \in X,\ M \in \mathscr{L}_r(X))$.

Observe that if we considered linear functionals on $C(X)$ instead of on $U_r(X)$, it could not be guaranteed, in general, that $\bar{L}$ maps $C(X)$ into itself, thus difficulties would arise with the very definition of convolution. Moreover, the estimate (3) will be essentially used in the sequel; this estimate has no analogon for functionals on $C(X)$, unless considerably restricting the scope of investigation (to tight functionals).

Of course, all the concepts introduced above have their "left" analogues: $U_l(X)$ is the Banach space of all bounded functions on X for which to every $\varepsilon > 0$ there exists $V \in \mathscr{V}$ such that $x_1^{-1}x_2 \in V$ implies $|f(x_1) - f(x_2)| < \varepsilon$. $\mathscr{L}_l(X)$ is the dual of $U_l(X)$, and if $L \in \mathscr{L}_l(X)$, an operator $\tilde{L}$: $U_l(X) \to U_l(X)$ corresponding to L is defined by $(\tilde{L}f)(x) = Lf_x$ where now $f_x(y) = f(yx)$. The convolution in $\mathscr{L}_l(X)$ is defined by $(LM)f = M(\tilde{L}f)$. On account of the complete symmetry, nothing will be lost by restricting attention to $\mathscr{L}_r(X)$ (this choice is the contrary of that having been made in [2] but it is more convenient for the present paper).

We shall be interested in the continuity of the operation of convolution, i.e., we look for a possibly weak condition under which $L_\alpha \to L$ and $M_\alpha \to M$ imply $L_\alpha M_\alpha \to LM$. The following example shows that this implication can not be expected to hold unconditionally.

Example 1. Let X be the additive group of real numbers with the usual topology and consider the sequences L_n and M_n defined by $L_n f = f(n)$, $M_n(f) = f(-n)$. As the unit sphere in the dual of a Banach space is weakly compact, there exist convergent subnets $L_\alpha \to L$ and $M_\alpha \to M$ of the sequences L_n and M_n, respectively (the underlying directed sets of both subnets may be assumed to be the same). If $f \in U_r(X)$ is such that $\lim_{x \to -\infty} f(x) = a$ exists, then $\lim_{x \to \infty} M_n f_x = a$ for every $x \in X$, implying $(\bar{M}f)(x) = Mf_x = a$ for every $x \in X$, thus $LMf = a$, as well. On the other hand $L_n M_n = N$ for every n, where $Nf = f(0)$, thus $L_\alpha M_\alpha \to N \neq LM$.

This example shows, too, that the operation of convolution need not be commutative even if X is Abelian. In fact if L and M are such as above, then for every $f \in U_r(X)$ for which both $\lim_{x \to -\infty} f(x) = a$ and $\lim_{x \to \infty} f(x) = b$ exist and $a \neq b$, we have $LMf = a \neq MLf = b$.

Our interest in the convolution algebra $\mathscr{L}_r(X)$ is motivated by that in the convolution of probability measures on the group X. A Borel-measure on X, i.e. a measure[1] μ defined on the σ-algebra $\mathscr{B}$ of Borel subsets[2] of X is called τ-regular if for every increasing net of open sets $G_\alpha \subset X$ we have

$$\lim \mu(G_\alpha) = \mu\left(\bigcup_\alpha G_\alpha\right). \tag{6}$$

Let $\mathscr{L}_r^+(X)$ denote the set of all positive linear functionals on $U_r(X)$. A positive linear functional $M \in \mathscr{L}_r^+(X)$ is called τ-continuous if for every decreasing net of functions $f_\alpha \in U_r(X)$ converging pointwise to 0 we have $\lim Mf_\alpha = 0$.

It is an easy consequence of a general measure-theoretic theorem (cf. [8], II—3, II—7) that there is a one-to-one correspondence between τ-continuous positive linear functionals on $U_r(X)$ and τ-regular Borel measures on X, established by

$$Mf = \int f(x)\mu(dx) \tag{7}$$

(see [2], lemma 1, with the unsignificant difference that in [2] attention was restricted to functionals with $M1 = 1$ thus the resulting Borel measures were probability measures).

Let us remark that a τ-regular Borel measure on a topological group X (or, more generally, on an arbitrary completely regular topological space) is a regular measure in the sense that $\mu(B) = \inf_{B \subset G} \mu(G)$ for every Borel set $B \subset X$ where G ranges over the open sets containing B, cf. [2], remark 1.

A sufficient (but not necessary) condition for a Borel measure to be τ-regular is

$$\mu(B) = \sup_{K \subset B} \mu(K) \quad \text{for every} \quad B \in \mathscr{B} \tag{8}$$

where K ranges over the compact subsets of B. The measures with the property (8) are called K-regular (or Radon) measures. A functional $M \in \mathscr{L}_r^+(X)$ can be represented by a K-regular measure μ if and only if M is tight, i.e. if to every $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset X$ such that $f \in U_r(X)$, $0 \leqq f \leqq 1$ and $f(x) = 0$ for $x \in K_\varepsilon$ imply $Mf < \varepsilon$ (by the aid of Dini's theorem easily follows that a tight functional is τ-continuous, thus it can be represented by a τ-regular Borel measure μ; from (7) one concludes that $\mu(X \setminus K_\varepsilon) < \varepsilon$, thus, as μ is regular, it is K-regular; conversely, if μ in (7) is K-regular, M is obviously tight).

Let us also remark that the convergence of a net of τ-continuous functionals $M_\alpha \in \mathscr{L}_r^+(X)$ to a τ-continuous functional $M \in \mathscr{L}_r^+(X)$ is equivalent to the usual (weak) convergence $\mu_\alpha \to \mu$ of the corresponding τ-regular measures. In fact, if $\int f(x)\mu_\alpha(dx) \to \int f(x)\mu(dx)$ for every $f \in U_r(X)$ and μ is τ-regular then the same limit relation holds for every $f \in C(X)$, where $C(X)$ is the Banach space of all bounded continuous functions on X (for a proof see e.g. [2], lemma 3).

---

[1] We restrict attention to finite measures, i.e. $\mu(X) < \infty$ will be always assumed without saying this explicitely.

[2] In the literature, the term "Borel set" is used in two different senses: it means a set belonging to the smallest σ-algebra containing either all compact sets or all closed (or, equivalently, all open) sets. In this paper, we have the latter definition in mind.

The one-to-one correspondence between τ-continuous positive linear functionals on $U_r(X)$ and τ-regular Borel measures on X enables one to define the convolution of τ-regular Borel measures as the convolution in $\mathscr{L}_r(X)$ (it is obvious that the convolution of τ-continuous positive linear functionals is again such a functional). The operation defined in this way deserves the name of convolution of measures because of the following proposition, having been proved in [2] (theorem 1 and its corollary).

PROPOSITION 1. *If μ and ν are τ-regular Borel measures on X, there is a unique τ-regular Borel measure $\mu\nu$ on X, with the property that*

$$\int f(x)\,\mu\nu(dx) = \int\Big(\int f(xy)\,\mu(dx)\Big)\,\nu(dy) = \int\Big(\int f(xy)\,\nu(dy)\Big)\,\mu(dx) \tag{9}$$

for every bounded Borel-measurable function f on X. This $\mu\nu$ is the convolution of μ and ν in the sense described above and it is uniquely determined already by the integrals of functions in $U_r(X)$.

Let us remark that when defining the convolution of τ-regular Borel measures on X one could have used instead of $\mathscr{L}_r(X)$ the convolution algebra $\mathscr{L}_l(X)$, as well; the above proposition says, in particular, that both approaches yield the same result.

According to the author's knowledge, no similar proposition is available for arbitrary Borel measures; this suggests that when defining the convolution of Borel measures on X one has to restrict attention to τ-regular measures.

§ 2. A general theorem on the continuity of convolution in $\mathscr{L}_r(X)$

It is well-know that if μ_α and ν_α are uniformly tight[3] nets of K-regular Borel measures on a topological group (or even a semigroup) X converging to (K-regular) Borel measures μ and ν, respectively, then $\mu_\alpha\nu_\alpha \to \mu\nu$. A proof of this statement is contained essentially in GRENANDER's book [3], where, however, the assertion is given in a less general form; as TORTRAT [11] has pointed out, GRENANDER's proof actually yields the result formulated above.

We are going to consider the more general problem of the continuity of convolution in $\mathscr{L}_r(X)$. As a corollary of theorem 1 below we shall see that for τ-regular Borel measures on a topological group X $\mu_\alpha \to \mu$ and $\nu_\alpha \to \nu$ always imply $\mu_\alpha\nu_\alpha \to \mu\nu$, without any additional assumption. In the proof, the group property of X will be essentially used, and it remains open whether such a general result holds for semigroups as well.*

Let us send forward the following lemma:

LEMMA 1. *For $L \in \mathscr{L}_r^+(X)$ (positive linear functional on $U_r(X)$) the following properties are equivalent:*

(i) *if f_α is a decreasing net of (right) uniformly equicontinuous functions on X converging pointwise to 0 then $Lf_\alpha \to 0$;*

[3] μ_α is uniformly tight if for every $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset X$ such that $\mu_\alpha(X \setminus K_\varepsilon) < \varepsilon$ for all α.

* Added in proof: F. Topsoe settled this question in the affirmative (private communication).

(ii) *to every* $V \in \mathscr{V}$ *and* $\varepsilon > 0$ *there exist a finite subset* $x_1, \dots, x_n$ *of* X *and a function* $g \in U_r(X)$ *such that* $0 \leqq g \leqq 1$, $g(x_i) = 0$, $i = 1, \dots, n$, $g(x) = 1$ *for* $x \notin \bigcup_{j=1}^{n} Vx_i$ *and* $Lg < \varepsilon$;

(iii) *if* f_α *is an arbitrary net of uniformly bounded and (right) uniformly equicontinuous function converging pointwise to* 0 *then* $Lf_\alpha \to 0$.

(Recall that $\mathscr{V}$ is the class of symmetric neighborhoods of the unit element e of X).

PROOF. Let $0 \leqq d(x) \leqq 1$ be a function in $U_r(X)$ vanishing at e and equaling 1 outside V; the existence of such a d is an easy consequence of the fact that the sets of form $\{X: f(x) < 1\}$, $f \in U_r(X)$ are a base at the identity. Let α range over the finite subsets of X and for

$$\alpha = (x_1, x_2, \dots, x_n) \quad \text{set} \quad g_\alpha(x) = \min_{1 \leqq i \leqq n} d(xx_i^{-1}).$$

Then the functions $g_\alpha(x)$ are (right) uniformly equicontinuous and they form a decreasing net converging pointwise to 0 (the ordering for the α's being the set-theoretical inclusion). Hence, if L has the property (i), $Lg_\alpha < \varepsilon$ for some $\alpha = (x_1, \dots, x_n)$. Here, by the definition of g_α, $g_\alpha(x_i) = 0$, $i = 1, \dots, n$ and $g_\alpha(x) = 1$ if $x \notin \bigcup_{i=1}^{n} Vx_i$; i.e., (i) implies (ii).

Let now f_α be an arbitrary net of uniformly bounded and (right) uniformly equicontinuous functions converging pointwise to 0; without any loss of generality, we assume that $\|f_\alpha\| \leqq 1$ for every α.

For given $\varepsilon > 0$ pick $V \in \mathscr{V}$ such that $x_1 x_2^{-1} \in V$ implies $|f_\alpha(x_1) - f_\alpha(x_2)| < \varepsilon$ for every α. If there exist $x_1, \dots, x_n$ and g with the properties in (ii), take an α_0 such that

$$(10) \qquad |f_\alpha(x_i)| < \varepsilon \qquad i = 1, \dots, n, \quad \alpha \succ \alpha_0;$$

then we have $|f_\alpha(x)| < 2\varepsilon$ for every $x \in \bigcup_{i=1}^{m} Vx_i$ and hence $\|f_\alpha(1-g)\| \leqq 2\varepsilon$. Since L is positive, $|L(f_\alpha g)| \leqq L|f_\alpha g| \leqq Lg < \varepsilon$ thus we obtain

$$(11) \qquad |Lf_\alpha| \leqq |L(f_\alpha g)| + \big|L\big(f_\alpha(1-g)\big)\big| \leqq \varepsilon + \|L\| \cdot 2\varepsilon,$$

proving the implication (ii) $\to$ (iii).

Finally, (iii) trivially implies (i), since if f_α is a decreasing net converging to 0 then $\|f_\alpha\| \leqq \|f_{\alpha_0}\|$ for $\alpha \succ \alpha_0$.

DEFINITION. A functional $L \in \mathscr{L}_r^+(X)$ satisfying the equivalent conditions (i)—(iii) will be called ϱ-continuous.

Of course, every τ-continuous $L \in \mathscr{L}_r^+(X)$ is ϱ-continuous (by (i)) but the converse is not true, in general.

Example 2. Let X be the group of rational numbers $0 \leqq x \leqq 1$ with addition mod 1 and with the usual metric topology (regarding the interval $[0, 1)$ as a circle). If $f \in U_r(X)$, let Lf denote the Lebesgue integral of the unique continuous extension of f to $[0, 1)$. Then $L \in \mathscr{L}_r^+(X)$ is ϱ-continuous, since for any decreasing net of uniformly equicontinuous functions converging to 0 on X, the net of continuous extensions to $[0, 1)$ also decreases to 0. On the other hand, L is not τ-continuous, nor can be represented by a (σ-additive) measure on X.

THEOREM 1. *Let $L_\alpha \to L$ be a convergent net in $\mathscr{L}_r^+(X)$ and $M_\alpha \to M$ a convergent net in $\mathscr{L}_r(X)$; then, if L is ϱ-continuous, we have $L_\alpha M_\alpha \to LM$.*

PROOF. Let $f \in U_r(X)$ and $\varepsilon > 0$ be arbitrary; pick $V \in \mathscr{V}$ such that $x_1 x_2^{-1} \in V$ implies $|f(x_1) - f(x_2)| < \varepsilon$. As $M_\alpha \to M$ in $\mathscr{L}_r(X)$, it follows from the Banach—Steinhaus theorem that there exists an α_0 and a $K > 0$ such that $\|M_\alpha\| \leqq K$ for $\alpha \succ \alpha_0$; then, of course, $\|M\| \leqq K$, as well. Thus from (3) we have for $\alpha \succ \alpha_0$

(12)
$$|(\overline{M}_\alpha f)(x_1) - (\overline{M}_\alpha f)(x_2)| \leqq K\varepsilon, \quad |(\overline{M} f)(x_1) - (\overline{M} f)(x_2)| \leqq K\varepsilon \quad \text{if} \quad x_1 x_2^{-1} \in V.$$

According to (4) one may write

(13)
$$L_\alpha M_\alpha f = L_\alpha(\overline{M}_\alpha f) = L_\alpha(\overline{M} f) + L_\alpha h_\alpha$$
with

(14)
$$h_\alpha = \overline{M}_\alpha f - \overline{M} f.$$

Let $x_1, \dots, x_n$ and g be chosen as in the lemma, (ii). Since $M_\alpha \to M$ means $(\overline{M}_\alpha f)(x) = M_\alpha f_x \to M f_x = (\overline{M} f)(x)$ for every fixed $x \in X$ (see (1), (2)), there exists an α_0 (which, without any loss of generality, may be assumed to be the same as in (12)) such that for $\alpha \succ \alpha_0$

(15)
$$|(\overline{M}_\alpha f)(x_i) - (\overline{M} f)(x_i)| < \varepsilon, \qquad i = 1, \dots, n.$$

(12) and (15) imply for $\alpha \succ \alpha_0$

(16)
$$|h_\alpha(x)| = |(\overline{M}_\alpha f)(x) - (\overline{M} f)(x)| < (2K+1)\varepsilon \quad \text{if} \quad x \in \bigcup_{i=1}^n V x_i;$$

hence, taking into account that $1 - g(x) = 0$ if $x \notin \bigcup_{i=1}^n V x_i$,

(17)
$$\|h_\alpha(1-g)\| \leqq (2K+1)\varepsilon \qquad (\alpha \succ \alpha_0).$$

From (14) follows also

(18)
$$\|h_\alpha\| \leqq (\|\overline{M}_\alpha\| + \|\overline{M}\|)\|f\| \leqq 2K\|f\| \quad (\alpha \succ \alpha_0)$$

(recall that the operator $\overline{M}$ has the same norm as the functional M).

Since $L_\alpha \in \mathscr{L}_r^+(X)$ and $g \geqq 0$, we have $|L_\alpha hg| \leqq \|h\| L_\alpha g$ for any $h \in U_r(X)$; thus (17) and (18) imply

(19)
$$|L_\alpha h_\alpha| \leqq |L_\alpha h_\alpha g| + |L_\alpha h_\alpha(1-g)| \leqq 2K\|f\| L_\alpha g + (2K+1)\varepsilon \|L_\alpha\| \quad \text{if} \quad \alpha \succ \alpha_0.$$

Here $L_\alpha g \to Lg < \varepsilon$ and $\|L_\alpha\| = L_\alpha 1 \to L1$, thus, as $\varepsilon > 0$ has been arbitrary, (19) shows that the second term on the right hand side of (13) converges to 0. Since, on the other hand, $L_\alpha(\overline{M}f) \to L(\overline{M}f) = LMf$, we arrive at

$$(20) \qquad L_\alpha M_\alpha f \to LMf \quad \text{for every} \quad f \in U_r(X)$$

completing the proof.

COROLLARY. *Let μ_α and ν_α be arbitrary nets of τ-regular Borel measures on X converging to τ-regular Borel measures μ and ν, respectively. Then $\mu_\alpha \nu_\alpha \to \mu\nu$.*

PROOF. The theorem applies for the τ-continuous functionals corresponding to the measures under consideration. Since the weak convergence of τ-regular measures is equivalent to the convergence of the corresponding τ-continuous functionals, the corollary follows.

Remarks. (i) As it has been pointed out in § 1, in the convolution algebra $\mathscr{L}_r(X)$ $L_\alpha \to L$ and $M_\alpha \to M$ do not imply $L_\alpha M_\alpha \to LM$ without any supplementary assumption, even if one restricts attention to $\mathscr{L}_r^+(X)$. In this respect, when considering $\mathscr{L}_r^+(X)$ only, theorem 1 can be regarded as very satisfactory, since the condition of ϱ-continuity of L is a rather weak one. Nevertheless, it would be interesting (although from the point of view of probability theory apparently irrelevant) to find a similar simple condition under which $L_\alpha \to L$ and $M_\alpha \to M$ would imply $L_\alpha M_\alpha \to LM$ also if L_α and L are not assumed to belong to $\mathscr{L}_r^+(X)$.

(ii) We have been interested in the joint continuity of the operation of convolution in $\mathscr{L}_r(X)$. As regards separate continuity, it is an immediate consequence of the definitions that $L_\alpha \to L$ always implies $L_\alpha M \to LM$. On the other hand, $M_\alpha \to M$ does not imply $LM_\alpha \to LM$ without any additional assumption. In fact, if X, L, M_α and M are the same as in example 1, and if for $f \in U_r(X)$ both $\lim_{x \to -\infty} f(x) = a$ and $\lim_{x \to +\infty} f(x) = b$ exist, then $LM_\alpha f = a$ for every α while $LMf = b$. It may be conjectured that if $L \in \mathscr{L}_r(X)$ is such that $LM_\alpha \to LM$ whenever $M_\alpha \to M$ then also $L_\alpha M_\alpha \to LM$ whenever $L_a \to L$, $M_\alpha \to M$.

§ 3. Some possible applications and generalizations

In [1] we have studied infinite convolutions of probability measures on a locally compact topological group. TORTRAT [11], [12] has considered similar problems on arbitrary topological groups but under a uniform tightness condition. SAZONOV and TUTUBALIN [9] pointed out that the methods and results of [1] are also valid under considerably weaker hypotheses than local compactness. Nevertheless, the study of infinite convolutions of arbitrary probability measures on arbitrary topological groups seems to be a very complex problem and in the general case no such nice results can be hoped for as those obtained in the mentioned papers. Of course, the general problem of infinite convolutions in the convolution algebra $\mathscr{L}_r(X)$ is even more complex.

The theorem proved in § 2 can be regarded as a first step in attacking this problem. Let us restrict attention to positive linear functionals on $U_r(X)$ with norm 1. Such functionals will be referred to as *means*; the class of all means over $U_r(X)$

will be denoted by $\mathcal{M}_r(X)$: $M \in \mathcal{M}_r(X)$ if and only if $M \in \mathcal{L}_r^+(X)$ and $\|M\| = M1 = 1$. The set of all ϱ-continuous means will be denoted by $\mathcal{M}_r^\varrho(X)$.

Observe that $\mathcal{M}_r(X)$ is a weak* closed subset of the unit sphere in the dual of the Banach space $U_r(X)$, thus $\mathcal{M}_r(X)$ is a compact subset of $\mathcal{L}_r(X)$; $\mathcal{M}_r(X)$ is a semigroup with respect to the operation of convolution, but it is no topological semigroup, in general. $\mathcal{M}_r^\varrho(X)$ is a topological semigroup (the convolution of ϱ-continuous means is again ϱ-continuous, on account of the fact that if $M \in \mathcal{M}_r^\varrho(X)$ and f_α is a net of (right) uniformly equicontinuous functions converging decreasingly to 0 then $\overline{M}f_\alpha$ shares these properties, by (3) and the definition of ϱ-continuity; the joint continuity of the operation of convolution in $\mathcal{M}_r^\varrho(X)$ follows from theorem 1) but $\mathcal{M}_r^\varrho(X)$ is not compact, in general.

The following is an easy consequence of theorem 1.

PROPOSITION 2. *Let $\mathcal{M} \subset \mathcal{M}_r^\varrho(X)$ be a set of ϱ-continuous means having the property that $L \in \mathcal{M}$ and $LM \in \mathcal{M}$ imply $M \in \mathcal{M}$.*

Let T be an arbitrary directed set which does not have a maximal element and let us be given means L_s^t (for every pair s, t in T with $s \prec t$) belonging to $\mathcal{M}$ and satisfying

$$L_r^s L_s^t = L_r^t \qquad (r, s, t \in T, r \prec s \prec t). \tag{21}$$

Then if $M \in \mathcal{M}_r(X)$ is any cluster point of the net L_r, $t \succ r$ (with $r \in T$ fixed), there exist a directed set A and a function $t(\alpha)$ on A such that to every $t_0 \in T$ there is an $\alpha_0 \in A$ with $t(\alpha) \succ t_0$ whenever $\alpha \succ \alpha_0$ for which the subnets $L_s^{t(\alpha)}$ of the nets L_s^t, $t \succ s$ (s being regarded as fixed) are convergent for all $s \in T$ and, denoting the limits of these subnets by M_s, we have $M_r = M$, moreover, also the subnet $M_{t(\alpha)}$ of the net M_t is convergent (to M_∞, say).

If $M = M_r \in \mathcal{M}$, we hawe $M_s \in \mathcal{M}$ for all $s \in T$, $M_\infty \in \mathcal{M}$ and

$$L_r^s M_s = M_r = M \quad (s \succ r), \; M_s M_\infty = M_s \quad (s \geqq r), \quad M_\infty M_\infty = M_\infty. \tag{22}$$

Moreover, if $t(\alpha)$ and $t'(\alpha)$ are nets[4] *with the above properties constructed to the (different or not) cluster points $M \in \mathcal{M}$ and $M' \in \mathcal{M}$ of the net L_r^t (for the same $r \in T$) and if M_s, M_s', M_∞ and M_∞' are the corresponding limits, we have*

$$M' = MN, \quad NN' = M_\infty \tag{23}$$

where N and N' are arbitrary cluster points of the nets $M_{t(\alpha)}'$ and $M_{t'(\alpha)}$ respectively. Such cluster points always exist and they belong to $\mathcal{M}$.

PROOF. The existence of the limits for a proper choice of $t(\alpha)$ is an immediate consequence of the compactness of $\mathcal{M}_r(X)$: consider in the topological product $Z = \bigtimes_{s \in T} Z_s$, $Z_s = \mathcal{M}_r(X)$ for every $s \in T$, the net z^t, $t \in T$ where the s'th coordinate of z^t is L_s^t if $s \prec t$ and some fixed $L_0 \in \mathcal{M}_r(X)$ if $s \succeq t$. Since Z is compact (by Tyhonov's theorem), there exists a convergent subnet of z^t for which r'th coordinate converges to the given M; taking again an appropriate subnet, it may be achieved that in addition to the already existing limits M_s also $\lim M_{t(\alpha)} = M_\infty$ exists.

[4] The underlying directed sets may be assumed to be identical without any loss of generality.

Since $L_r^s \in \mathcal{M}$ is ϱ-continuous, from (21) follows by theorem 1 $L_r^s M_s = M_r$ for every $r \prec s$. Thus the first relation of (22) is true, whence in case of $M_r = M \in \mathcal{M}$, by the assumption on $\mathcal{M}$, follows that $M_s \in \mathcal{M}$ $(s \succ r)$. Rewriting the relation $L_r^s M_s = M_r$ $(r \prec s)$ as $L_s^t M_t = M_s$ $(s \prec t)$ and taking limits as $t = t(\alpha)$ (this is legitimate by theorem 1, if $s \succeq r$, since $\lim L_s^{t(\alpha)} = M_s \in \mathcal{M}$ is ϱ-continuous) we obtain the second relation in (22). This relation, on the other hand, implies $M_\infty \in \mathcal{M}$ by the assumption on $\mathcal{M}$.

Finally, the third relation of (22) obviously follows from the second one and from $\lim M_{t(\alpha)} = M_\infty$, (even without referring to theorem 1).

To prove the remaining assertion, first observe that by the compactness of $\mathcal{M}_r(X)$, the nets $M'_{t(\alpha)}$ and $M_{t'(\alpha)}$ surely have cluster points N and N' (say). Let $t(\beta)$ and $t'(\beta)$ be subnets of $t(\alpha)$ and $t'(\alpha)$ converging to N and N' respectively. In view of the first relation in (22), we have $L_r^s M'_s = M'$; taking limits as $s = t(\beta)$ and having in mind that $t(\beta)$ is a subnet of $t(\alpha)$, we obtain by theorem 1 the first relation of (23). This relation implies, by the assumption on $\mathcal{M}$, also $N \in \mathcal{M}$ and, by symmetry, $N' \in \mathcal{M}$ as well.

To prove the second relation of (23) take limits in $L_s^t M_t = M_s$ $(s \prec t)$ as $t = t'(\beta)$ to obtain $M'_s N' = M_s$; hence, by taking limits as $s = t(\beta)$ and having in mind that $t(\beta)$ is a subnet of $t(\alpha)$, the assertion follows.

The proof is complete.

A typical example of means with the property (21) is obtained by considering a sequence $N_1, N_2, \ldots$ of means in $\mathcal{M}$ and setting $L_k^n = N_k N_{k+1} \ldots N_{n-1}$ (in this case T is the set of positive integers with the natural ordering). If $\mathcal{M}$ is a semigroup with respect to the operation of convolution, $N_n \in \mathcal{M}$ $(n = 1, 2, \ldots)$ implies $L_k^n \in \mathcal{M}$ for every $k < n$ and proposition 2 applies. Observe that this particular case of proposition 2 is just the specialization to the present problem of theorem 2. 1 of [1], which has been used there as a starting point for the investigation of infinite convolutions of probability measures on a (locally compact) topological group. Here we preferred to formulate the proposition for a general set of indices, having possible applications to the study of X-valued continuous parameter stochastic processes or random fields with independent increments in mind.

Our first concern is to see whether proposition 2 applies already for $\mathcal{M} = \mathcal{M}_r^\varrho(X)$.

Let $\mathcal{M}_r^p(X)$ denote the set of all such means $L \in \mathcal{M}_r(X)$ for which the operator $\bar{L}$ defined by (2) is strictly positive in the sense that $f \in U_r(X)$, $f \geqq 0$, $\bar{L}f = 0$ implies $f = 0$.

LEMMA 2. *Let L and M be such means that $L \in \mathcal{M}_r^p(X)$ and $LM \in \mathcal{M}_r^\varrho(X)$. Then $M \in \mathcal{M}_r^\varrho(X)$.*

PROOF. Let f_α be a decreasing net of (right) uniformly equicontinuous functions converging to 0. Then the functions $f_{\alpha x}$ for fixed $x \in X$ (where $f_{\alpha x}(y) = f_\alpha(xy)$, see (2)) also form a decreasing net of (right) uniformly equicontinuous functions converging to 0. Thus, as LM is ϱ-continuous,

$$(24) \qquad (\overline{LM}f_\alpha)(x) = LMf_{\alpha x} \to 0 \quad \text{for every} \quad x \in X.$$

On the other hand, the (right) uniform equicontinuity of f_α implies that of $\bar{M}f_\alpha$, too, in view of (3), and as the net f_α is decreasing, the net $\bar{M}f_\alpha$ is also decreasing.

Hence follows that $h(y)=\lim(\overline{M}f_\alpha)(y)$ exists and $h\in U_r(X)$. Furthermore, $\overline{M}f_\alpha \geqq h$ imples (for arbitrary $L\in\mathscr{L}_r^+(X)$) that for every α

$$(25)\qquad (\bar{L}\overline{M}f_\alpha)(x)\geqq(\bar{L}h)(x)\quad\text{for every}\quad x\in X.$$

In view of the second relation of (4), from (24) and (25) follows that $\bar{L}h$ is identically 0; in view of the assumption $L\in\mathscr{M}_r^p(X)$, h must be identically 0 as well. In particular, $\lim Mf_\alpha=h(e)=0$, proving that M is ϱ-continuous.

LEMMA 3. *Suppose that the topological group X has the property*

(P) *to every $U\in\mathscr{V}$ there exists $V\in\mathscr{V}$ such that to every $x_0\in X$ there exists a finite set $\{z_1, z_2, \dots, z_m\}\subset X$ for which $\bigcup_{i=1}^m z_iU\subset Vx_0$.*

Then $\mathscr{M}_r^\varrho(X)\subset\mathscr{M}_r^p(X)$.

PROOF. Suppose that there exists a nonnegative $f\in U_r(X)$ which is not identically 0 while $(\bar{L}f)(x)=Lf_x=0$ for every $x\in X$. Without any loss of generality, we may assume that $f(e)=1$. Then, for some $U\in\mathscr{V}$, $f(x)>1/2$ if $x\in U$. Pick a $V\in\mathscr{V}$ satisfying (P). If L is ϱ-continuous, consider a function g with the properties (ii) in Lemma 1; then $0\leqq 1-g\leqq 1$, $1-g$ vanishes outside $\bigcup_{i=1}^n Vx_i$ and $L(1-g)>0$. From the assumption (P) follows the existence of a finite set $\{z_1, \dots, z_m\}\subset X$ such that $\bigcup_{i=1}^m z_iU\supset \supset\bigcup_{i=1}^n Vx_i$ implying

$$(26)\qquad \sum_{i=1}^m f(z_i^{-1}x)\geqq(1-g(x))\sum_{i=1}^m f(z_i^{-1}x)\geqq\frac{1}{2}(1-g(x));$$

thus, see (2), we have arrived at the contradiction

$$(27)\qquad \sum_{i=1}^m(\bar{L}f)(z_i^{-1})\geqq\frac{1}{2}L(1-g)>0,$$

completing the proof.

The lemmas 2 and 3 show that proposition 2 certainly applies to $\mathscr{M}=\mathscr{M}_r^\varrho(X)$ if X has the property (P), in particular, if X is Abelian or locally compact (or locally precompact).

The question whether a ϱ-continuous mean L always gives rise to a strictly positive operator $\bar{L}$, i.e. if $\mathscr{M}_r^\varrho(X)\subset\mathscr{M}_r^p(X)$ unconditionally, remains open. Of course, in the case this question were answered in the affirmative, it would be possible to assert, in general, that proposition 2 applies for $\mathscr{M}=\mathscr{M}_r^\varrho(X)$. Anyway, the choice $\mathscr{M}=\mathscr{M}_r^\varrho(X)\cap\mathscr{M}_r^p(X)$ is certainly feasible. In fact, as $LM\in\mathscr{M}_r^p(X)$ obviously implies $M\in\mathscr{M}_r^p(X)$ (even if $L\notin\mathscr{M}_r^p(X)$), the set $\mathscr{M}=\mathscr{M}_r^\varrho(X)\cap\mathscr{M}_r^p(X)$ surely fulfills the assumption of proposition 2, in view of lemma 2.

Unfortunately, proposition 2 alone does not seem sufficient to prove such desirable results as (with the notations of proposition 2)

CONJECTURE 1. If $M\in\mathscr{M}$ and $M'\in\mathscr{M}$ are two different cluster points of the net L_r^t, $t\succ r$ (with $r\in T$ fixed) then $M'=M\Delta_a$ for some $a\in X$ where Δ_a denotes the functional $\Delta_af=f(a)$ $(f\in V_r(X))$.

CONJECTURE 2. If the nets L_s^t, $t \succ s$ are convergent for every fixed $s \geqq r : L_s^t \to M_s$ and if $M_r = M \in \mathcal{M}$, then the net M_t $(t \succ r)$ is convergent as well: $M_t \to M_\infty$, and M_∞ is an idempotent: $M_\infty M_\infty = M_\infty$.

To prove these conjectures in the above generality, one would need some further information, first of all a characterization of the idempotents in $\mathcal{M}$.

Such a characterization is available for tight idempotent means, or, in other words, for idempotent K-regular Borel measures (cf. the paragraph containing (8)). A well known theorem proved first for compact X by WENDEL [10] then for locally compact X by HEYER [5] and in the form cited below by TORTRAT [11] says namely that a tight Borel measure μ on an arbitrary topological group X is idempotent if and only if $\mu = \varkappa_H$, the (normed) Haar measure on some compact subgroup H of X.

Using this fact and the well-known formula

$$S_{\mu\nu} = \overline{S_\mu S_\nu} \tag{28}$$

valid for the supports[5] of arbitrary τ-regular Borel measures on X it is easy to conclude from proposition 2 that conjectures 1 and 2 are true if $\mathcal{M}$ denotes the class of all tight means. Of course, one has to check that this class satisfies the condition on $\mathcal{M}$ of proposition 2.

LEMMA 4. *If $L \in \mathcal{M}_r(X)$ and $M \in \mathcal{M}_r(X)$ are such that both L and LM are tight then M is tight as well.*

PROOF. Let K_1 and K_2 be compact subsets of X such that if $f \in U_r(X)$, $0 \leqq f \leqq 1$ and f vanishes on K_1 or on K_2 then $Lf < \varepsilon_1$ or $LMf < \varepsilon_2$, respectively. Then if $f \in U_r(X)$, $0 \leqq f \leqq 1$ and $f(x) = 1$ for $x \in K_1$, we have $L(1-f) \leqq \varepsilon_1$, i.e. $Lf \geqq 1 - \varepsilon_1$. This immediately implies for $h \in U_r(X)$, $h \geqq 0$

$$Lh \geqq (1-\varepsilon_1)a \quad \text{if} \quad h(x) \geqq a \quad \text{for} \quad x \in K_1. \tag{29}$$

Let us suppose that $f \in U_r(X)$, $0 \leqq f \leqq 1$, and f vanishes on $K_1^{-1}K_2$; then

$$h(y) = \max_{x \in K_1} f(x^{-1}y) \tag{30}$$

vanishes on K_2.

We have to show that $h \in U_r(X)$; if $V \in \mathcal{V}$ is such that $y_1 y_2^{-1} \in V$ implies $|f(y_1) - f(y_2)| < \varepsilon$ and $V' \in \mathcal{V}$ is such that $V'x \subset xV$ if $x \in K_1$ (the existence of such V' follows from the compactness of K_1, see e.g. [1], lemma 3. 1) then for every $x \in K_1$

$$(x^{-1}y_1)(x^{-1}y_2)^{-1} = x^{-1}y_1 y_2^{-1} x \in x^{-1}V'x \subset V$$

whenever $y_1 y_2^{-1} \in V'$, implying $|f(x^{-1}y_1) - f(x^{-1}y_2)| < \varepsilon$ if $y_1 y_2^{-1} \in V'$, for every $x \in K_1$, thus $|h(y_1)| - h(y_2) < \varepsilon$, as well.

Now, since $h \in U_r(X)$, $0 \leqq h \leqq 1$, and h vanishes on K_2, we have by assumption $LMh < \varepsilon_2$. Hence follows in view of $LMh = L(\overline{M}h)$ and of (29) that there exists at least one $x_0 \in K_1$ for which

$$Mh_{x_0} = (\overline{M}h)(x_0) < \frac{\varepsilon_2}{1-\varepsilon_1} \qquad (x_0 \in K_1). \tag{31}$$

[5] The support S_μ of a measure μ is the smallest closed set with $\mu(S_\mu) = \mu(X)$; the existence of the support is an easy consequence of the definition of τ-regularity, see (6).

In view of (30) we have $h_{x_0}(y) = \max_{x \in K_1} f(x^{-1}x_0 y) \geqq f(y)$, thus (31) implies $Mf < \frac{\varepsilon_2}{1-\varepsilon_1}$, as well. Since $\frac{\varepsilon_2}{1-\varepsilon_1} = \varepsilon > 0$ is arbitrary, we have arrived at the conclusion that to every $\varepsilon > 0$ there exists a compact set K_ε ($K_\varepsilon = K_1^{-1}K_2$) such that the assumptions $f \in U_r(X)$, $0 \leqq f \leqq 1$, $f(x) = 0$ for $x \in K_\varepsilon$ imply $Mf < \varepsilon$, i.e. M is tight.

THEOREM 2. *Let T be a directed set without maximal element and let λ_s^t ($s, t \in T$, $s \prec t$) be a family of K-regular Borel probability measures on an arbitrary topological group X, satisfying*

$$(32) \qquad \lambda_r^s \lambda_s^t = \lambda_r^t \qquad (r \prec s \prec t).$$

Then for any two K-regular cluster points μ and μ' of the net λ_r^t, $t \succ r$ (with $r \in T$ fixed) there exists an $a \in X$ such that $\mu' = \mu\delta_a$, where δ_a denotes the point mass at a.

Furthermore, if the nets λ_s^t ($t \succ s$) are convergent[6] *for every fixed $s \succeq r$, and $\mu_r = \lim_t \lambda_r^t$ is K-regular, then so are $\mu_s = \lim_t \lambda_s^t$ ($s \succ r$), moreover, μ_s converges to the Haar measure $\varkappa_H$ on come sompact subgroup H of X.*

PROOF. Applying proposition 2 to the class of tight means with L_s^t corresponding to the measures λ_s^t, from (23) follows

$$(33) \qquad \mu' = \mu\nu, \qquad \nu\nu' = \varkappa_H$$

(having in mind that the means N, N' and M_∞ in (23) must be also tight thus N and N' correspond to K-regular measures ν and ν' while the idempotent M_∞ corresponds to a Haar measure $\varkappa_H$).

In view of (28) we have $S_\nu S_{\nu'} \subset H$ and hence $S_\nu \subset Ha$ for some $a \in X$, implying[7] $\varkappa_H \nu = \varkappa_H \delta_a$. Since $\mu = \mu\varkappa_H$, by the second relation in (22), $\mu' = \mu\delta_a$ follows now from the first relation of (33).

If the nets λ_s^t ($t \succ s$) are convergent for every fixed $s \succeq r$ and $\mu_r = \lim_t \lambda_r^t$ is K-regular, proposition 2 implies that the measures $\mu_s = \lim_t \lambda_s^t$, $s \succ r$, corresponding to the means $M_s = \lim_t L_s^t$ are also K-regular (see footnote 6.) We claim that every convergent subnet of the net M_s has the same limit.

In fact, if $M_{t(\alpha)} \to M_\infty$ and $M_{t'(\alpha)} \to M'_\infty$, the N and N' of (23) are equal to M_∞ and M'_∞, respectively (as now $M_s^r = M_s$ for every $s \succeq r$). As M_∞ and M'_∞ must correspond to Haar measures $\varkappa_H$ and $\varkappa_{H'}$, respectively, the second relation of (23) becomes $\varkappa_H \varkappa_{H'} = \varkappa_H$, implying $H' \subset H$. By symmetry, we also have $H \subset H'$, i.e. $M'_\infty = M_\infty$. However, since $\mathcal{M}_r(X)$ is compact, the fact that every convergent subnet of M_s has the same limit means that the net M_s is convergent; the proof is complete.

[6] By this it is enough to understand that $\int f(x)\lambda_s^t(dx)$ is convergent for every $f \in U_r(X)$; then the statement that $\mu_s = \lim \lambda_s^t$ is K-regular, is interpreted so that there exists a K-regular measure μ_s with $\int f(x)\mu_s(dx) = \lim_t \int f(x)\lambda_s^t(dx)$ for every $f \in U_r(X)$ (and then for every $f \in C(X)$ as well).

[7] Cf. (9), taking into account that $\int f(xy)\varkappa_H(dx) = \int f(xa)\varkappa_H(dx)$ if $y \in Ha$.

REMARK. Both statements of theorem 2 are well-known under some more restrictive conditions, see e.g. [1], where the method used here was first developped. The first statement is related to KLOSS' "convergence principle" [7] and its most general known formulation is apparently that of TORTRAT [12], where, however, a uniform tightness assumption is imposed. A particular case of the second statement (with T being the sequence of positive integers) has appeared in [1] as a step in the proof of the main theorem 3. 2, and in a similar role but under less restrictive conditions, in [9].

The novel feature of theorem 2 consists mainly in eliminating the condition of uniform tightness; this has been made possible by theorem 1 and lemma 4.

Of course, one expects that theorem 2 remains true also if the measures λ_s^t as well as the cluster points μ and μ' are supposed to be τ-regular only. Since the characterization of τ-regular idempotent measures has been obtained in [2] (corollary of theorem 4), and the relation (28) holds for arbitrary τ-regular Borel measures, this generalization of theorem 2 would be immediate, if one knew that the set of τ-regular means satisfies the condition on $\mathcal{M}$ of proposition 2. Unfortunately, I have not succeeded in proving this*, thus the problem, as well as the more general problem of finding less restrictive conditions under which the conjectures 1 and 2 are true, remains open.

Finally, I should like to point out that the results of § 2 remain true even for func-
such convolutions where the left factor is a functional on $U_r(X)$ and the right one a
tional on $U(Y)$, where Y is a "homogeneous space" and X is the group "acting on Y".

More exactly, let X be a topological group, let H be a subgroup of X and let $Y = X|H$ denote the class of left cosets of H endowed with the quotient topology; while Y is not a group in general (unless H is a normal subgroup of X), the multiplication of elements of Y by elements of X can be defined: if $x \in X$ and $y \in Y$ is a left coset of H ($y = y_0 H$, say) then xy is again a left coset of H (namely $xy = xy_0 H$). Clearly, the class of sets[8] $Vy \subset Y$, $V \in \mathcal{V}$ is a base for the neighborhoods of $y \in Y$; moreover $\mathcal{Y}$ can be considered as a uniform space, a base for the uniformity being the sets of pairs $\{(y_1, y_2): y_1 \in Vy_2\}$, $V \in \mathcal{V}$ (as $V = V^{-1}$ for $V \in \mathcal{V}$, $y_1 \in Vy_2$ is equivalent to $y_2 \in Vy_1$). Let $U(Y)$ be the Banach space of bounded and uniformly continuous (real valued) functions on Y; let $\mathcal{L}(Y)$ be the dual of $U(Y)$ i.e. the set of all bounded linear functionals on $U(Y)$, endowed with the weak* topology.

If $M \in \mathcal{L}(Y)$, the linear operator $\overline{M}$ defined by

$$(34) \qquad (\overline{M}f)(x) = Mf_x, \quad f_x(y) = f(xy), \quad (f \in U(Y), x \in X)$$

maps $U(Y)$ into $U_r(X)$ and $\overline{M}f$ is (right) uniformly equicontinuous with f in the sense that if $V \in \mathcal{V}$ is such that $y_2 = Vy_1$ implies $|f(y_1) - f(y_2)| < \varepsilon$ then $x_1 x_2^{-1} \in V$ implies $|(\overline{M}f)(x_1) - (\overline{M}f)(x_2)| \leqq \|L\| \varepsilon$.

Thus, if $L \in \mathcal{L}_r(X)$ and $M \in \mathcal{L}(Y)$, the convolution $LM \in \mathcal{L}(Y)$ can be defined as

$$(35) \qquad MNf = M(\overline{N}f) \qquad (f \in U(Y)).$$

One can prove quite similarly to theorem 1 that if $L_\alpha \to L$ in $\mathcal{L}_r^+(X)$ and $M_\alpha \to M$ in $\mathcal{L}(Y)$ then $L_\alpha M_\alpha \to LM$ (in $\mathcal{L}(Y)$) provided that $L \in \mathcal{L}_r^+(X)$ is ϱ-continuous.

[8] where $Vy = \{vy : v \in V\}$

* Added in prof: A. Tortrat has given a counterexamle (private communication).

REFERENCES

[1] CSISZÁR, I.: On infinite products of random elements and infinite convolutions of probability distributions on locally compact groups. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete,* **5** (1966), 279—295.
[2] CSISZÁR, I.: Some problems concerning measures on topological spaces and convolutions of measures on topological groups. *Les Probabilités sur les Structures Algébriques*, Clermont-Ferrand, 1969, 75—96. Colloques Internationaux du CNRS, Paris, 1970.
[3] GRENANDER, U.: *Probabilities on Algebraic Structures.* Almquist and Wiksell, Stockholm, 1963.
[4] HEWITT, E.—ROSS, K. A.: *Abstract Harmonic Analysis.* Springer, 1963.
[5] HEYER, H.: *Untersuchung zur Theorie der Wahrscheinlichkeitsverteilungen auf lokalkompakten Gruppen.* Dissertation, Hamburg, 1963.
[6] KELLEY, J.: *General Topology.* Van Nostrand, Princeton, 1955.
[7] Клосс, Б. М.: О вероятностных распределениях на бикомпактных топологических группах. *Теория Вероятностей и ее Применения* **41** (1959) 255—290.
[8] NEVEU, J.: *Bases mathématiques du Calcul des Probabilités.* Masson, Paris, 1964.
[9] Сазонов, В. В.—Тутубалин, В. Н.: Распределения вероятностей на топологических группах. *Теория Вероятностей и ее Применения,* **11** (1966) 3—55.
[10] WENDEL, J. G.: Haar measure and the semi-group of measures on a compact group. *Proceedings of the American Mathematical Society.* **5** (1954), 923—929.
[11] TORTRAT, A.: Lois de probabilité sur un espace topologique complètement régulier et produits infinis â termes indépendants dans un groupe topologique. *Annales de l'Institut Henri Poincaré,* **1** (1965), 217—237.
[12] TORTRAT, A.: Lois tendues et convolutions dénombrables dans un groupe topologique X. *Annales de l'Institut Henri Poincaré*, **2** (1966), 279—298.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received January 21, 1970.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 41—48.

THE SERIES QUEUE $M/G/1 \to /M/1$ WITH FINITE WAITING ROOM IN THE FIRST STAGE

by

C. PEARCE

1. Introduction

Although the problem of ascertaining queue length distribution in an unrestricted $M/G/1 \to \ldots/M/1$ tandem queue is still unsolved, a number of papers treat of this problem in the practically important case when one or other of the service points has finite waiting room (see AVI-ITZHAK and YADIN [1], NEUTS [4], PRABHU [5] and SUZUKI [6] for the queue with a restricted second stage and ÇINLAR [2] for a restricted first stage). The analysis tends to become involved and is usually based on our being able to regard the restricted queue as a semi-Markov process.

Our purpose is to give a simple alternative approach based on the imbedded chain method for single stage queues. We develop this for the queue with restricted waiting room at the first stage.

In section three we extend our results to cover Erlangian services at the second station and in section five we prove that the equilibrium queue length distribution in the unrestricted $M/G/1 \to \ldots/M/1$ queue can be obtained as the limit of distributions in the corresponding finite systems as the size of the finite waiting room increases without bound.

We denote by λ, μ respectively the parameters of the Poisson input stream and the service time process in the second stage and by $F(\cdot)$ the probability distribution function for service times at the first server. The maximum number of customers which the first stage can tolerate is $k<\infty$. So long as the first stage contains k customers further arrivals are lost. The system is said to be in state (i,j) at any instant if the first stage contains i customers and the second j.

2. The equilibrium joint queue length in the restricted queue $M/G/1 \to \ldots/M/1$

Consider the state (i,j) of the system at departure instants of the first stage, where we take i as the number of customers left behind in the first stage by the departure and j the number he finds in the second stage.

It follows from Feller's theory of recurrent events that to demonstrate the existence of a joint limiting equilibrium distribution of queue length independent of the initial state, it suffices to find a non-trivial solution with $\sum_{i,j} |\pi_{ij}| < \infty$ for the equations

$$\pi_{i0} = \sum_{r=0}^{i+1} \sum_{j=0}^{\infty} \pi_{rj} \sum_{m=j+1}^{\infty} p_m(r,i), \qquad 0 \leqq i < k-1, \tag{2.1}$$

$$\pi_{ij} = \sum_{r=0}^{i+1} \sum_{m=0}^{\infty} \pi_{r,j-1+m} p_m(r,i), \qquad 0 \leqq i < k-1,\ j>0, \tag{2.2}$$

$$\pi_{k-1,0} = \sum_{r=0}^{k-1} \sum_{j=0}^{\infty} \pi_{rj} \sum_{m=j+1}^{\infty} p_m(r, i), \tag{2. 3}$$

$$\pi_{k-1,j} = \sum_{r=0}^{k-1} \sum_{m=0}^{\infty} \pi_{r,j-1+m} p_m(r, k-1), \qquad j > 0, \tag{2. 4}$$

where the transition probabilities $p_m(r, s)$ are

$$p_m(r, s) = \tag{2. 5}$$

$$= \begin{cases} \int\limits_0^\infty \exp(-(\lambda+\mu)x)(\lambda x)^{s+1-r}/(s+1-r)!(\mu x)^m/m!\, dF(x), & 0 < r \leqq s+1 < k \\ \int\limits_0^\infty \exp(-(\lambda+\mu)x) \sum\limits_{j=s}^{\infty} (\lambda x)^{j+1-r}/(j+1-r)!(\mu x)^m/m!\, dF(x), & 0 < r \leqq s+1 = k \\ \int\limits_0^\infty \int\limits_0^\infty \exp(-(\lambda+\mu)(x+y))(\lambda x)^s/s!(\mu(x+\lambda))^m/m!\lambda\, dy\, dF(x), & 0 = r < s+1 < k, \\ \int\limits_0^\infty \int\limits_0^\infty \exp(-(\lambda+\mu)(x+y)) \sum\limits_{j=k-1}^{\infty} (\lambda x)^j/j!(\mu(x+y))^m/m!\lambda\, dy\, dF(x), & r = 0,\ s = k-1, \\ 0, \text{ otherwise.} \end{cases}$$

The π_{ij} normalized to sum unity then provide that equilibrium distribution.

We first seek a solution for (2. 2) and (2. 4) of the form

$$\pi_{ij} = a_i T^j, \qquad j \geqq 0. \tag{2. 6}$$

If $|T| < 1$ this solution will satisfy the requirement $\sum_j |\pi_{ij}| < \infty$ for each i. Under the condition that the traffic intensity for the second stage is less than unity we shall find there are k (in general distinct) such values of T. It then only remains to verify that by taking a suitable linear combination of these basic solutions we can also satisfy (2. 1) and (2. 3).

Substitution of (2. 6) into (2. 2) and use of (2. 5) gives, for all $j > 0$,

$$a_i T = \sum_{r=0}^{i} a_{i+1-r} \psi_r + \lambda(\lambda+\mu(1-T))^{-1} a_0 \psi_i \qquad 0 \leqq i < k-1, \tag{2. 7}$$

where

$$\psi_i = \int\limits_0^\infty \exp(-\lambda x)(\lambda x)^i/i! \exp(-\mu x(1-T))\, dF(x).$$

Similarly (2. 4) yields

$$a_{k-1} T = \sum_{r=1}^{k-1} a_r \sum_{s=k-r}^{\infty} \psi_s + \lambda(\lambda+\mu(1-T))^{-1} a_0 \sum_{r=k-1}^{\infty} \psi_r. \tag{2. 8}$$

The k simultaneous equations (2. 7), (2. 8) in the a_i possess a non-trivial solution if and only if T satisfies the determinantal equation

$$(2.9)\quad \begin{vmatrix} \lambda(\lambda+\mu(1-T))^{-1}\psi_0-T & \psi_0 & 0 & \dots & 0 \\ \lambda(\lambda+\mu(1-T))^{-1}\psi_1 & \psi_1-T & \psi_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda(\lambda+\mu(1-T))^{-1}\psi_{k-2} & \psi_{k-2} & \psi_{k-1} & \cdots & \psi_0 \\ \lambda(\lambda+\mu(1-T))^{-1}\sum_{r=k-1}^{\infty}\psi_r & \sum_{r=k-1}^{\infty}\psi_r & \sum_{r=k-2}^{\infty}\psi_r & \cdots & \sum_{r=1}^{\infty}\psi_r-T \end{vmatrix} = 0.$$

Removal of the obvious solution $T=0$ leads to

$$(2.10)\quad \lambda\psi_0(\lambda+\mu(1-T))^{-1}D_{k-2}+D_{k-1}=0,$$

where

$$(2.11)\quad D_p=\begin{vmatrix} \psi_1-T & \psi_0 & \cdots & 0 \\ \psi_2 & \psi_1-T & \cdots & 0 \\ \vdots & \vdots & & \\ \sum_{r=p}^{\infty}\psi_r & \sum_{r=p-1}^{\infty}\psi_r & \cdots & \sum_{r=1}^{\infty}\psi_r-T \end{vmatrix},\qquad p\geqq 2$$

$$D_1=\sum_{r=1}^{\infty}\psi_r-T,$$

$$D_0=1.$$

Expansion of the determinant in (2. 11) gives a recurrence relation for the D_p, from which their generating function $\sum_{p=1}^{\infty}D_pz^p$ is found to be

$$(2.12)\quad \sum_{p=1}^{\infty}D_pz^p=[Tz+\psi_0^{-1}\Psi(-z\psi_0)]^{-1}[-Tz+z(1+z\psi_0^{-1})(\Psi(1)-\Psi(-z\psi_0))],$$

where

$$(2.13)\quad \begin{aligned}\Psi(t)&=\sum_{s=0}^{\infty}\psi_st^s\\ &=\int_0^{\infty}\exp[-\lambda x(1-t)-\mu x(1-T)]\,dF(x).\end{aligned}$$

We use (2. 12) to rewrite (2. 10) as

$$(2.14)\quad \begin{gathered}\lambda(\lambda+\mu(1-T))^{-1}\cdot\ \text{coeff. of } z^{k-2} \text{ in p. s. expansion of}\\ z(1-z)^{-1}[1+(T-\Psi(1))/(-Tz+\Psi(z))]\\ -\text{coeff. of } z^{k-1} \text{ in p. s. expansion of}\\ z(1-z)^{-1}[1+(T-\Psi(1))/(-Tz+\Psi(1))]=0.\end{gathered}$$

By virtue of Rouche's theorem and (2. 13), (2. 14) considered as an equation in T can be shown to possess $k-1$ roots inside the unit circle in the complex plane provided that the condition

$$\mu^{-1} < \int_0^\infty x\, dF(x) + \lambda^{-1}/(\text{coeff. of } z^{k-2} \text{ in p.s. expansion of } \left[-z+\int_0^\infty \exp(-\lambda x(1-z))\, dF(x)\right]^{-1}) \tag{2.15}$$

is satisfied.

(2. 15) is, however, simply the condition that the traffic intensity for the second stage be less than unity. This can be seen as follows:

For the first stage the equilibrium queue length distribution values q_j $(0 \leqq j < k)$ taken on the imbedded chain of departure points satisfy the recursive relations

$$q_j = \sum_{r=0}^{j+1} q_r p(r, j), \tag{2.16}$$

wher :

$$p(r, s) = \begin{cases} \int_0^\infty \exp(-\lambda x)(\lambda x)^{s+1-r}/(s+1-r)!\, dF(x), & 0 < r \leqq s+1, \\ \int_0^\infty \exp(-\lambda x)(\lambda x)^s/s!\, dF(x), & s \geqq 0, \\ 0 \text{ otherwise.} \end{cases}$$

By defining q_j for $j \geqq k$ recursively through extending (2. 16) to hold for all positive integral j we find on taking generating functions that

$$q_j = q_0 \cdot \text{coeff. of } z^j \text{ in formal p.s. expansion of} \tag{2.17}$$

$$(1-z)\Big/\left(1-z\left[\int_0^\infty \exp(-\lambda x(1-z))\, dF(x)\right]^{-1}\right), \qquad 0 \leqq j < k-1,$$

and the q_j are fixed through the normalization

$$\sum_{j=0}^{k-1} q_j = 1. \tag{2.18}$$

Thus the mean inter-departure time d for the first stage is given by

$$d = \text{mean service time in first stage} + \lambda^{-1} \cdot q_0 =$$

$$= \int_0^\infty x\, dF(x) + \lambda^{-1}\Big/\left(\text{coeff. of } z^{k-2} \text{ in p.s. expansion of } \left[-z+\int_0^\infty \exp(-\lambda x(1-z))\, dF(x)\right]^{-1}\right).$$

Hence by comparison with (2. 15), we have that the traffic intensity condition $\mu^{-1}<d$ ensures that (2. 2), (2. 4) have $k-1$ solutions

$$\pi_{ij}=a_i(p)T_p^j, \qquad 1\leqq p\leqq k-1.$$

A further solution

$$\pi_{ij}=a_i(k)\delta_{j,0}$$

is provided by the root $T=0$. We note that (2. 2) implies

$$a_i(k) = -\lambda(\lambda+\mu)^{-1}a_0(k)\delta_{i1}.$$

As the $a_i(p)$ is determined only to a (common) scalar multiplier for each p

$$\pi_{ij} = \sum_{p=1}^{k-1} s(p)a_i(p)T_p^j+\delta_{j,0}s(k)a_0(k)[\delta_{i,0}-\lambda(\lambda+\mu)^{-1}\delta_{i,1}]$$

will also satisfy (2. 2), (2. 4) for any constants $s(p)$. For suitable choices of the k quantities $s(p)$, (2. 1), (2. 3) can also be satisfied. (2. 1), (2. 3) impose k simultaneous homogeneous linear conditions on the $s(p)$, but one of these derives from the others and (2. 2), (2. 4), since the relation obtained by adding (2. 1)—(2. 4) is an identity.

The working in practical calculations is much simplified by use of

(2. 19) $$\pi_{i\cdot} = \sum_{r=0}^{i+1} \pi_{r\cdot}p(r,i), \qquad 0\leqq i<k-1$$

(2. 20) $$\pi_{k-1\cdot} = \sum_{r=0}^{k-1} \pi_{r\cdot} \sum_{s=k-1}^{\infty} p(r,s),$$

where

$$\pi_{j\cdot} = \sum_{m=0}^{\infty} \pi_{jm}, \qquad 0\leqq j\leqq k-1.$$

Equations (2. 19) are simply the recurrence relations (2. 16) for the imbedded chain equilibrium distribution of queue length in the first stage, (2. 20) can be verified by addition to be deducible from (2. 19).

3. Extension to Erlangian services in the second stage

Suppose now the services in the second stage are Erlangian of order m. We adopt the standard simple device of replacing these services by negative exponential services with the same parameter and simultaneously replacing the customers by batches of m units which arrive and are served collectively at the first stage but receive individual services at the second stage.

If $\{\pi_{ij},\ 0\leqq i\leqq k-1,\ j\geqq 0\}$ denotes the joint equilibrium probability (when it exists) that a batch leaves i batches in the first stage and finds j units in the second stage, we have analogously to (2. 1)—(2. 4)

$$\pi_{i0} = \sum_{r=0}^{i+1} \sum_{j=0}^{\infty} \pi_{rj} \sum_{n=j+m}^{\infty} p_n(r, i), \qquad 0 \leqq i < k-1, \tag{3. 1}$$

$$\pi_{ij} = \sum_{r=0}^{i+1} \sum_{n=0}^{\infty} \pi_{rn} p_{n+m-j}(r, i), \qquad 0 \leqq i < k-1, \tag{3. 2}$$

$$\pi_{ij} = \sum_{r=0}^{i+1} \sum_{n=0}^{\infty} \pi_{r,j-m+n} p_n(r, i), \qquad 0 \leqq i < k-1,\ j \geqq m, \tag{3. 3}$$

$$\pi_{k-1,0} = \sum_{r=0}^{k-1} \sum_{j=0}^{\infty} \pi_{rj} \sum_{n=j+m}^{\infty} p_n(r, k-1), \tag{3. 4}$$

$$\pi_{k-1,j} = \sum_{r=0}^{k-1} \sum_{n=0}^{\infty} \pi_{rn} p_{n+m-j}(r, k-1), \qquad 1 \leqq j \leqq m-1, \tag{3. 5}$$

$$\pi_{k-1,j} = \sum_{r=0}^{k-1} \sum_{n=0}^{\infty} \pi_{r,j-m+n} p_n(r, k-1), \qquad j \geqq m \tag{3. 6}$$

where the p_r are defined as before. When the traffic intensity condition is satisfied for the second stage it turns out that (3. 3), (3. 6) possess geometric solutions

$$\pi_{ij} = a_i(p) T_p^j \qquad 0 \leqq i \leqq k-1,\ j \geqq 0$$

for $(k-1)m$ values T_p, $1 \leqq p \leqq (k-1)m$. Corresponding to the root $T=0$ removed from (2. 9) there is an m-fold root $T=0$. These roots lead to a solution to equations (3. 1)—(3. 6) inclusive of the form

$$\pi_{ij} = \sum_{p=1}^{(k-1)m} s(p) a_i(p) T_p^j + \begin{cases} b_j & i = 0,\ 0 \leqq j < m \\ -\lambda(\lambda+\mu)^{-1} b_j, & i = 1,\ 0 \leqq j < m. \end{cases}$$

4. Repeated roots

In sections two and three we have assumed that, apart from the root $T=0$ of multiplicity m in section three, the roots T_p are distinct. It seems likely that this is, in general, true, but as with the single server queue no simple proof for such a result is apparent. Our methods can, however, be extended to cover such a contingency along the lines of WISHART [7] for the $G/E_m/1$ queue.

In the special case of exponential services (with parameter ν, say) in the first stage, the root $T=0$ is multiple even for $m=1$. This can be seen as follows:

Directly from its definition

$$\psi_i = \nu\lambda^i/(\lambda+\mu(1-T)+\nu)^{i+1}, \qquad i \geqq 0,$$

from which an expansion of the determinant D_p yields

$$D_p = -TD_{p-1} - T\lambda\nu(\lambda+\mu(1-T)+\nu)^{-2} D_{p-2}, \qquad p \geqq 2.$$

By an easy induction D_p has a factor $T^{[p/2]}$ so that by (2. 10) the determinant in (2. 9) has a factor of $T^{[k/2]}$.

5. The queue length distribution when the waiting room is unrestricted

It would be interesting to be able to extend the foregoing results to unrestricted queues, but this seems quite difficult to do. We can, however, prove the following result:

THEOREM *Let $\pi(i,j;m)$ ($\pi(i,j)$) be the equilibrium queue length probabilities on the imbedded chain for the system with first stage waiting room of size m (unrestricted). We assume the traffic intensity condition*

$$\lambda^{-1} > \max\left(\mu^{-1}, \int_0^\infty x\,dF(x)\right) \tag{5.1}$$

for the unrestricted queue. Then for all $i, j \geqq 0$,

$$\pi(i,j;m) \to \pi(i,j) \quad as \quad m \to \infty.$$

PROOF. We first label the joint queue lengths in the two stages in the case of an unrestricted first stage by the non-negative integers. This can be done in a 1-1 manner by denoting the successive points in the path sketched below 0, 1, 2, ... We shall use the same label to refer to a given joint queue size in the case when the waiting room in the first stage is restricted.

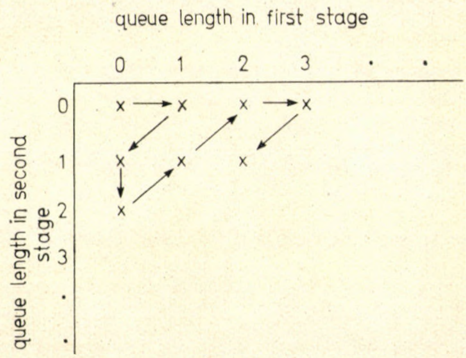


Fig. 1

Take any two (fixed) joint queue lengths 'i', 'j' in the unrestricted queue. If the size m of the waiting room in the first stage of the corresponding restricted queue is sufficiently large, 'i' and 'j' will be states of this queue also, and the one-step transition probabilities $P_{ij}(m)$, P_{ij} will have the same values. In particular $\lim_{m\to\infty} P_{ij}(m)$ exists for every ordered pair (i,j) and has the value P_{ij}.

Let us denote the limiting (joint) queue length distributions in the unrestricted queue and the queue with finite waiting room m by $\{\pi_i\}$, $\{\pi_i(m)\}$, respectively. Then by a standard result on Markov chains,

$$\pi_i(m) = \sum_{j=0}^{\infty} \pi_j(m) P_{ji}, \qquad i \geqq 0, \tag{5.2}$$

for each $m \geqq 1$.

Under our ergodicity assumption (5.1) each of the series queues under consideration will be positive recurrent and so the limiting probability that a customer leaving the first stage empty also finds the second stage empty will be non-zero. (5.2) may therefore be written as

$$v_i(m) = \sum_{j=0}^{\infty} v_j(m) P_{ji}(m), \qquad i \geqq 0,$$

for each m, where $v_i(m)$ is defined as $\pi_i(m)/\pi_0(m)$. Letting $m \to \infty$ we find that

$$\underline{\lim}\, v_i(m) \geqq \sum_{j=0}^{\infty} \underline{\lim}\, v_j(m) P_{ji}, \qquad i \geqq 0.$$

We note that

$$\underline{\lim}\, (v_i(m)) \geqq 0, \qquad i \geqq 1$$

and that

$$\underline{\lim}\, (v_0(m)) = 1.$$

But by a well known result in the theory of Markov chains (see, for example, KARLIN [3]), the system

$$(5.\ 3) \qquad \begin{cases} v_i \geqq \sum_{j=0}^{\infty} v_j P_{ji}, & i \geqq 0 \\ v_0 = 1, & \\ v_i \geqq 0, & i \geqq 1 \end{cases}$$

has a unique solution if the Markov chain with 1-step transition probabilities (P_{ij}) is recurrent and irreducible.

As the quantities $v_i = \pi_i/\pi_0$, $i \geqq 0$, obviously satisfy the relations (5. 3) with equality, we thus have

$$\underline{\lim}\, (v_i(m)) = v_i, \qquad i \geqq 0.$$

By symmetry we shall also have

$$\overline{\lim}\, (v_i(m)) = v_i, \qquad i \geqq 0,$$

so that $\lim_{m \to \infty} (v_i(m))$ exists and has value v_i for each $i \geqq 0$. This establishes that

$$\pi_i(m) \to \pi_i \quad \text{for each} \quad i \geqq 0,$$

which gives the required result.

REFERENCES

[1] AVI—ITZHAK, B., and YADIN, M.: A sequence of two servers with no intermediate queue, *Management Sci.*, **11** (1965), 553—64.
[2] ÇINLAR, E.: Queues with semi-Markovian arrivals, *J. Appl. Prob.*, **4** (1967), 365—79.
[3] KARLIN, S.: *A first course in stochastic processes*, Academic Press (New York and London) (1966).
[4] NEUTS, M. F.: Two queues in series with a finite, intermediate waiting room, *J. Appl. Prob.*, **5** (1968), 123—42.
[5] PRABHU, N. U.: Transient behaviour of a tandem queue, *Management Sci.*, **13** (1966), 631—9.
[6] SUZUKI, T.: On a tandem queue with blocking, *J. Oper. Res. Soc. Japan*, **6** (1964), 137—57.
[7] WISHART, D. M. G.: A queueing system with χ^2 service-time distribution, *Ann. Math. Statist.*, **27** (1956), 768—79.

Department of Probability and Statistics, The University, Sheffield, England.
Now at Department of Mathematics, University of Adelaide, Australia
(Received: April 20, 1967. Revised January 24, 1970.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 49—53.

EINIGE METRISCHE ERGEBNISSE IN DER THEORIE DER CANTORSCHEN REIHEN UND BAIRESCHE KATEGORIEN VON MENGEN

von

T. ŠALÁT

Es sei $x = \sum_{n=1}^{\infty} \frac{C_n(x)}{g^n}$ die g-adische Entwicklung der Zahl $x \in \langle 0, 1)$ (also g ist ganz, $g > 1$, $C_n(x)$ sind ganze Zahlen, $0 \leqq C_n(x) < g$ $(n = 1, 2, 3, \ldots)$ und für unendlich viele n ist $C_n(x) < g-1$). Die Zahl $x \in \langle 0, 1)$ nennt man g-adisch normal, wenn für jede $r = 0, 1, \ldots, g-1$ die Beziehung

$$\lim_{n \to \infty} \frac{N_n(r, x)}{n} = \frac{1}{g}$$

gilt, dabei $N_n(r, x) = \sum_{k \leqq n, C_k(x) = r} 1$. Bekanntlich sind fast alle Zahlen $x \in \langle 0, 1)$ g-adisch normal (siehe [1] S. 125—128). Also die Menge N_g aller g-adisch normalen Zahlen des Intervalles $\langle 0, 1)$ ist vom metrischen Standpunkt eine reiche Menge. Eine ganz andere Situation tritt ein, wenn wir die Menge N_g vom Standpunkt der Baireschen Kategorien von Mengen in dem metrischen Raum $\langle 0, 1)$ (mit euklidischer Metrik) studieren. Dann kann man beweisen, dass die Menge N_g nur eine Menge von erster Kategrie in $\langle 0, 1)$ ist (siehe [2]).

A. Rényi hat gezeigt, dass der Begriff der normalen Zahl auch auf die Cantorschen Reihen erweitert werden kann. Es sei $\{q_k\}_{k=1}^{\infty}$ eine Folge von natürlichen Zahlen, $q_k > 1$ $(k = 1, 2, 3, \ldots)$. Es sei

$$x = \sum_{k=1}^{\infty} \frac{\varepsilon_k(x)}{q_1 q_2 \cdots q_k} \tag{1}$$

die Cantorsche Reihe der Zahl $x \in \langle 0, 1)$, also $\varepsilon_k(x)$ $(k = 1, 2, \ldots)$ sind ganze Zahlen, sogenannte Ziffern von x, $0 \leqq \varepsilon_k(x) < q_k$ $(k = 1, 2, 3, \ldots)$ und für unendlich viele k ist $\varepsilon_k(x) < q_k - 1$ (siehe [3] S. 113). Setzen wir

$$N_n(r, x) = \sum_{k \leqq n,\, \varepsilon_k(x) = r} 1.$$

Die Zahl x nennt man normal in bezug auf die „Grundfolge" $\{q_k\}_{k=1}^{\infty}$, $\sum_{k=1}^{\infty} 1/q_k = +\infty$, wenn für jede ganze Zahl $r \geqq 0$, für welche $s_n(r) = \sum_{k \leqq n, r < q_k} 1/q_k \to +\infty$ ist, $\lim_{n \to \infty} \frac{N_n(r, x)}{s_n(r)} = 1$ gilt. Aus den Arbeiten [4], [5] folgt, dass (im Falle $\sum_{k=1}^{\infty} 1/q_k = +\infty$) fast alle Zahlen $x \in \langle 0, 1)$ in bezug auf $\{q_k\}_{k=1}^{\infty}$ normal sind. In der Arbeit [6] ist ein topologisches Ergebnis bewiesen, aus welchem folgt, dass (im Falle $\sum_{k=1}^{\infty} 1/q_k = +\infty$)

die Menge $N(q_1, q_2, \ldots)$ aller in bezug auf $\{q_k\}_{k=1}^{\infty}$ normalen Zahlen eine Menge von erster Kategorie in $\langle 0, 1)$ ist.

Im Zusammenhang mit den erwähnten Ergebnissen wollen wir in dieser Arbeit noch einige andere Fragen über die Verteilung der Ziffern in Cantorschen Reihen vom Standpunkt der Baireschen Kategorien von Mengen studieren.

Definitionen und Bezeichnungen

1. $\{a_n\}_n'$ bezeichnet die Menge aller Häufungswerte der Folge $\{a_n\}_{n=1}^{\infty}$.
2. $|M|$ bezeichnet das Lebesguesche Mass der Menge M.
3. Wenn A eine Menge der natürlichen Zahlen ist, dann setzen wir $A(n) = \sum_{a \leq n, a \in A} 1$. Die Zahl $\delta(A) = \lim_{n \to \infty} \frac{A(n)}{n}$, sobald sie existiert, nennt man die asymptotische Dichte der Menge A.
4. Die Intervalle

$$i_n^{(k)} = \left\langle \frac{k}{q_1 q_2 \cdots q_n}, \frac{k+1}{q_1 q_2 \cdots q_n} \right) \qquad (k = 0, 1, \ldots, q_1 \cdots q_n - 1)$$

nennem wir die Intervalle der n-ten Ordnung. Offenbar $\langle 0, 1) = \bigcup_{k=0}^{q_1 \cdots q_n - 1} i_n^{(k)}$. Weiter werden wir kurz sagen, dass das Intervall $i_n^{(k)}$ zur Folge $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ gehört, wenn

$$\frac{k}{q_1 q_2 \cdots q_n} = \frac{\varepsilon_1}{q_1} + \frac{\varepsilon_2}{q_1 q_2} + \ldots + \frac{\varepsilon_n}{q_1 q_2 \cdots q_n}$$

ist, wo $0 \leq \varepsilon_j < q_j$ $(j = 1, 2, \ldots, n)$, ε_j $(j = 1, 2, \ldots, n)$ ganze Zahlen sind. Offenbar gehört die Zahl x (siehe (1)) zum Intervall $i_n^{(k)}$ dann und nur dann, wenn $\varepsilon_j(x) = \varepsilon_j$ $(j = 1, 2, \ldots, n)$.

5. T bezeichnet die Menge aller Endpunkte aller Intervalle $i_k^{(k)}$, $n = 1, 2, 3, \ldots$, $k = 0, 1, \ldots, q_1 q_2 \ldots q_n - 1$. Weiter setzen wir $X = \langle 0, 1) - T$. Im weiteren betrachten wir $\langle 0, 1)$ (und auch X) als einen metrischen Raum mit der euklidischen Metrik.

Im Zusammenhang mit dem erwähnten topologischen Ergebnis der Arbeit [6] beweisen wir folgende zwei Sätze.

SATZ 1. *Es sei $q_k \to +\infty$. Dann gilt für alle $x \in \langle 0, 1)$ bis auf eine Menge von erster Kategorie in $\langle 0, 1)$ die Beziehung*

$$(*) \qquad \left\{ \frac{N_n(r, x)}{n} \right\}_n' = \langle 0, 1 \rangle,$$

$(r = 0, 1, 2, 3, \ldots)$.

BEWEIS. Es sei $\zeta \in (0, 1)$, $r \geq 0$, k, s seien natürliche Zahlen. Bezeichen wir mit $M(\zeta, r, k, s)$ die Menge aller

$x \in X$, für welche $\left| \frac{N_s(r, x)}{s} - \zeta \right| < \frac{1}{k}$ ist. Setzen wir

$$(2) \qquad M(\zeta, r) = \bigcap_{k=1}^{\infty} \bigcap_{p=1}^{\infty} \bigcup_{s=p}^{\infty} M(\zeta, r, k, s).$$

$M(\zeta, r)$ ist ersichtlich die Menge aller derjenigen $x \in X$, für welche $\zeta \in \left\{\frac{N_n(r, x)}{n}\right\}_n'$ ist. Erwägen wir, dass $M(\zeta, r, k, s)$ gleich der Vereinigungsmenge einer endlichen Anzahl der Mengen der Form $X \cap i_s^{(l)}$ ist, also $M(\zeta, r, k, s)$ eine in X offene Menge ist. Aus (2) folgt, dass $M(\zeta, r)$ eine G_δ-Menge in X ist.

Wir zeigen, dass $M(\zeta, r)$ in X dicht ist. Es genügt zu beweisen, dass für jedes Intervall $(a, b) \subset \langle 0, 1)$

$$M(\zeta, r) \cap [(a, b) \cap X] \neq \emptyset \tag{3}$$

ist. Es sei also $(a, b) \subset \langle 0, 1)$. Wählen wir l so, dass $\frac{1}{q_1 q_2 \dots q_l} < \frac{b-a}{2}$ und $q_m > r+1$ für jedes $m \geqq l$. Die Existenz einer solchen Zahl l folgt aus der Bedingung $q_n \to \infty$. Dann existiert ein Intervall $i_s^{(p)}$ s-ter Ordnung derart, dass $s \geqq l$, $i_s^{(p)} \subset (a, b)$. Es gehöre $i_s^{(p)}$ zur Folge $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_s$. Es sei $A = \{m_1 < m_2 < m_3 < \dots\}$ eine Menge von natürlichen Zahlen mit $m_1 > s$ und $\delta(A) = \zeta$ (siehe [9] S. 194—195). Definieren wir die Zahl

$$x_0 = \sum_{j=1}^{\infty} \frac{\varepsilon_j(x_0)}{q_1 q_2 \dots q_j} \tag{4}$$

folgendermassen:

$$\varepsilon_j(x_0) = \varepsilon_j \quad \text{für} \quad j = 1, 2, \dots, s; \tag{5}$$

$$\varepsilon_{m_j}(x_0) = r \quad \text{für} \quad j = 1, 2, 3, \dots; \tag{6}$$

$$\varepsilon_j(x_0) = r+1 \quad \text{für} \quad j > s, \quad j \neq m_k \quad (k = 1, 2, 3, \dots). \tag{7}$$

Dann ist offenbar (4) die Cantorsche Reihe der Zahl x_0, $x_0 \in \langle 0, 1)$ und da alle Zahlen der Menge T endliche Cantorsche Entwicklungen haben, ist $x_0 \in X$. Aus (5) folgt $x_0 \in i_s^{(p)} \subset (a, b)$, also $x_0 \in (a, b) \cap X$. Weiter mit Rücksicht auf (6), (7) und auf die Wahl der Menge A ist $x_0 \in M(\zeta, r)$. Also $x_0 \in M(\zeta, r) \cap [(a, b) \cap X]$ und so gilt (3).

Da die Menge $M(\zeta, r)$ eine dichte G_δ-Menge in X ist, ist diese Menge residual in X (siehe [7] S. 49). Dann ist auch $M(\zeta) = \bigcap_{r=0}^{\infty} M(\zeta, r)$ eine in X residuale Menge und wenn R die Menge aller rationalen Zahlen des Intervalles $(0, 1)$ bedeutet, dann ist $M = \bigcap_{\zeta \in R} M(\zeta)$ infolge der Abzählbarkeit der Menge R eine in X residuale Menge und infolge der Abzählbarkeit der Menge T auch eine in $\langle 0, 1)$ residuale Menge. M ist aber ersichtlich die Menge aller derjenigen $x \in \langle 0, 1)$, für welche (*) gilt. Damit ist der Beweis des Satzes beendet.

SATZ 2. *Es sei* $\limsup_{k \to \infty} q_k = +\infty$. *Dann gilt für alle* $x \in \langle 0, 1)$ *bis auf eine Menge von erster Kategorie in* $\langle 0, 1)$ $\{\varepsilon_n(x)\}_n' \supset \{0, 1, 2, \dots, n, \dots\}$.

BEWEIS. Setzen wir bei natürlichem n und ganzem $r \geqq 0$

$$H(n, r) = \{x \in X;\ \varepsilon_n(x) = r\}.$$

Weiter sei

$$H(r) = \bigcap_{p=1}^{\infty} \bigcup_{n=p}^{\infty} H(n, r). \tag{8}$$

Wenn $x \in H(r)$, dann ist offenbar $r \in \{\varepsilon_n(x)\}'_n$. Setzen wir noch $H = \bigcap_{r=0}^{\infty} H(r)$. Für $x \in H$ gilt dann

$$\{\varepsilon_n(x)\}'_n \supset \{0, 1, 2, \ldots, n, \ldots\}.$$

Es genügt zu beweisen, dass $H(r)$ eine in $\langle 0, 1)$ residuale Menge ist. Die Menge $H(n, r)$ ist entweder leer (dieser Fall tritt ein, wenn $q_n \leqq r$ ist) oder sie ist gleich der Vereinigungsmenge einer endlichen Anzahl der Mengen der Form $i_n^{(l)} \cap X$ (dieser Fall tritt ein, wenn $q_n > r$ ist). In beiden Fällen ist $H(n, r)$ eine in X offene Menge. Auf Grund von (8) ist also $H(r)$ eine G_δ-Menge in X.

Es genügt noch zu zeigen, dass $H(r)$ in X dicht ist (siehe den vorigen Beweis). Um die Dichtigkeit der Menge $H(r)$ zu beweisen, genügt es zu zeigen, dass, wenn m eine beliebige natürliche Zahl ist und l ganz, $0 \leqq l \leqq q_1 q_2 \ldots q_m - 1$, dann ist

$$i_m^{(l)} \cap H(r) \neq \emptyset. \tag{9}$$

Es haben also m, l die vorige Bedeutung. Es gehöre $i_m^{(l)}$ zur Folge $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m$. Mit N_0 bezeichnen wir die Menge aller derjenigen $j > m$, für welche $q_j > r + 1$ ist. Auf Grund der Voraussetzung des Satzes ist N_0 unendlich. Definieren wir die Zahl $x_0 = \sum_{j=1}^{\infty} \frac{\varepsilon_j(x_0)}{q_1 q_2 \ldots q_j}$ folgendermassen: Es sei $N_0 = N_1 \cup N_2$, $N_1 \cap N_2 = \emptyset$, N_1, N_2 seien unendlich. Dann setzen wir $\varepsilon_j(x_0) = \varepsilon_j (j = 1, 2, \ldots, m)$, $\varepsilon_j(x_0) = r$ für $j \in N_1$, $\varepsilon_j(x_0) = r$ für $j \in N_2$, und $\varepsilon_j(x_0) = 0$ für $j > m$, $j \notin N_0$. Dann ist ersichtlich $x_0 \in i_m^{(l)} \cap \cap H(r)$, also gilt (9). Damit ist der Beweis des Satzes beendet.

BEMERKUNG 1. In der Arbeit [8] ist bewiesen, dass im Falle $\sum_{n=1}^{\infty} 1/q_n < +\infty$ für fast alle $x \in \langle 0, 1)$ die Beziehung $\lim_{n \to \infty} \varepsilon_n(x) = +\infty$ gilt. Im Zusammenhang mit diesem Ergebnis bemerken wir, dass aus dem Satz 2 das folgende Ergebnis folgt: Die Menge

$$\{x \in \langle 0, 1);\quad \lim_{n \to \infty} \varepsilon_n(x) = +\infty\}$$

ist eine Menge von erster Kategorie in $\langle 0, 1)$, wenn $\limsup_{k \to \infty} q_k = +\infty$ ist.

Aus der Konstruktion der Cantorschen Entwicklungen der Zahlen $x \in \langle 0, 1)$ folgt leicht, dass im Falle $\limsup_{k \to \infty} q_k = +\infty$ die Menge $M_\infty = M_\infty(q_1, q_2, \ldots)$ aller $x \in \langle 0, 1)$ (siehe (1)) mit den beschränkten Folgen von Ziffern $\{\varepsilon_n(x)\}_{n=1}^{\infty}$ das Lebesguesche Mass 0 hat. An dieses Ergebnis knüpft der folgende Satz an.

SATZ 3. *Es sei* $\limsup_{k \to \infty} q_k = +\infty$. *Dann ist die Menge* $M_\infty = M_\infty(q_1, q_2, \ldots)$ *von erster Kategorie in* $\langle 0, 1)$.

BEWEIS. Offenbar ist

$$M_\infty = \bigcup_{s=1}^{\infty} M_s,$$

wo $M_s = \{x \in \langle 0, 1);\ \varepsilon_k(x) \leqq s,\ k = 1, 2, 3, \ldots\}$ ist. Es genügt also zu beweisen, dass jede der Mengen M_s $(s = 1, 2, \ldots)$ nirgends dicht in $\langle 0, 1)$ ist.

Es sei I ein Intervall, $I \subset \langle 0, 1)$. Dann existiert eine natürliche Zahl n und eine ganze Zahl l, $0 \leqq l \leqq q_1 \ldots q_n - 1$ so, dass $i_n^{(l)} \subset I$ ist. Es gehöre $i_n^{(l)}$ zur Folge $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$.

Wählen wir $m > n$ derart, dass $q_m > s+1$. Das ist infolge der Voraussetzung des Satzes möglich. Konstruieren wir jetzt das Intervall der m-ten Ordnung, welches zur Folge

$$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n, 0, 0, \ldots, 0, s+1$$

gehört. Bezeichnen wir dieses Intervall mit I^*. Dann ist offenbar $I^* \subset i_n^{(l)} \subset I$ und $I^* \cap M_s = \emptyset$. Daraus folgt, dass die Menge M_s nirgends dicht in $\langle 0, 1)$ ist. Damit ist der Beweis des Satzes beendet.

LITERATUR

[1] HARDY, G. H. and WRIGHT E. M.: *An introduction to the theory of numbers*, Oxford, 1954.

[2] ŠALÁT, T.: A remark on normal numbers, *Revue roumaine de math. pures et appl.* XI (1966), 53—56.

[3] PERRON, O.: *Irrationalzahlen*, De Gruyter, Berlin—Leipzig, 1921.

[4] RÉNYI, A.: A számjegyek eloszlása valós számok Cantor-féle előállításaiban, *Mat. Lap.* 7 (1956), 77—100.

[5] ERDŐS, P. and RÉNYI A.: Some further statistical properties of the digits in Cantor's series, *Acta Math. Acad. Sci. Hung.* **10** (1959), 21—29.

[6] ŠALÁT, T.: Über die Cantorschen Reihen, *Czechosl. Math. J.* **18** (93) (1968), 25—56.

[7] KURATOWSKI, K.: *Topologie I*, Warszawa, 1958.

[8] ERDŐS, P. and RÉNYI A.: On Cantor's series with convergent $\Sigma \frac{1}{q_n}$, *Ann. Univ. Sci. Budap. de Rol. Eötvös nom.* **2** (1959), 93—109.

[9] OSTMANN, *H. H. Additive Zahlentheorie I*, Springer-Verlag, Berlin—Heidelberbg—Göttingen, 1956.

Bratislava

(Eingegangen: 7 September 1967.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 55—61.

ON THE CONVERGENCE OF EMPIRICAL DISTRIBUTION AND DENSITY FUNCTIONS

by

G. DÁVID

1. §. Introduction

We wish to estimate the distribution function $F(x)$ and density function $f(x)$ of some random variable ξ by means of the ordered sample $\xi_1^*, \xi_2^*, \ldots, \xi_m^*$ (that is $\xi_k^* \leqq \xi_{k+1}^*$ for $k=1, 2, \ldots, m-1$). The empirical distribution function is defined by the formula

$$F_m(x) = \frac{k_x}{m}$$

where k_x is the number of elements of sample, which are equal or less than x. In the Monte-Carlo-methods we are in need of samples whit great number of elements, so we have no possibility reserving all elements of the sample because of the capacity of the memory of computers. In this paper we give estimators which on one hand are uniformly convergent with probability one, and on other hand which are applicable in case of a great number of trials.

In 2 § and 3 § we shall give the new definitions of the empirical density and distribution functions respectively, and we shall prove theorems for the rate of convergence, and a similar theorem for the $F_m(x)$.

2. §. Theorems for the density function

Having the ordered sample $\xi_1^*, \xi_2^*, \ldots, \xi_n^*$, let us denote by $I^n(x)$ that interval from the intervals $(-\infty, \xi_1^*], (\xi_1^*, \xi_2^*], \ldots, (\xi_n^*, \infty)$ which contains the point x. For arbitrary $m(>n)$ let us define the empirical density function by means of the first n elements of sample with the following formula

$$f_n^m(x) = \frac{\mathsf{P}_m(I^n(x))}{\lambda(I^n(x))}$$

where $\lambda(A)$ ist the Lebesgue-measure and $\mathsf{P}_m(A)$ is the empirical probability measure generated by $F_m(x)$-of a Borel-set A.

THEOREM 1. *If* $[a, b]$ *is a closed interval in which* $f(x)$ *and the inverse of* $F(x)$ *in* $[F(a), F(b)]$ *are continuous then*

$$\lim_{n\to\infty} \lim_{m\to\infty} \sup_{x\in[a,b]} |f_n^m(x) - f(x)| = 0$$

with probability one.

PROOF. Since

$$f_n^m(x) = \frac{\mathsf{P}_m(I^n(x))}{\mathsf{P}(I^n(x))} \cdot \frac{\mathsf{P}(I^n(x))}{\lambda(I^n(x))}$$

and the Glivenko—Cantelli-lemma gives for fixed n, that $\mathsf{P}_m(I^n(x)) \to \mathsf{P}(I^n(x))$ uniformly with probability 1; so it is sufficient to prove that

$$f_n(x) = \frac{\mathsf{P}(I^n(x))}{\lambda(I^n(x))} \to f(x)$$

uniformly with probability one.

We have to prove only the uniformity of this convergence which depends on the uniformity of the convergence of lengths $\lambda(I^n)$, but this fact follows from on one hand the Glivenko—Cantelli-lemma applied to the ordered sample of the uniformly distributed random variable η in [0, 1] received by the transformation $\eta = F(x)$ and on other hand the uniform continuity of $F^{-1}(x)$.

REMARK 1. From the proof we can conclude the pointwise convergence of the sample-density $f_n^m(x)$ to $f(x)$ with probability one without the assumption in Theorem 1.

Let the function $f(x)$ be differentiable, positive and

$$f(x) > c > 0, \quad |f'(x)| < c_1 < +\infty, \quad f'(x) \neq 0$$

with suitable constans c and c_1.

Let us define

$$\omega_1(n, \alpha) = \frac{c}{4c_1}\left(1 - e^{-\frac{1}{n}(1+\alpha)\log n}\right)^{-}$$

(1) $$\omega_2(n, m, x) = \lambda(I^n(x)) \cdot \left\{\frac{2}{m}\mathsf{P}(I^n(x))[1 - \mathsf{P}(I^n(x))]\log\log m\right\}^{-\frac{1}{2}},$$

$$\omega(n, m, \alpha, x) = \frac{1}{2}\min\{\omega_1(n, \alpha),\ \omega_2(n, m, x)\},$$

for all n, m $(m > n)$ and for arbitrary small $\alpha > 0$.

THEOREM 2. *If the point x has a neighbourhood $[a, b]$ satisfying the assumption of the Theorem 1, $f(x) > c > 0$ and $f(x)$ has first derivate $f'(x) \neq 0$, $|f'(x)| < c_1 < +\infty$ then*

$$\mathsf{P}\left(\overline{\lim_{n\to\infty}}\ \overline{\lim_{m\to\infty}}\ \omega(n, m, \alpha, x)|f_n^m(x) - f(x)| \leqq 1\right) = 1$$

where the function $\omega(n, m, \alpha, x)$ is defined in (1).

PROOF. It is enough to prove that

$$\sum_{n=1}^{\infty} \mathsf{P}\left(\overline{\lim_{m\to\infty}}\ \omega(n, m, \alpha, x)|f_n^m(x) - f(x)| > 1 + \varepsilon\right) < +\infty$$

for arbitrary $\alpha > 0$. But

$$\left\{\overline{\lim_{m\to\infty}}\, \omega(n, m, \alpha, x)\,|f_n^m(x) - f(x)| > 1+\varepsilon\right\} \subseteqq$$

$$\subseteqq \left\{\overline{\lim_{m\to\infty}}\, \omega(n, m, \alpha, x)\,|f_n^m(x) - f_n(x)| > \frac{1+\varepsilon}{2}\right\} \cup \tag{2}$$

$$\cup \left\{\overline{\lim_{m\to\infty}}\, \omega(n, m, \alpha, x)\,|f_n(x) - f(x)| > \frac{1+\varepsilon}{2}\right\}$$

where

$$f_n(x) = \frac{1}{\lambda(I^n(x))} \int\limits_{I^n(x)} f(t)\,dt$$

Now for the second event

$$\left\{\overline{\lim_{m}}\, \omega(n, m, \alpha, x)\,|f_n(x) - f(x)| > \frac{1+\varepsilon}{2}\right\} \subseteqq \{\omega_1(n, \alpha)\,|f_n(x) - f(x)| > 1+\varepsilon\}$$

follows.

We shall estimate the latest event by means of the chain

$$\frac{|F(x)-F(y)|}{|f(x)-f(y)|} = \frac{|F(x)-F(y)|}{|x-y|} \cdot \frac{|x-y|}{|f(x)-f(y)|} = \tag{3}$$

$$= \frac{f(x)+o(x-y)}{|f'(x)+O(x-y)|} \geqq \frac{f(x)+o(x-y)}{2C_1}$$

and this holds except for a set of Lebesgue-measure 0, and on the other hand

$$\left\{|f_n(x)-f(x)| > \frac{1}{\omega_1(n, \alpha)}\right\} \subseteqq \left\{\sup_{x, y \in I^n} |f(x)-f(y)| > \frac{1}{\omega_1(n, \alpha)}\right\}$$

and this implies the existence of points $x_0, y_0 \in [\xi_k^1, \xi_{k+1}^*] = I^n(x)$ for which $|f(x_0) - f(y_0)| > \dfrac{1}{\omega_1(n, \alpha)}$ and applying the estimation (3) we have

$$|F(\xi_k^*) - F(\xi_{k+1}^*)| \geqq |F(x_0) - F(y_0)| \gg \frac{f(x_0)+o(x_0-y_0)}{2C_1} \cdot \frac{1}{\omega_1(n, \alpha)}$$

Using this:

$$\mathsf{P}\left(|f_n(x)-f(x)| > \frac{1}{\omega_1(n, \alpha)}\right) \leqq$$

$$\leqq \mathsf{P}\left(|F(\xi_k^*) - F(\xi_{k+1}^*)| > \frac{f(x_0)+o(x_0-y_0)}{2C_1} \frac{1}{\omega_1(n, \alpha)}\right) \leqq$$

$$\leqq \mathsf{P}\left(F(\xi_{k+1}^*) - F(\xi_k^*) > \left(1 - e^{-\frac{1}{n}(1+\alpha)\log n}\right)\right) = \tag{4}$$

$$= \left[1 - \left(1 - e^{-\frac{1}{n}(1+\alpha)\log n}\right)\right]^n = \frac{1}{n^{1+\alpha}}$$

where we used the Beta-distribution with parameter-pair $(1, n)$ for $\eta^*_{k+1}-\eta^*_k$ (for $\eta=F(\xi)$), — that is we have for arbitrary $\varepsilon>0$

$$\sum_{n=1}^{\infty} \mathsf{P}\big(|f_n(x)-f(x)|\,\omega_1(n,\alpha) > 1+\varepsilon\big) < +\infty \tag{5}$$

The probability of the first event in the right side of (2)

$$\mathsf{P}\left[\overline{\lim_{m}} \frac{\lambda(I^n(x))}{\sqrt{\frac{2}{m}\mathsf{P}(I^n(x))(1-\mathsf{P}(I^n(x)))\log\log m}}\left|\frac{\mathsf{P}_m(I^n(x))}{\lambda(I^n(x))}-\frac{\mathsf{P}(I^n(x))}{\lambda(I^n(x))}\right| > 1+\varepsilon\right] = 0$$

from the theorem of the iterated logarithm. This and (5) prove the Theorem 2.

It is easy to see if the inequality

$$\frac{m\left[1-\exp\left(-\frac{1}{n}(1+\alpha)\log n\right)\right]^2}{\log\log m} \geqq \frac{2\mathsf{P}(I^n(x))(1-\mathsf{P}(I^n(x)))}{\lambda(I^n(x))}\left(\frac{4C_1}{C}\right)^2 \tag{6}$$

holds for m and n then

$$\omega(n,m,\alpha,x) = \frac{C}{8C_1}\left\{1-\exp\left[-\frac{1}{n}(1+\alpha)\log n\right]\right\}^{-1}$$

So we have

REMARK 2. *If* (6) *holds for n and m then, under the assumptions of Theorem 2*

$$\mathsf{P}\left(\lim_{n\to\infty}\lim_{m\to\infty}\frac{C}{8C_1}\frac{|f_n^m(x)-f(x)|}{1-\exp\left(-\frac{1+\alpha}{n}\log n\right)} \leqq 1\right) = 1$$

The proof is immediate.

This remark and (6) make possible the appropriate choise of the function $m(n)$.

In the sequel let the function $\omega_1^*(n,\alpha)$ be

$$\omega_1^*(n,\alpha) = \frac{C}{4C_1}\left\{1-\exp\left[-\frac{2+\alpha}{n}\log n\right]\right\}^{-1}$$

and

$$\omega^*(n,m,\alpha,x) = \frac{1}{2}\min\{\omega_1^*(n,\alpha),\ \omega_2(n,m,x)\}$$

with arbitrary small $\alpha>0$. Let be true for m and n the following inequality:

$$\frac{m\left\{1-\exp\left[-\frac{2+\alpha}{n}\log n\right]\right\}^2}{\log\log m} \geqq \frac{2\mathsf{P}(I^n(x))[1-\mathsf{P}(I^n(x))]}{\lambda(I^n(x))} \tag{7}$$

THEOREM 3. *Let* $[a, b]$ *be such an interval for which* $F^{-1}(x)$ *is continuous in* $[F(a), F(b)]$, $f(x)$ *has bounded first derivate* $f'(x)$ *in* $[a, b]$ *and* $f'(x) \neq 0$ *for* $x \in [a, b]$, *so*

$$\mathsf{P}\left(\overline{\lim_n}\,\overline{\lim_m}\,\frac{1}{2}\,\omega_1^*(n, \alpha) \sup_{x\in[a,b]} |f_n^m(x) - f(x)| \leq 1\right) = 1$$

PROOF. Repeating the decomposition applied in the proof of Theorem 2 and taking $m = g(n)$ when the equality holds in (7), we get the estimation

$$\sum_{n=1}^{\infty} \mathsf{P}\left(\overline{\lim_{m>g(n)}}\,\frac{1}{2}\,\omega_1^*(n, \alpha) \sup_{x\in[a,b]} |f_n^m(x) - f(x)| > 1+\varepsilon\right) \leq$$

(8)
$$\sum_{n=1}^{\infty}\sum_{i=1}^{m} \mathsf{P}\left(\overline{\lim_m}\,\omega^*(n, m, \alpha, \xi_i^*)\,|f_n^m(\xi_i^*) - f(\xi_i^*)| > 1+\varepsilon\right)$$

$$\leq \sum_{n=1}^{\infty} n \max_{i=1,2,\dots,n} \mathsf{P}\left(\overline{\lim_m}\,\omega^*(n, m, \alpha\, \xi_i^*)\,|f_n^m(\xi_i^*) - f(\xi_i^*)| > 1+\varepsilon\right)$$

where we used the monotonity of function $f(x)$.

We shall only deal with the probabilities

$$\mathsf{P}\left(\overline{\lim_m}\,\omega_1^*(n, \alpha)\,|f_n(x) - f(x)| > 1+\varepsilon\right) = \mathsf{P}\left(\omega_1^*(n, \alpha)\,|f_n(x) - f(x)| > 1+\varepsilon\right)$$

as in the proof of Theorem 2. Repeating the argument used there we have (see (4)) that this probability is less or equal than $n^{-(2+\alpha)}$, and taking into account (8), we proved the theorem.

3 § Theorems for empirical distribution functions

The following theorem is an obvious one:

THEOREM 4. *Let* $\tilde{F}_{n,m}(x) = \int_a^x f_n^m(t)\,dt + F_n(a)$ *be the modified empirical distribution function, then*

$$\mathsf{P}\left(\overline{\lim_{n\to\infty}}\,\overline{\lim_{m\to\infty}} \sup_{x\in[a,b]} |\tilde{F}_{n,m}(x) - F(x)| = 0\right) = 1$$

The proof is routine.

THEOREM 5. *Under conditions of the Theorem 3 and denoting* $\Omega(n, \alpha) = \dfrac{\omega_1^*(n, \alpha)}{2(b-a)}$ *we have*

$$\mathsf{P}\left(\overline{\lim_n}\,\overline{\lim_m}\,\Omega(n, \alpha) \sup_{x\in[a,b]} |\tilde{F}_{n,m}(x) - F(x)| \leq 1\right) = 1\,.$$

The proof: this theorem is an immediate consequence of Theorem 3.

In the sequel we shall deal with the groupped empirical distribution function defined by the following manner:

Let us divide the real line by points a_i^n for all n and $i = 1, 2, \dots, n$ such that $a_0^n = -\infty$, $a_i^n < a_{i+1}^n$, $a_{n+1}^n = +\infty$ and $|a_i^n - a_{i-1}^n| \to 0$ $i = 2, 3, \dots, n$; $a_1^n \to -,\infty$ $a_n^n \to +\infty$ if $n \to \infty$, let us denote by m_i the number of elements of the sample $\xi_1^*, \xi_2^*, \dots,$

..., ξ_n^* with $n \leqq m$, which are contained in the interval $(a_i^n, a_{i+1}^n]$, so put

(9)
$$G_m(x) = \frac{\sum_{i=0}^{j(x)} m_i}{m} \quad \text{if} \quad x \in [a_{j(x)}^n, a_{j(x)+1}^n]$$

Suppose that $[a, b]$ is such an interval in which $0 < C_2 < F(x) < C_3 < 1$ and the function $\Omega(m)$ is such that $\frac{1}{\Omega(m)} < F(x)[1-F(x)]$ in $[a, b]$.

Using the monotonity of the distribution function $F(x)$ we have

$$\mathsf{P}\Big(\Omega(m) \sup_{x \in [a,b]} |G_m(x) - F(x)| > 1+\varepsilon\Big) \leqq$$

$$\leqq \sum_{k=1}^{n} \mathsf{P}\left(\Omega(m)\left|\frac{\sum_{i=0}^{k-1} m_i}{m} - F(a_k^n)\right| > 1+\varepsilon\right)$$

It is well known that if ζ_m is the frequency of the event A in m trials and $\mathsf{P}(A) = p$ and $0 < \varepsilon < p(1-p)$ then

$$\mathsf{P}\left(\left|\frac{\zeta_m}{m} - p\right| \geqq \varepsilon\right) \leqq 2\exp\left\{-\frac{\varepsilon^2 m}{2pq\left(1+\frac{\varepsilon}{2pq}\right)^2}\right\}.$$

Hence

$$\mathsf{P}\left(\left|\frac{\sum_{i=1}^{k-1} m_i}{m} - F(a_k^n)\right| \geqq \frac{1+\varepsilon}{\Omega(m)}\right) \leqq$$

$$\leqq 2\exp\left\{-\frac{\left(\frac{1+\varepsilon}{\Omega(m)}\right)^2 m}{2F(a_k^n)[1-F(a_k^n)]\left[1+\frac{1+\varepsilon}{2F(a_k^n)(1-F(a_k^n))\Omega(m)}\right]^2}\right\}.$$

Because of $\max_{x \in [a,b]} F(x)(1-F(x)) = \frac{1}{4}$ and $\frac{1+\varepsilon}{2F(x)(1-F(x))\Omega(m)} < \frac{1+\varepsilon}{2} < 1$ we have

$$\mathsf{P}\left(\Omega(m)\left|\frac{\sum_{i=1}^{k-1} m_i}{m} - F(a_k^n)\right| > 1+\varepsilon\right) \leqq 2e^{-\frac{m}{2\Omega^2(m)}}.$$

Hence

$$\sum_{m=1}^{\infty}\sum_{k=1}^{n} \mathsf{P}(\Omega(m)|G_m(a_k^n) - F(a_k^n)| > 1+\varepsilon) \leqq \sum_{m=1}^{\infty} 2me^{-\frac{m}{2\Omega^2(m)}} < +\infty.$$

If $\exp\left(-\frac{m}{2\Omega^2(m)}\right) = \frac{1}{m^{2+\alpha}}$, from which the equality

(10)
$$\Omega(m) = \sqrt{\frac{m}{(4+\alpha)\log m}}$$

follows with arbitrary small $\alpha > 0$.

Applying the Borel—Cantelli-lemma we have the following theorem:

THEOREM 6. *If* $0<c_1<F(a)<F(b)<c_2<1$ *in* $[a, b]$ *and* $\Omega(m)$ *is defined by (10) then*

$$\mathsf{P}\left(\lim_m \Omega(m) \sup_{x\in[a,b]} |G_m(x)-F(x)| \leqq 1\right) = 1.$$

THEOREM 7. *If the distribution function* $F(x)$ *is continuous then for arbitrary* α $(\alpha>0)$ *and interval* $[a, b]$ *in which* $0<F(a)<F(b)<1$ *we have*

$$\mathsf{P}\left(\lim_{m\to\infty} \frac{1}{2}\sqrt{\frac{m}{(4+\alpha)\log m}} \sup_{x\in[a,b]} |F_m(x)-F(x)| \leqq 1\right) = 1.$$

PROOF. Let $a_0^m=a<a_1^m<\ldots<a_{n(m)}^m=b$ be points of $[a, b]$ such that — denoting $(a_i^m, a_{i+1}^m]$ by I_i^m —

$$p_m = \max_{i=0,1,\ldots,n-1} \{\mathsf{P}(I_i^m)\} \leqq \frac{1}{\sqrt{m^{2+\alpha}}}$$

and $G_m(x)$ is defined in (9), we have

$$\mathsf{P}\left(\overline{\lim_m}\sqrt{\frac{m}{(4+\alpha)\log m}} \sup_{x\in[a,b]} |F_m(x)-F(x)| > 2\right) \leqq$$

$$\leqq \mathsf{P}\left(\overline{\lim_m}\sqrt{\frac{m}{(4+\alpha)\log m}} \sup_{x\in[a,b]} |G_m(x)-F_m(x)| > 1\right) +$$

$$+\mathsf{P}\left(\overline{\lim_m}\sqrt{\frac{m}{(4+\alpha)\log m}} \sup_{x\in[a,b]} |G_m(x)-F(x)| > 1\right).$$

The latest probability equals 0 because of Theorem 6. On the other hand for arbitrary fixed ε $(1>\varepsilon>0)$ we have estimation

$$\sum_{m=1}^{\infty}\mathsf{P}\left(\sqrt{\frac{m}{(4+\alpha)\log m}} \sup_{x\in[a,b]} |G_m(x)-F_m(x)| > 1+\varepsilon\right) \leqq$$

$$\leqq \sum_{m=1}^{\infty}\mathsf{P}\left(m \sup_{x\in[a,b]} |G_m(x)-F_m(x)| > 1+\varepsilon\right) \leqq \sum_{m=1}^{\infty}\mathsf{P}\left(m \sup_{x\in[a,b]} |G_m(x)-F_m(x)| \geqq 2\right)$$

$$\leqq \sum_{m=1}^{\infty} m \sum_{k=2}^{m} \mathsf{P}(m_i=k) \leqq \sum_{m=1}^{\infty} mp_m^2 \sum_{k=0}^{\infty} p_m^k \leqq C\sum_{m=1}^{\infty} mp_m^2 < +\infty.$$

From the Borel—Cantelli-lemma the theorem follows.

AKNOWLEDGEMENT. I should like to express my thank to Mr. P. RÉVÉSZ for his valuable ideas.

REFERENCES

[1] PARZEN, E.: On the estimation of a probability density function and modes, *Ann. Math. Stat.* **33** (1962), 1065—1076.

[2] RÉVÉSZ P.: *The laws of large numbers*, Publ. House of Hung. Acad. 1967.

Computing Centre of the Hungarian Academy of Sciences
(Received September 4, 1968.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 63—65.

SPHERICAL FUNCTIONS ON COMPACT RIEMANNIAN SYMMETRIC SPACES

by

R. G. LAHA

The main aim of the present note is to give a simple, alternative proof of the theorem on integral representation of spherical functions on compact Riemannian symmetric spaces (cf. [1], p. 426).

THEOREM. *Let* (G, K) *be a Riemannian symmetric pair of the compact type such that the group* G *is a compact connected Lie group. Let* dk *be the Haar measure on* K *such that*

$$\int_K dk = 1.$$

Let φ *be a spherical function on* G. *Then there exists a finite dimensional irreducible unitary representation* π *of group* G *of class 1 with character* χ *such that the function* φ *has the integral representation*

$$\varphi(g) = \int_K \chi(g^{-1}k)\,dk \qquad (g \in G). \tag{1}$$

Moreover the function φ *is an elementary positive definite function on* G *associated with the representation* π.

Conversely let π *be an arbitrary finite dimensional irreducible unitary representation of* G *of class 1 with character* χ. *Then the function* φ *on* G *defined by the formula* (1) *is an elementary positive definite spherical function on* G.

PROOF: Let φ be a spherical function on G. Let $\tau_g: x \to g^{-1}x$ $(x \in G)$ be the left translation of the group G associated with the element $g \in G$. Let $\mathfrak{L}_2(G)$ be the Hilbert space of all complex valued measurable functions on G which are square summable with respect to the Haar measure dg on G. Let $\mathfrak{H}_\varphi$ be the closure (in the $\mathfrak{L}_2$-norm) of the linear hull of the set $\{\varphi\tau_g : g \in G\}$ in the space $\mathfrak{L}_2(G)$. Then the space $\mathfrak{H}_\varphi$ is a Hilbert space. Let $\pi: g \to \pi(g)$ $(g \in G)$ be a representation of G in the space $\mathfrak{H}_\varphi$ defined by the formula

$$\pi(g)\theta = \theta\tau_g \tag{2}$$

for $g \in G$, $\theta \in \mathfrak{H}_\varphi$. Then we see easily that π is a unitary representation of G in the space $\mathfrak{H}_\varphi$. Moreover the relation $\pi(k)\varphi = \varphi$ holds for all $k \in K$.

Let $\theta, \psi \in \mathfrak{H}_\varphi$. We now set

$$B(\theta, \psi) = \int_K (\pi(k)\theta, \psi)\,dk. \tag{3}$$

Then we can verify easily that B is a bounded bilinear form in the space $\mathfrak{H}_\varphi$ so that there exists a bounded linear operator P in the space $\mathfrak{H}_\varphi$ such that the relation

$$(4)\qquad (P\theta, \psi) = B(\theta, \psi) = \int_K (\pi(k)\theta, \psi)\, dk$$

holds for all $\theta, \psi \in \mathfrak{H}_\varphi$.

We now set

$$(5)\qquad P = \int_K \pi(k)\, dk.$$

Then it can be easily verified that P satisfies the relation

$$(6)\qquad P^2 = P^* = P = P\pi(k) = \pi(k)P$$

for all $k \in K$. Hence P is a projection operator in the space $\mathfrak{H}_\varphi$: We also note that the relation $P\varphi = \varphi$ holds.

Let $g \in G$, $\psi \in \mathfrak{H}_\varphi$. Then we have

$$(7)\qquad (P\pi(g)\varphi, \psi) = \int_K (\pi(kg)\varphi, \psi)\, dk = \varphi(g^{-1})(\varphi, \psi).$$

Hence we conclude from (7) that the relation

$$(8)\qquad P\varphi\tau_g = P\pi(g)\varphi = \varphi(g^{-1})\varphi$$

holds for all $g \in G$. Then it follows from the definition of the space $\mathfrak{H}_\varphi$ that

$$(9)\qquad P\mathfrak{H}_\varphi = C\varphi$$

where C is the field of complex numbers.

Let $\mathfrak{H}_1$ be the closure of the sum of all closed subspaces $\mathfrak{M} \subset \mathfrak{H}_\varphi$ which are invariant with respect to π and satisfies the condition $P\mathfrak{M} = (0)$. Let $\mathfrak{H}_2 = \mathfrak{H}_\varphi \ominus \mathfrak{H}_1$ be the orthogonal complement of $\mathfrak{H}_1$ in $\mathfrak{H}_\varphi$. Then both $\mathfrak{H}_1$ and $\mathfrak{H}_2$ are closed subspaces of $\mathfrak{H}_\varphi$ which are invariant with respect to π and moreover $P\mathfrak{H}_1 = (0)$ while $P\mathfrak{H}_2 \neq (0)$. Let $\mathfrak{N} \neq (0)$ be an arbitrary closed subspace of the space $\mathfrak{H}_2$ which is invariant with respect to π. Then we have $P\mathfrak{N} \neq (0)$ so that we can verify easily from (9) that $\varphi \in P\mathfrak{N} \subset \mathfrak{N}$. Hence we conclude that $\pi(g)\varphi \in \mathfrak{N}$ for every $g \in G$, so that $\mathfrak{N} = \mathfrak{H}_\varphi$. Therefore the only closed subspaces of the space $\mathfrak{H}_\varphi$ which are invariant with respect to π are (0) and $\mathfrak{H}_\varphi$. Consequently π is an irreducible unitary representation of G in the space $\mathfrak{H}_\varphi$. But since G is compact, we conclude that the space $\mathfrak{H}_\varphi$ is finite dimensional. Let $\dim \mathfrak{H}_\varphi = n$.

We set $\|\varphi\|_2 = (\varphi, \varphi)^{\frac{1}{2}}$. Let $\dfrac{\varphi}{\|\varphi\|_2} = \varphi_1, \varphi_2, \ldots, \varphi_n$ be an orthonormal basis of the space $\mathfrak{H}_\varphi$. Then it follows immediately from (8) and (9) that

$$P\pi(g)P\varphi_1 = \varphi(g^{-1})\varphi_1; \quad P\pi(g)P\varphi_j = 0 \quad \text{for} \quad j = 2, 3, \ldots, n$$

so that we conclude that

$$(10)\qquad \begin{aligned} \varphi(g^{-1}) &= \operatorname{tr}(P\pi(g)P) = \operatorname{tr}(\pi(g)P) \\ &= \operatorname{tr}\int_K \pi(gk)\, dk = \int_K \chi(gk)\, dk. \end{aligned}$$

Moreover in this case we have

$$(11)\qquad \begin{aligned}(\varphi_1, \pi(g)\varphi_1) &= (\pi(g^{-1})\varphi_1, P\varphi_1)\\ &= (P\pi(g^{-1})\varphi_1, \varphi_1) = \varphi(g).\end{aligned}$$

Consequently φ is an elementary positive definite function on G associated with the representation π.

Since φ is bi-invariant with respect to K, we conclude that π is a representation of G of class 1.

Conversely let π be a finite dimensional irreducible unitary representation of G of class 1 in a Hilbert space $\mathfrak{H}$ and let χ be the corresponding character. Let $\xi_0 \in \mathfrak{H}$ such that $(\xi_0, \xi_0) = 1$ and moreover the relation $\pi(k)\xi_0 = \xi_0$ holds for all $k \in K$. Then the function φ on G defined by

$$(12)\qquad \varphi(g) = (\xi_0, \pi(g)\xi_0 \qquad (g \in G)$$

is an elementary positive definite function on G which is bi-invariant with respect to K. Hence in view of Theorem 4. 5 ([1], p. 414) we conclude that φ is a spherical function on G. We now set

$$P = \int_K \pi(k)\, dk$$

and we note the P satisfies the relation (6).

Therefore we have

$$(13)\qquad \operatorname{tr}(P\pi(g)P) = \int_K \chi(gk)\, dk.$$

On the other hand in view of (6) the relation

$$(14)\qquad \pi(k)P\pi(g)\xi_0 = P\pi(g)\xi_0$$

holds for all $g \in G$, $k \in K$. Then using Lemma 4. 7 ([1], p. 416), we conclude that

$$(15)\qquad P\mathfrak{H} = C\xi_0.$$

Let $\xi_0 = e_1, e_2, \ldots, e_n$ be an orthonormal basis of the space $\mathfrak{H}$. Then in view of (15) we have $P\pi(g)Pe_j = 0$ for $j = 2, 3, \ldots, n$.

Hence we conclude from (12) that

$$(16)\qquad \begin{aligned}\operatorname{tr}(P\pi(g)P) &= (P\pi(g)\xi_0, \xi_0)\\ &= (\pi(g)\xi_0, \xi_0) = \varphi(g^{-1}).\end{aligned}$$

Then combining (13) and (16) we conclude that the function φ defined by the formula (1) is an elementary positive definite spherical function on G.

REFERENCE

[1] HELGASON, S.: *Differential Geometry and Symmetric Spaces*. Academic Press, New York, 1962.

The Catholic University of America
(Received August 23, 1968.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 67—73.

ON THE THRESHOLD DISTRIBUTION FUNCTION OF CYCLES IN A DIRECTED RANDOM GRAPH

by

I. PALÁSTI

Let $\Gamma_{n,N}$ be a random graph having n given labelled vertices and N edges chosen at random, so that all the $\binom{\binom{n}{2}}{N}$ possible choices are supposed to be equiprobable.

P. ERDŐS and A. RÉNYI investigated in their paper [1] the evolution of random graphs. They treated the "typical" structures arising at a given stage of evolution, if N increases together with n, as a given function of n. By a typical structure they mean such a structure the probability of which tends to 1 for $n \to \infty$. If A is a structural property which a random graph may or may not posses, and $\mathsf{P}_{n,N(n)}(A)$ denotes the probability that the random graph $\Gamma_{n,N(n)}$ posesses the property A where $N(n)$ is a given function of n, and if

$$\lim_{n\to\infty} \mathsf{P}_{n,N(n)}(A) = 1, \tag{1}$$

then we say that "almost all" random graphs $\Gamma_{n,N(n)}$ posses this property. If it is also true that there exists a probability distribution function $F(x)$ such that

$$\lim_{n\to\infty} \mathsf{P}_{n,N}(A) = F(x) \quad \text{if} \quad \lim_{n\to\infty} \frac{N(n)}{A(n)} = x \tag{2}$$

then we say that $A(n)$ is a regular threshold function, and $F(x)$ is the threshold distribution function concerning to the property A. In such a way they obtained "threshold functions" and "threshold distribution functions" for certain structural properties. Among a lot of theorems they proved the following two concerning cycles. (A cycle is called of order k if it has exactly k points and k edges.)

Theorems of P. Erdős and A. Rényi: (Theorems 3a. and 3b. of [1])

a) Let us suppose that

$$N(n) \sim cn \quad \text{where} \quad c > 0, \tag{3}$$

and let us denote by γ_k the number of cycles of order k contained in the random graph $\Gamma_{n,N}$ (where $k = 3, 4, \ldots$). Then

$$\lim_{n\to\infty} \mathsf{P}_{n,N(n)}(\gamma_k = j) = \frac{\lambda^j e^{-\lambda}}{j!} \qquad (j = 0, 1, \ldots) \tag{4}$$

where

$$\lambda = \frac{(2c)^k}{2k}. \tag{5}$$

Thus the threshold distribution corresponding to the threshold function $A(n)=n$ for the property that the random graph contains a cycle of order k is $1-e^{-\frac{(2c)^k}{2k}}$.

b) Suppose that (3) holds. Let γ_k^* denote the number of isolated cycles of order k contained in $\Gamma_{n,N(n)}$; $k=3,4,\ldots$. Then

$$\lim_{n\to\infty} \mathsf{P}_{n,N(n)}(\gamma_k^* = j) = \frac{\mu^j e^{-\mu}}{j!} \qquad (j=0,1,\ldots) \tag{6}$$

where

$$\mu = \frac{(2ce^{-2c})^k}{2k}. \tag{7}$$

Further on we shall deal with directed random graphs.

As a graph is called directed, if all of its edges are directed; a directed random graph $\vec{\Gamma}_{n,N}$ is obtained, if we chose N different edges among the possible n^2 (directed) ones connecting n given vertices so that each of the $\binom{n^2}{N}$ possible choices are equiprobable. We allow now loops, that is edges (P_i, P_i) too. We suppose furthermore that there are no parallel edges with the same direction connecting two points.

In our paper [2] we were intent to know how large directed cycle would appear with probability 1 in such a directed random graph $\vec{\Gamma}_{n,N_c}$ at the evolution stage for which the number of the chosen edges is $N_c = [n\log n + cn]$, where c is an arbitrary fixed real number and $[x]$ denotes the integral part of the number x in the brackets, and found the following.

A directed random graph $\vec{\Gamma}_{n,N_c}$ contains a directed cycle of order $[\log\log n]$ "almost surely" if $N_c = [n\log n + cn]$; by other words, the probability of the directed random graph $\vec{\Gamma}_{n,N_c}$ containing such a cycle tends to 1 for $n\to\infty$.

This result was used as a lemma for proving the strong connectedness of the directed random graphs. After having all these it is easy to state the theorems which are analogous to the above mentioned theorems of P. ERDŐS and A. RÉNYI, but concerning to the directed random graphs. That is to say to the distribution of the numbers of directed cycles arising in the directed random graphs. The results may be formulated as follows:

THEOREM 1. *Let us denote by δ_l the number of directed cycles of order l, that is the number of cycles $P_{i_1}\to P_{i_2}\to\cdots\to P_{i_l}\to P_{i_1}$ in a directed random graph $\vec{\Gamma}_{n,N}$ (where $i_1, i_2, \ldots, i_l$ are all different and $l=1,2,3,\ldots$). Then the limiting distribution of the number of edge-independent directed cycles is a Poisson distribution with the parameter*

$$\lambda = \frac{c^l}{l}. \tag{8}$$

That is

$$\lim_{n\to\infty} \mathsf{P}_{n,N(n)}(\delta_l = j) = \frac{\lambda^j e^{-\lambda}}{j!} \qquad (j=0,1,\ldots) \tag{9}$$

supposing that, $N(n) \sim cn$ *and* $c > 0$. *Thus the threshold distribution function of the directed random graph* $\vec{\Gamma}_{n, N(n)}$ *with* $N(n) \sim cn$ *containing a cycle of order* l *is* $1 - e^{-\frac{c^l}{l}}$.

THEOREM 2. *Let us suppose again that* $N(n) \sim cn$ *and let us now denote* δ_l^* *the number of isolated directed cycles of order* l *contained in* $\vec{\Gamma}_{n, N(n)}$, *where* $l = 1, 2, 3, \ldots$. *Then the number* δ_l^* *has a Poisson distribution, for* $n \to \infty$, *with the parameter*

$$\mu = \frac{(ce^{-2c})^l}{l}. \tag{10}$$

First let us consider *the proof of theorem 2*. We shall prove it in an analogous way as it was done in paper [1].

Let us denote by $\varkappa = (i_1, i_2, \ldots, i_l)$ an ordered l-tuple, where the different numbers $i_1, i_2, \ldots, i_l$ are arbitrary chosen from the numbers $1, 2, \ldots, n$. We do not distinguish between the l-tuple $\varkappa$ and its cyclic permutations. Let I_l be the set of all possible such l-tuples $\varkappa$.

Let us introduce now the random variable

$$\varepsilon_\varkappa = \begin{cases} 1 & \text{if the isolated directed cycle } P_{i_1} \to P_{i_2} \to \cdots \to P_{i_l} \to P_{i_1} \\ & \text{belongs to the random graph } \vec{\Gamma}_{n,N}, \\ 0 & \text{otherwise.} \end{cases}$$

Then the mean value of $\varepsilon_\varkappa$ is

$$\mathsf{M}(\varepsilon_\varkappa) = \frac{\binom{(n-l)^2}{N-l}}{\binom{n^2}{N}} \sim \frac{N^l}{n^{2l}} \prod_{j=0}^{N-l-1} \frac{(n-l)^2 - j}{n^2 - l - j}. \tag{11}$$

That is

$$\mathsf{M}(\varepsilon_\varkappa) \sim \frac{N^l}{n^{2l}} e^{-2l\frac{N}{n}}, \tag{12}$$

where $\frac{N}{n} \to c$. Accordingly if $\varkappa_1, \varkappa_2, \ldots, \varkappa_r$ are disjoint ordered l-tuples formed from the numbers $1, 2, \ldots, n$

$$\mathsf{M}(\varepsilon_{\varkappa_1} \varepsilon_{\varkappa_2} \ldots \varepsilon_{\varkappa_r}) = \frac{\binom{(n-rl)^2}{N-rl}}{\binom{n^2}{N}} \sim \frac{N^{rl}}{n^{2rl}} e^{-2rl\frac{N}{n}}. \tag{13}$$

Let be $\varepsilon = \sum_{\varkappa \in I_l} \varepsilon_\varkappa$; then ε is clearly equal to the total number of isolated cycles of order l in $\vec{\Gamma}_{n,N}$. On the other hand the number of elements $\varkappa$ of I_l is equal to $\binom{n}{l}(l-1)!$ In view of the asymptotical formula

$$\binom{n}{l}\binom{n-l}{l} \cdots \binom{n-(r-1)l}{l} = \left(\frac{n^l}{l!}\right)^r \left[1 + O\left(\frac{r}{n}\right)\right] \tag{14}$$

for $n \to \infty$, and for a fixed value of l we obtain

$$(15) \qquad \sum_{\substack{\varkappa_i \in I_l \\ (i=1,2,\dots,r)}} \mathsf{M}(\varepsilon_{\varkappa_1}\varepsilon_{\varkappa_2}\dots\varepsilon_{\varkappa_r}) \sim \left(\frac{N}{n}\right)^{rl}\left(\frac{1}{l}\right)^{r} e^{-2rl\frac{N}{n}}.$$

In such a way it is obtained that uniformly in r

$$(16) \qquad \lim_{\frac{N}{n}\to c} \sum_{\substack{\varkappa_i \in I_l \\ (i=1,2,\dots,r)}} \mathsf{M}(\varepsilon_{\varkappa_1}\varepsilon_{\varkappa_2}\dots\varepsilon_{\varkappa_r}) = \lambda^r,$$

where

$$(17) \qquad \lambda = \frac{(ce^{-2c})^l}{l}.$$

Now clearly each r-tuple $\varkappa_1, \dots, \varkappa_r$ occurs $r!$ times in the sum on the left hand side of (16). If we extended the summation out over all different combinations of order r of elements of $\varkappa$, then the sum is equal that in (16) divided by $r!$. Thus if $\sum\limits_{(\varkappa_1,\dots,\varkappa_r)}$ denotes that the sum is extended out over all different combination of order r of elements of $\varkappa$ we get

$$\lim_{\frac{N}{n}\to c} \sum_{(\varkappa_1,\dots,\varkappa_r)} \mathsf{M}(\varepsilon_{\varkappa_1}\dots\varepsilon_{\varkappa_r}) = \frac{\lambda^r}{r!}.$$

The Lemma 1 of P. ERDŐS—A. RÉNYI in [1] states the following:

Let $\varepsilon_{n_1}, \varepsilon_{n_2}, \dots, \varepsilon_{n_{b_n}}$ be a set of random variables such that ε_{n_i} $(1 \leqq i \leqq b_n)$ takes on only the value 1 or 0. If

$$(18) \qquad \lim_{n\to\infty} \sum_{1\leqq i_1<i_2<\dots<i_r\leqq b_n} \mathsf{M}(\varepsilon_{n_{i_1}}\varepsilon_{n_{i_2}}\dots\varepsilon_{n_{i_r}}) = \frac{\lambda^r}{r!}$$

uniformly in r, $r=1,2,\dots$, where $\lambda>0$, and the summation is extended over all combinations of order r of the integers $1, 2, \dots, b_n$, then

$$(19) \qquad \lim_{n\to\infty} \mathsf{P}\left(\sum_{i=1}^{b_n} \varepsilon_{n_i} = j\right) = \frac{\lambda^j e^{-\lambda}}{j!} \qquad (j = 0, 1, \dots).$$

Applying this Lemma, Theorem 2 follows immediately from (18).

Let us now consider *the proof of Theorem 1.* It can be proved as follows.

As from l given points $(l-1)!$ directed cycles of order l can be formed, we have for fixed l evidently that

$$(20) \qquad \mathsf{M}(\delta_l) = \binom{n}{l}(l-1)!\frac{\binom{n^2-l}{N-l}}{\binom{n^2}{N}} \sim \frac{n^l}{l}\left(\frac{N}{n^2}\right)^l = \frac{1}{l}\left(\frac{N}{n}\right)^l.$$

If we denote by J_l the set of all the possible cycles of order l and to every ν cycles (where $\nu \in J_l$) let us associate a random variable η_ν such that η_ν is equal to 1 or 0 according to that ν is a subgraph of the directed random graph $\vec{\Gamma}_{n,N(n)}$ or it is not.

On the other hand K. JORDÁN's formula states the following: if π_j denotes the probability of that exactly j events occour among the possible $A_1, A_2, \ldots, A_n$ events then

$$\sum_{r=0}^{2s}(-1)^r\binom{j+r}{j}S_{j+r} \leqq \pi_j \leqq \sum_{r=0}^{2s+1}(-1)^r\binom{j+r}{j}S_{j+r}, \tag{21}$$

where

$$S_{j+r} = \sum_{1=i_1<i_2<\cdots<i_{j+r}\leqq n} \mathsf{P}(A_{i_1}A_{i_2}\ldots A_{i_{j+r}}), \tag{22}$$

where the summation is taken over all different combinations of order $j+r$ of the integers $1, 2, \ldots, n$ (for the proof and applications of (21) see [3] or [4]). As in our case the events A_k $(k=1, 2, \ldots, n)$ are equivalent thus the probability $\mathsf{P}(A_{i_1}A_{i_2}\ldots A_{i_{j+r}})$ depends only on $j+r$, that is $\mathsf{P}(A_{i_1}A_{i_2}\ldots A_{i_{j+r}})=\mathsf{P}(A_1A_2\ldots A_{j+r})$. Therefore

$$\mathsf{P}(A_{i_1}A_{i_2}\ldots A_{i_{j+r}}) = \frac{\binom{n^2-(j+r)l}{N-(j+r)l}}{\binom{n^2}{N}} \sim \left(\frac{N}{n^2}\right)^{(j+r)l}. \tag{23}$$

For calculating S_{j+r} we deal with the case when the directed cycles of order l have no common edges, then taking into consideration the asymptotical formula

$$\frac{1}{(j+r)!}\binom{n}{l}\binom{n-l}{l}\cdots\binom{n-(j+r-1)l}{l} \sim \frac{\left(\frac{n^l}{l!}\right)^{(j+r)}}{(j+r)!}. \tag{24}$$

Thus by (22) we obtain

$$S_{j+r} \sim \frac{\left[\frac{1}{l}\left(\frac{N}{n}\right)^l\right]^{j+r}}{(j+r)!}. \tag{25}$$

Put $\frac{N}{n} \sim c$ into (25) thus we have

$$S_{j+r} \sim \frac{c^{l(j+r)}}{l^{(j+r)}(j+r)!}. \tag{26}$$

As for every fixed value of s (21) holds, consequently

$$\lim_{n\to\infty}\pi_j = \sum_{r=0}^{\infty}(-1)^r\binom{r+j}{j}\frac{c^{l(j+r)}}{l^{(j+r)}(j+r)!} = \left(\sum_{r=0}^{\infty}(-1)^r\frac{c^{lr}}{r!\,l^r}\right)\frac{c^{lj}}{j!\,l^j}. \tag{27}$$

Since evidently

$$\sum_{r=0}^{\infty}(-1)^r\frac{\left(\frac{c^l}{l}\right)^r}{r!} = e^{-\frac{c^l}{l}}, \tag{28}$$

thus we obtain

$$(29)\qquad \lim_{n\to\infty} \pi_j = \frac{\left(\frac{c^l}{l}\right)^j}{j!} e^{-\frac{c^l}{l}}.$$

By using the notation $\pi_j = \mathsf{P}_{n,N(n)}\,(\delta_j = j)$ we obtain the desired result, that is

$$(30)\qquad \lim_{n\to\infty} \mathsf{P}_{n,N(n)}(\delta_j = j) = \frac{\left(\frac{c^l}{l}\right)^j}{j!} e^{-\frac{c^l}{l}}.$$

Let us now compare these two couples of theorems. Putting $l=k$ into (9) and (10), then it can be seen that the asymptotical mean values (5) and (7) of the numbers of undirected cycles being the 2^{k-1}-fold of those of directed ones (9) and (10) respectively, supposing that the number of the chosen edges (3) is always the same. This is clear enough if we considerate that in the case of directed random graphs we do not enumerate those configurations in which the directions of the consecutive edges are not suitable. Since a cycle of order k has k edges, and evidently every edge can be directed in two different ways, and since among the whole 2^k possibilities only two yield a well directed k-cycle, thus the number of unsuitable directed cycles is 2^{k-1}. This fact is reflected in the asymptotical mean values too.

Finally we shall prove the following theorem which is analogous to the theorem 5b. of [1].

THEOREM 3. *Let L denote the property that a graph contains at least one directed cycle and if* $\mathsf{P}_{n,N(n)}(L)$ *denotes the probability that the directed random graph* $\vec{\Gamma}_{n,N(n)}$ *possess the property L then we have*

$$(31)\qquad \lim_{n\to\infty} \mathsf{P}_{n,N(n)}(L) = c,$$

supposing that $N(n) \sim cn$ *holds with* $c \leqq 1$. *That is for* $c=1$ *it is "almost sure" that the directed random graph* $\vec{\Gamma}_{n,N(n)}$ *contains at least one directed cycle, but for* $c<1$ *this is not true.*

PROOF. We shall determine the probability of the opposite event and subtract it from 1.

Let $H_{n,N}$ denote the number of all directed cycles contained in $\vec{\Gamma}_{n,N(n)}$. Then for $H_{n,N}$ we obtain

$$(32)\qquad H_{n,N(n)} = \sum_{l=1}^{n} \delta_l,$$

where according to (20)

$$(33)\qquad \mathsf{M}(\delta_l) = \frac{1}{l}\left(\frac{N}{n}\right)^l.$$

By using the supposition $N(n) \sim cn$, thus for $\mathsf{M}(H_{n,N})$ we obtain

$$(34)\qquad \lim_{n\to\infty} \sum_{l=1}^{n} \frac{c^l}{l} = -\log(1-c), \quad \text{for} \quad 0 < c < 1$$

and

$$\sum_{l=1}^{n} \frac{1}{l} \sim \log n, \quad \text{if} \quad c = 1. \tag{35}$$

The probability of $\vec{\Gamma}_{n,N(n)}$ having two directed cycles with common points is negligibly small as we have supposed in theorem 1, therefore we have

$$\lim_{n\to\infty} \mathsf{P}_{n,N(n)}(\bar{L}) = e^{-\lim\limits_{n\to\infty} \mathsf{M}(H_{n,N(n)})} = 1-c, \tag{36}$$

where $\bar{L}$ denotes the opposite of L.

Thus theorem 3 is proved.

REFERENCES

[1] Erdős, P. and Rényi, A.: On the evolution of random graphs. *Publ. of the Math. Inst. of the Hung. Acad. Sci.* **5** (1960) 17—61.

[2] Palásti, I.: On the strong connectedness of directed random graphs. *Studia Scientiarum Mathematicarum Hungarica* **1—2** (1966) 205—214.

[3] Rényi, A.: Egy általános módszer valószínűségszámítási tételek bizonyítására és annak néhány alkalmazása. *A Magyar Tudományos Akadémia Matematikai és Fizikai Tudományok Osztályának Közleményei.* **11** (1961) 79—105.

[4] Rényi, A.: *Valószínűségszámítás*, Tankönyvkiadó, Budapest, 1966.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received December 20, 1968.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 75—79.

ON THE CONNECTIONS BETWEEN APPROXIMATION BY CONVOLUTION AND THE MULTIPLIER PROBLEM

by

L. MÁTÉ

At the end of the fifties a series of papers written by ALEXEIEWITZ and SEMADENI appeared concerning the various types of linear functionals of a two-norm space. A two-norm space is defined as *a linear space X with a limit* generated by two norms $\|\cdot\|$ and $\|\cdot\|_*$ for which $\langle X, \|\cdot\|\rangle$ is a Banach space, $\|x\| \geqq \|x\|_*$ for every $x \in X$ and the sequence $\{x_n\}$ converges to $x \in X$ iff $\{\|x_n\|, n=1,2,\dots\}$ is bounded and $\lim_{n\to\infty} \|x_n - x\|_* = 0$. There are three natural classes of linear functionals in a two-norm space $\langle X, \|\cdot\|, \|\cdot\|_*\rangle$: The dual X^* of the Banach space $\langle X, \|\cdot\|\rangle$, the dual Y^* of the normed space $\langle X, \|\cdot\|_*\rangle$ and the mixed dual $\langle X, \|\cdot\|, \|\cdot\|_*\rangle^*$ i.e. the linear functionals f of X for which $x_n \to x$ in the two-norm limit implies $f(x_n) \to f(x)$; obviously

$$Y^* \subset \langle X, \|\cdot\|, \|\cdot\|_*\rangle^* \subset X^*.$$

If the unit sphere in $\langle X, \|\cdot\|\rangle$ is closed in the topology generated by the weaker norm $\|\cdot\|_*$ then the two-norm space $\langle X, \|\cdot\|, \|\cdot\|_*\rangle$ is called normal.

It follows from the GROETENDICK Theorem ([2] Thm 6. p. 148) that in this case $\langle X, \|\cdot\|, \|\cdot\|_*\rangle^*$ is the closure of Y^* in X^*. The main efforts in the papers of ALEXEIEWITZ and SEMADENI are directed to find conditions for which $Y^* = \langle X, \|\cdot\|, \|\cdot\|_*\rangle^*$ resp. $\langle X, \|\cdot\|, \|\cdot\|_*\rangle^* = X^*$.

The subject of this paper is a continuation of the above mentioned investigations of ALEXEIEWITZ and SEMADENI and the application of the same to a particular two-norm space consisting of continuous functions which seems to be a natural setting for the multiplier problem in Lebesgue spaces.

A bounded linear operator from $L^p(G)$ into $L^q(G)$ where $1 \leqq p, q < \infty$ and G is a locally compact group is called (p, q)-multiplier if it is commuting with translation. A (p, p)-multiplier is called shortly a *p-multiplier*.

If $L^p \otimes L^{p'}$ is the projective (greatest) tensor product of $L^p(G)$ and $L^{p'}(G)$ $\left(\frac{1}{p} + \frac{1}{p'} = 1;\ 1 < p < \infty\right)$, $C_0(G)$ the usual Banach space of the continuous functions on G tending to zero at infinity, then *in the case of Abelian G* there is a continuous homomorphism from $L^p \otimes L^{p'}$ into C_0 defined by

$$(*) \qquad f \hat{\otimes} g \to f * g \qquad (* \text{ is the convolution})$$

The range of this homomorphism with the quotient norm will be denoted by A_p. FIGA-TALAMANCA [6] studied the algebra M_p of the p-multipliers $(1 < p < \infty)$ and established that $M_p = A_p^*$ i.e. there is an isometric isomorphism from M_p onto the dual space of A_p.

If $\|\cdot\|_\infty$ is the uniform norm then it follows from [7] Thm 1 that $\langle A_p, \|\cdot\|, \|\cdot\|_\infty\rangle$ is a normal two-norm space. In this case $Y^* = M(G)$, $X^* = M_p$ and $\langle A_p, \|\cdot\|, \|\cdot\|_\infty\rangle^*$ is the closure of $M(G)$ in M_p i.e. in the operator norm. Hence $\langle A_p, \|\cdot\|, \|\cdot\|_\infty\rangle$ is a natural setting to investigate how to connect bounded measures and p-multipliers.

The organization of this paper is the following: In § 1 new conditions for $Y^* = \langle X, \|\cdot\|, \|\cdot\|_*\rangle^*$ and a new proof for ALEXEIEWITZ and SEMANEDI's original results are given; in § 2 the p-multipliers represented by bounded measures and the subsets of $C_0(G)$ represented in the form of the convolution series

$$\sum_{k=1}^{\infty} f_k * g_k : \sum_{k=1}^{\infty} \|f_k\|_p \|g_k\|_{p'} < \infty \; f_k \in L^p, g_k \in L^{p'}$$

are characterized; in § 3, with the same two-norm space technique, a proof for the impossibility of the COHEN—HEWITT factorization for the convolution algebra $L^p(G)$ (G compact and $p>1$) is given; in § 4 a characterization of amenable groups by p-multipliers is given by an easy adaptation of the results of H. LEPTIN's paper [12].

§ 1. Let Y be the completion of $\langle X, \|\cdot\|_*\rangle$. There is a continuous injection T from $\langle X, \|\cdot\|\rangle$ into Y hence, the dual operator T^* is a continuous injection with $\sigma(X^*, X)$-dense range. A two-norm space is called *trivial* if $X = Y$ evidently, in this case both T and T^* are topological isomorphisms.

THEOREM. *The following assertions are equivalent:*

I. *The two-norm space $\langle X, \|\cdot\|, \|\cdot\|_*\rangle$ is trivial.*

II. *Y is the completion of $\langle X, \|\cdot\|, \|\cdot\|_*\rangle$.*

III. *T^*Y^* is closed in X^*.*

IV. *$T^*Y^* = X^*$.*

PROOF. I$\Rightarrow$II is obvious. II$\Rightarrow$III: It follows from II. that for every $y \in Y$ there exists $K>0$ and sequence $\{x_\alpha\}$ such that $\|x_\alpha\| < K$ and $Tx_\alpha \to y$. Now, let $\{T^* y_n^*\}$ be convergent; then from

$$|\langle y_n^* - y_m^*, y\rangle| \leqq \sup_\alpha |\langle y_n^* - y_m^*, Tx_\alpha\rangle| = \sup_\alpha |\langle T^*(y_n^* - y_n^*), x_\alpha\rangle| \leqq K\|T^* y_n^* - T^* y_m^*\|$$

it follows that there exists a $y^* \in Y$ such that $\sigma - \lim y_n^* = y^*$ and hence $\sigma - \lim T^* y_n^* = T^* y^*$. Consequently, $\lim T^* y_n^* = T^* y^*$ also in the norm topology.

III$\Rightarrow$IV: Since T^*Y^* is σ-dense, from the KREIN—SMULJAN theorem ([2] p. 152.) it follows that we have only to show that $T^*Y^* \cap S^*$ is σ-closed if S^* is the unit sphere in X^*.

If $T^* y^* \in S^*$ and $\{\langle x, T^* y_\alpha^*\rangle\}$ is convergent for every $x \in X$, then from

$$\langle x, T^* y_\alpha^*\rangle = \langle Tx, y_\alpha^*\rangle$$

it follows that $\{y_\alpha^*\}$ is pointwise convergent on a dense subset of Y. Moreover, $\{y_\alpha^*\}$ is bounded since it follows from the Banach homeomorphism theorem that T^{*-1} carries bounded set into bounded set. Consequently, from the Banach—Steinhaus theorem it follows that there exists $y^* \in Y^*$ such that $\sigma - \lim y_\alpha^* = y^*$. Moreover,

$$\lim_\alpha \langle x, T^* y_\alpha^*\rangle = \lim_\alpha \langle Tx, y_\alpha^*\rangle = \langle Tx, y^*\rangle = \langle x, T^* y^*\rangle$$

for every $x \in X$.

IV⇒I: It follows from IV. and the Banach homeomorphism theorem that the continuous linear operator T^* is a topological isomorphism and hence $T^{**}X = TX$ is closed in Y. On the other hand TX is dense in Y by definition.

Remark 1. It would seem that T^*Y^* is dense in X^* in any two-norm space, however this is not the case at all. Indeed, the closure of T^*Y^* is included in the dual of the locally convex space $\langle X, \|\cdot\|, \|\cdot\|_*\rangle$ which is identical with X^* only in particular cases. (See [11] p. 290.)

Remark 2. The *Theorem* is valid for locally convex and metrizable X and Y too with similar proof.

§ **2.** Let G be a locally compact Abelian group.

If $M(G)$ is the usual Banach space of bounded measures on a locally compact G then evidently $M \subset A_p^*$ moreover, $M_1 = M$ (see e.g. [7]) i.e. every multiplier in L^1 is a convolution with a certain $\mu \in M$ and there is an isometric isomorphism from M onto M_1 defined by $\mu \to T_\mu$ (T_μ is the operator presented as the convolution with μ). If $1 < p < \infty$ then the situation is quite different. It is shown in [8] that if $1 \leqq p < q < \infty$ then $M_p \subset M_q$ and the inclusion is proper. Particularly, $M \neq M_p$ for any $1 < p < \infty$. Thus, it is natural to ask the following question motivated also by [4]:

Let G be a locally compact group and $1 < p < \infty$. Let us characterize the closed subsets $P \subset G$ for which every $F \in A_p^*$ such that supp $F \subset P$ is a bounded measure.

The main subject of this § is to give an answer to this problem by a simple application of the *Theorem*.

Let A_p^P be the range with quotient norm of the continuous homomorphism from $L^p \hat{\otimes} L^{p'}$ into $C_0(P)$ defined by $f \otimes g \to f * g|_P$ where $f * g|_P$ is the restriction of $f * g$ to P. If we apply the *Theorem* to the two-norm space $\langle A_p^P, \|\cdot\|, \|\cdot\|_\infty\rangle$ then we obtain

Proposition: The following assertions are equivalent:

I. If $h \in C_0(P)$ then there exist $f_k \in L^p$, $g_k \in L^{p'}$ $k = 1, 2, \dots$ such that

$$\sum_{k=1}^{\infty} \|f_k\|_p \|g_k\|_{p'} < \infty \quad \text{and} \quad h = \sum_{k=1}^{\infty} f_k * g_k \quad \text{for } t \in P.$$

II. If $h \in C_0(P)$ then there exist $K > 0$ and sequence $\{h_n\}$ $h_n \in A_p$ such that $\|h_n\| \leqq K$ and $\{h_n\}$ converges on P *uniformly* to h.

III. $\{F_\mu;\ \mu \in M(P)\}$ i.e. $M(P)$ as a subspace of A_p^* is closed.

IV. If $F \in A_p^*$ and supp $F \subset P$ then F is a bounded measure I.e. $(A_p^P)^* = M(P)$.

Remarks: 1. The assertion II⇒III is essentially the same as Theorem 1. 4. in [4]. Moreover, we learn from the Proposition that III implies the stronger assertion I., too. I.e. in this case every $h \in C_0(P)$ is not only uniformly approximated on P by convolutions but h is given in the form of a convolution series.

2. If $p = 2$ then we get [1] Theorem *5. 6. 3.* on the characterization of Helson sets. Indeed* $A(G)$ consists precisely of the convolutions $f_1 * f_2$ with f_1 and f_2 in

* for the notations see [1]

$L^2(G)$ ([1] Thm 1. 6. 3.), hence $A \subset A_2$. On the other hand from

$$\|f_1 * f_2\|_A = \|\hat{f}_1 \hat{f}_2\|_1 \leqq \|f_1\|_2 \|f_2\|_2$$

it follows that if $\sum_{k=1}^{\infty} \|f_k\|_2 \|g_k\|_2 < \infty$ then $\sum_{k=1}^{\infty} f_k * g_k$ is a Cauchy sequence in A and hence $\sum_{k=1}^{\infty} f_k * g_k \in A$. Consequently, A and A_2 consist of the same functions.

It follows from a similar argument that also

$$\left\| \sum_{k=1}^{\infty} f_k * g_k \right\|_A \leqq \left\| \sum_{k=1}^{\infty} f_k * g_k \right\|_{A_2}$$

hence it follows from the Banach homeomorphism theorem that the norms $\|\cdot\|_A$ and $\|\cdot\|_{A_2}$ are equivalent.

Finally, it follows from the definition of $A(G)$ that $A(G)^*$ is isometrically isomorphic with $L^\infty(\Gamma)$.

3. The results of this § is valid also for amenable groups (i.e. for locally compact groups with invariant mean) with a slight modification.

RIEFFEL [9] proved that $M_p = A_p^*$ for any compact group G if — instead of (*) — A_p is the range of the continuous homomorphism

$$f \hat{\otimes} g \to \tilde{f} * g \quad (\tilde{f}(t) = f(t^{-1})).$$

Moreover, C. S. HERZ proved that $M_p = A_p^*$ also for any amenable group G if A_p is the range of the continuous homomorphism

$$f \hat{\otimes} g \to \Delta(\cdot)^{1/p'} \tilde{f} * g$$

where Δ is the modular function and $\tilde{f}(t) = \Delta(t^{-1}) f(t^{-1})$ (see e.g. [9]). Hence, *for the case of an amenable group G any of I—III in the Proposition characterizes p-multipliers represented by bounded measures.*

§ 3. Let G be compact and $1 \leqq p, q < \infty$. Then each L^q is an L^p-modul and it follows from the COHEN—HEWITT factorization theorem (see e.g. [10]) that every $h \in L^q$ can be written in the form

$$h = f * g \qquad f \in L^1, \quad g \in L^q.$$

In the following it will be proved as an application of the I⇒IV part of the Theorem that *a similar factorization is far from being true for* $p \neq 1$.

If A_p^q is the range — supplied with the quotient norm — of the continuous homomorphism from $L^p \hat{\otimes} L^q$ into L^q defined by $f \hat{\otimes} g \to \tilde{f} * g$ then by applying the *Theorem* to the two-norm space $\langle A_p^q, \|\cdot\|, \|\cdot\|_q \rangle$ the following assertions are found to be equivalent:

I′. If $h \in L^q$ then there exist $f_k \in L^p, g_k \in L^q$ $k = 1, 2, \ldots$ such that $\sum_{k=1}^{\infty} \|f_k\|_p \|g_k\|_q < \infty$ and $h = \sum_{k=1}^{\infty} f_k * g_k$ in the sense of L^q-convergence.

IV′. Every $F \in (A_p^q)^*$ can be represented by a certain $f_F \in L^{q'}$.

Now, in order to prove the impossibility of the COHEN—HEWITT factorization it is enough to prove that IV′ is not true for $p \neq 1$. Moreover, on the basis of a result of RIEFFEL [9] — asserting that the Banach space M_p^q of (p, q) multipliers is isometrically isomorphic to $(A_p^q)^*$ — in the case when G is compact, it is enough to prove that there are (p, q')-multipliers which cannot be represented as a convolution by a $L^{q'}$-function. Indeed, it is well-known (see e.g. [9]) that if $f \in L^p$, $g \in L^r$ and $\frac{1}{p}+\frac{1}{r}-\frac{1}{q'}=1$ then $f*g\ L^{q'}$; on the other hand, it is evident that $\frac{1}{r}-\frac{1}{q'}>0$ hence $r<q'$. Consequently there exists an f, $f \notin L^r$, $f \notin L^{q'}$ which is a (p, q')-multiplier. Q.E.D.

Finally we remark that a similar assertion as the *Proposition* can be established for (p, q)-multipliers also in the case of $p \neq q$ and for any locally compact G.

§ 4. It follows easily from the results of H. LEPTIN [12] that the following assertions are equivalent for a locally compact group G:

I. G is amenable i.e. there is an invariant mean in G.

II. For every positive definit function $h \in C_0(G)$ there exists $K>0$ and sequence $\{h_n\}$, $h_n = f_n * \tilde{f}_n$ where f_n is a continuous function with compact support, such that $\|f_n\|_2 < K$ and $\{h_n\}$ converges to h uniformly on every compact subset.

III. If $\|T_f\|_p$ is the norm of $f \in L^1(G)$ as a convolution operator in $L^p(G)$ then

$$\|T_f\|_p = \|f\| \quad \text{for every positive} \quad f \in L^1(G).$$

IV. If T_μ is an operator of $L^p(G)$ represented by the convolution with a positive measure μ, then μ is a bounded measure.

If we compare I and IV with a theorem of BRAINERD and EDWARDS which asserts that every positive p-multiplier is represented as the convolution with a positive measure, then we obtain the following characterization of amenable groups:

The locally compact group G is amenable iff every positive p-multiplier on G is represented by the convolution with bounded *measure.*

REFERENCES

[1] RUDIN, W.: *Fourier Analysis on Groups*, Interscience Publ. 1962.

[2] SCHAEFER, H. H.: *Topological Vector Spaces*, Macmillan Co. 1966.

[3] ALEXEIEVITZ, A. and SEMADENI, W.: Linear functionals on two-norm spaces, *Studia Math.* **17** (1958), 121—140.

[4] EDWARDS, R. E.: Approximation by Convolutions, *Pac. J. Math.* **15** (1965), 85—95.

[5] EYMARD, P.: L'algébre de Fourier d'un groupe localement compact, *Bull. Soc. Math. France* **92** (1964), 181—236.

[6] FIGA-TALAMANCA, A.: Translation-invariant operators in L^p, *Duke Math. J.* **32** (1965), 495—502.

[7] FIGA-TALAMANCA, A. and GAUDRY, G. I.: Density and representation theorems for multipliers of Type (p, q), *J. Austr. Math. Soc.* **7** (1967), 1—9.

[8] FIGA-TALAMANCA, A. and GUDRY, G. I.: Multipliers and set of uniqueness of L^p, *to appear.*

[9] RIEFFEL, M. A.: Multipliers and tensor products of L^p-spaces of locally compact groups. *Studia Math.* **38** (1969), 71—82.

[10] CURTIS, P. C. and FIGA-TALAMANCA, A.: Factorization Theorems for Banach algebras. *Functions Algebras* (Editor: F. T. Birtel) p. 169—186.

[11] ALEXEIEVITZ, A. and SEMADENI, W.: The two-norm spaces and their conjugate spaces. *Studia Math.* **18** (1959), 275—296.

[12] LEPTIN, H.: On locally compact groups with invariant means. *Proc. Amer. Math. Soc.* **19** (1968), 489—494.

Polytechnical University of Budapest
(Received April 18, 1969. Revised April 13, 1970.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 81—90.

VERALLGEMEINERUNG EINES SATZES VON A. RÉNYI

von

G. MÖNCH

Es sei $[A, \varrho_A]$ ein vollständiger separabler metrischer Raum, $\mathfrak{A}$ die σ-Algebra der Borelmenge, $\mathfrak{B}$ der Ring der beschränkten X aus $\mathfrak{A}$ und N die Menge derjenigen Maße ν auf $\mathfrak{A}$, die auf $\mathfrak{B}$ endlich sind.

Ein Φ aus N heißt Punktfolge, wenn Φ als Summe einer abzählbaren Familie $(\delta_{a_i})_{i\in I}$ von δ-Maßen dargestellt werden kann. Sind in einer und damit jeder Darstellung dieser Art die a_i paarweise verschieden, so heißt Φ einfach. Es bezeichne M die Menge aller Punktfolgen und $\mathfrak{M}$ die kleinste σ-Algebra von Teilmengen von M, bezüglich der für alle X aus $\mathfrak{B}$ die reelle Funktion

$$\Phi \curvearrowright \Phi(X) \qquad (\Phi \in M)$$

meßbar ist. Die Menge E aller einfachen Punktfolgen liegt in $\mathfrak{M}$. Greifen wir uns nämlich zu jedem natürlichen m eine abzählbare Überdeckung $(X_{m,n})_{n=1,2,\dots}$ von A durch paarweise disjunkte, in $\mathfrak{A}$ liegende Mengen mit $\frac{1}{m}$ nicht übersteigenden Durchmessern heraus, so gilt

$$\bar{E} = \bigcup_{k=1}^{\infty} \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \{\Phi : \Phi \in M; \Phi(X_{m,n} \cap S_k(z)) \geqq 2\},$$

wobei $S_k(z)$ die offene Vollkugel um den beliebigen, aber festen Punkt z aus A mit dem Radius k bezeichnet. Jedem ν aus N kann in Gestalt des Poissonschen Verteilungsgesetzes P_ν mit dem Intensitätsmaß ν ein wohlbestimmtes Verteilungsgesetz auf $\mathfrak{M}$ zugeordnet werden.

Ein Poissonsches Verteilungsgesetz genügt genau dann der Bedingung

$$P_\nu(\Phi \text{ ist einfach}) = 1,$$

wenn

$$\nu(\{a\}) = 0 \qquad (a \in A)$$

erfüllt ist.

Für jede endliche Folge $X_1, \dots, X_n$ von Mengen aus $\mathfrak{B}$ und jedes Verteilungsgesetzt P auf $\mathfrak{M}$ bezeichnen wir das bezüglich P gebildete Verteilungsgesetz von $[\Phi(X_1), \dots, \Phi(X_n)]$ mit $P_{X_1,\dots,X_n}$. Zwei Verteilungsgesetze P, Q auf $\mathfrak{M}$ sind gleich, falls für alle endlichen Folgen $X_1, \dots, X_n$ paarweise disjunkter Mengen aus $\mathfrak{B}$ die Gleichung

$$P_{X_1,\dots,X_n} = Q_{X_1,\dots,X_n}$$

erfüllt ist.

A. RÈNYI bewies in [1], daß im Spezialfall $P=P_\nu$, $\nu\in N$ schon aus der Gültigkeit von

$$P_X=Q_X$$

für hinreichend viele X aus $\mathfrak{B}$ auf $P=Q$ geschlossen werden kann, falls

$$\nu(\{a\})=0 \qquad (a\in A)$$

gilt. Es zeigt sich nun, daß diese Aussage nicht an eine spezielle Gestalt von P gebunden ist:[1]

SATZ. *Es seien P, Q Verteilungsgesetze auf $\mathfrak{M}$. Gilt nun*

$$P_X=Q_X$$

für alle X aus einem Teilring $\mathfrak{X}$ von $\mathfrak{B}$, so kann auf

$$P=Q$$

geschlossen werden, falls folgende Bedingungen erfüllt sind:

a) *Es ist* $P(\Phi$ ist einfach$)=1$,

b) *Der kleinste $\mathfrak{X}$ umfassende σ-Teilring von $\mathfrak{B}$ ist $\mathfrak{B}$.*

Hierbei nennen wir einen Teilring $\mathfrak{v}$ von $\mathfrak{B}$ einen σ-Teilring von $\mathfrak{B}$, wenn für alle Folgen (Y_n) von Mengen aus $\mathfrak{v}$ mit der Eigenschaft $Y=\bigcup_{n=1}^{\infty} Y_n\in\mathfrak{B}$ die Beziehung $Y\in\mathfrak{v}$ erfüllt ist.

Den Beweis führen wir in mehreren Schritten.

1. Es bezeichne $\mathfrak{S}$ das System derjenigen X aus $\mathfrak{B}$, für die

$$P_X=Q_X$$

gilt. Für jede absteigende Folge (X_n) von Mengen aus $\mathfrak{S}$ erhalten wir für alle Φ aus M

$$\Phi(X_n)\xrightarrow[n\to\infty]{}\Phi(X),$$

wenn $X=\bigcap_{n=1}^{\infty} X_n$ gesetzt wird. Für $m=0, 1, \dots$ kann hieraus auf

$$\begin{aligned} P_X(\{m\}) &= P\big(\Phi(X)=m\big)=\lim_{n\to\infty} P\big(\Phi(X_n)=m\big)\\ &=\lim_{n\to\infty} P_{X_n}(\{m\})=\lim_{n\to\infty} Q_{X_n}(\{m\})\\ &=\lim_{n\to\infty} Q\big(\Phi(X_n)=m\big)=Q\big(\Phi(X)=m\big)\\ &=Q_X(\{m\}),\ \text{d.h. auf}\end{aligned}$$

$$\bigcap_{n=1}^{\infty} X_n\in\mathfrak{S}$$

geschlossen werden.

Entsprechend zeigt man für jede in $\mathfrak{B}$ nach oben beschränkte, aufsteigende Folge (Y_n) von Mengen aus $\mathfrak{S}$

$$\bigcup_{n=1}^{\infty} Y_n\in\mathfrak{S}.$$

---

[1] Vgl [1], [2], [3].

Das $\mathfrak{X}$ umfassende System $\mathfrak{S}$ ist also in $\mathfrak{B}$ monoton abgeschlossen und fällt daher[1] auf Grund von b) mit $\mathfrak{B}$ zusammen. Wir können daher von jetzt ab o. B. d. A. zusätzlich annehmen, es sei $\mathfrak{X} = \mathfrak{B}$.

2. Wir wollen in diesem Beweisschritt zwei spezielle Abbildungen von $[M, \mathfrak{M}]$ in sich untersuchen, die wir im weiteren Beweis benötigen.

Wir ordnen jedem Φ aus M und jedem $\eta > 0$ die durch „Verdünnung" entstehende einfache Punktfolge

$$_\eta\Phi = \Phi - \sum_{\substack{\Phi(\{a\})>0 \\ \Phi(S_\eta(a))>1}} \Phi(\{a\})\delta_a$$

zu, wobei $S_\eta(a)$ die offene Vollkugel um a mit dem Radius η bezeichnet. Für alle $\eta > 0$ vermittelt

$$\Phi \curvearrowright {_\eta\Phi} \qquad (\Phi \in M)$$

eine meßbare Abbildung von $[M, \mathfrak{M}]$ in sich. Auf Grund der Definition der σ-Algebra $\mathfrak{M}$ reicht es aus, die Meßbarkeit der reellen Funktion

$$\Phi \curvearrowright {_\eta\Phi}(X)$$

für alle X aus $\mathfrak{B}$ nachzuweisen. Wir zerlegen den Phasenraum $[A, \varrho_A]$ in abzählbar viele, paarweise disjunkte Mengen $(Z_{n,k})_{k=1,2,\dots}$ aus $\mathfrak{A}$ mit $\frac{1}{n}$ unterschreitenden Durchmessern.

Es sei X eine beliebige Menge aus $\mathfrak{B}$. Wir können o. B. d. A. die Mengen $Z_{n,k}$ so wählen, daß $Z_{n,k} \cap X$ gleich $Z_{n,k}$ oder die leere Menge für alle n, k ist. Die Menge dieser $Z_{n,k}$ bezeichnen wir mit $\mathfrak{z}_n$. Nun bilden wir die Folge $(f_n(\Phi))_{n=1,2,\dots}$ der Funktionen $f_n(\Phi) = \sum_{\{k: Z_{n,k} \subseteq X\}} f_{n,k}(\Phi)$ mit

$$f_{n,k}(\Phi) = \begin{cases} 0 & \text{für} \quad \Phi(Z_{n,k}) = 0 \\ 0 & \text{für} \quad \Phi(Z_{n,k}) > 1 \\ 0 & \text{für} \quad \Phi(Z_{n,k} = 1) \quad \text{und} \quad \Phi(Z) > 0 \text{ für mindestens} \\ & \text{ein} \quad Z \in \mathfrak{z}_n \quad \text{mit} \quad \varrho_A(Z_{n,k}, Z) < \eta - \frac{2}{n} \\ 1 & \text{für} \quad \Phi(Z_{n,k}) = 1 \quad \text{und} \quad \Phi(Z) = 0 \quad \text{für alle} \quad Z \in \mathfrak{z}_n \\ & \text{mit} \quad \varrho_A(Z_{n,k}, Z) < n - \frac{2}{n}, \end{cases}$$

wobei mit $\varrho_A(Z_{n,k}; Z)$ der Abstand der Mengen $Z_{n,k}$ und Z bezeichnet wird. Die Funktionen $f_{n,k}(\Phi)$ sind für alle n, k meßbar, und damit ist auch $f_n(\Phi)$ für $n = 1, 2, \dots$ meßbar. Es gilt aber

$$f_n(\Phi) \to {_\eta\Phi}(X) \quad \text{für} \quad n \to \infty \quad \text{und festes } \Phi.$$

[1] Aussage und Beweis von Theorem 2 aus [2], § 6 bleiben ja gültig, wenn an Stelle des vollen Potenzmengenverbandes der bedingt σ-vollständige Ring $\mathfrak{B}$ benutzt wird.

Daraus folgt aber, daß die reelle Funktion

$$\Phi \curvearrowright_\eta \Phi(X)$$

für alle X aus $\mathfrak{B}$ meßbar ist.

Wir ordnen jedem Φ aus M die einfache Punktfolge

$$\Phi^* = \sum_{\Phi(\{a\})>0} \delta_a$$

zu. Die Abbildung

$$\Phi \curvearrowright \Phi^*$$

von $[M, \mathfrak{M}]$ in sich ist meßbar.

Es reicht wiederum aus, die Meßbarkeit der reellen Funktion

$$\Phi \curvearrowright \Phi^*(X)$$

für alle X aus $\mathfrak{B}$ nachzuweisen. Es sei X eine beliebige Menge aus $\mathfrak{B}$. Zu jedem natürlichen n wählen wir eine abzählbare Überdeckung $(X_{n,k})_{k=1,2,\dots}$ von X durch paarweise disjunkte Mengen aus $\mathfrak{B}$ mit $\frac{1}{n}$ unterschreitenden Durchmessern aus. Wir bilden die Folge $\big(f_n(\Phi)\big)_{n=1,2,\dots}$ von Funktionen

$$f_n(\Phi) = \sum_{k=1}^{\infty} f_{n,k}(\Phi)$$

mit

$$f_{n,k}(\Phi) = \begin{cases} 0 & \text{für} \quad \Phi(X_{n,k}) = 0 \\ 1 & \text{für} \quad \Phi(X_{n,k}) > 0. \end{cases}$$

Jede dieser Funktionen $f_{n,k}(\Phi)$ ist meßbar und damit auch $f_n(\Phi)$ für $n=1, 2, \dots$. Nun gilt aber $f_n(\Phi) \to \Phi^*$ für $n \to \infty$, und unsere Behauptung ist bewiesen.

3. Für jede nichtleere endliche Familie $(X_i)_{i\in I}$ paarweise disjunkter Mengen aus $\mathfrak{B}$ und jede Teilmenge J von I gilt

$$\begin{aligned} &P\big(\Phi(X_i) \geqq 1 \quad \text{für} \quad i \in J, \ \Phi(X_i) = 0 \quad \text{für} \quad i \in I \setminus J\big) \\ &= Q\big(\Phi(X_i) \geqq 1 \quad \text{für} \quad i \in J, \ \Phi(X_i) = 0 \quad \text{für} \quad i \in I \setminus J\big). \end{aligned}$$

Wir erbringen den Nachweis durch vollständige Induktion nach der Anzahl n von J.

Für $n=0$ ist

$$\begin{aligned} &P\big(\Phi(X_i) \geqq 1 \quad \text{für} \quad i \in J, \Phi(X_i) = 0 \quad \text{für} \quad i \in I \setminus J\big) = \\ &= P\Big(\Phi\Big(\bigcup_{i\in I} X_i\Big) = 0\Big) = P_{\bigcup_{i\in I} X_i}(\{0\}) = Q_{\bigcup_{i\in I} X_i}(\{0\}) = Q\Big(\Phi\Big(\bigcup_{i\in I} X_i\Big) = 0\Big) = \\ &= Q\big(\Phi(X_i) \geqq 1 \quad \text{für} \quad i \in J, \Phi(X_i) = 0 \quad \text{für} \quad i \in I \setminus J\big). \end{aligned}$$

Wir wollen jetzt annehmen, unsere Behauptung sei für n_0 gültig. Es sei J eine Teilmenge von I mit der Anzahl $n_0 + 1$ .Wir greifen nun irgendein j_0 aus J heraus. Voraussetzungsgemäß ist dann

$$\begin{aligned} &P\big(\Phi(X_i) \geqq 1 \quad \text{für} \quad i \in J, \Phi(X_i) = 0 \quad \text{für} \quad i \in I \setminus J\big) = \\ &= P\big(\Phi(X_i) \geqq 1 \quad \text{für} \quad i \in J \setminus \{j_0\}, \Phi(X_i) = 0 \quad \text{für} \quad i \in (I \setminus \{j_0\}) \setminus (J \setminus \{j_0\})\big) - \\ &- P\big(\Phi(x_i) \geqq 1 \quad \text{für} \quad i \in J \setminus \{j_0\}, \Phi(X_i) = 0 \quad \text{für} \quad i \in I \setminus (J \setminus \{j_0\})\big) \end{aligned}$$

gleich

$$Q\big(\Phi(X_i) \geqq 1 \quad \text{für} \quad i \in J \setminus \{j_0\}, \Phi(X_i) = 0 \quad \text{für} \quad i \in (I \setminus \{j_0\}) \setminus (J \setminus \{j_0\})\big) -$$
$$- Q\big(\Phi(X_i) \geqq 1 \quad \text{für} \quad i \in J \setminus \{j_0\}, \Phi(X_i) = 0 \quad \text{für} \quad i \in I \setminus (J \setminus \{j_0\})\big),$$

d.h., es gilt

$$P\big(\Phi(X_i) \geqq 1 \quad \text{für} \quad i \in J, \Phi(X_i) = 0 \quad \text{für} \quad i \in I \setminus J\big) =$$
$$= Q\big(\Phi(X_i) \geqq 1 \quad \text{für} \quad i \in J, \Phi(X_i) = 0 \quad \text{für} \quad i \in I \setminus J\big).$$

4. Für alle X aus $\mathfrak{B}$ und alle $\varepsilon > 0$ existiert ein $\eta > 0$, so daß für jede abzählbare Überdeckung $(U_i)_{i \in I}$ von X durch Mengen aus $X \cap \mathfrak{B}$ mit η unterschreitenden Durchmessern die Ungleichung

$$P\big(\Phi(U_i) > 1 \text{ für mindestens ein } i \in I\big) < \varepsilon$$

erfüllt ist.

Zum Beweis ordnen wir jedem Φ aus M und jedem $\eta > 0$ die durch „Verdünnung" entstehende einfache Punktfolge

$${}_\eta\Phi = \Phi - \sum_{\substack{\Phi(\{a\}) > 0 \\ \Phi(S_\eta(a)) > 1}} \Phi(\{a\})\delta_a$$

zu. Die Meßbarkeit der Abbildung

$$\Phi \curvearrowright {}_\eta\Phi \qquad (\Phi \in M)$$

für alle $\eta > 0$ wurde unter 2. gezeigt.

Für alle einfachen Φ gilt

$$\Phi(X) = \lim_{\eta \to 0+0} {}_\eta\Phi(X)$$

Somit ist auf Grund von a)

$$P\big({}_\eta\Phi(X) \neq \Phi(X)\big) \xrightarrow[\eta \to 0+0]{} 0$$

erfüllt. Für jede abzählbare Überdeckung $(U_i)_{i \in I}$ von X durch Mengen aus $X \cap \mathfrak{B}$ mit η unterschreitenden Durchmessern gilt aber

$$P\big(\Phi(U_i) > 1 \text{ für mindestens ein } i \in I\big) \leqq P\big({}_\eta\Phi(X) \neq \Phi(X)\big),$$

und unsere Aussage ist bewiesen.

5. Unter der zusätzlichen Voraussetzung

$$Q(\Phi \text{ ist einfach}) = 1$$

gilt

$$P = Q.$$

Es sei $X_1, \ldots, X_n$ eine endliche Folge paarweise disjunkter Mengen aus $\mathfrak{B}$. Eine abzählbare Überdeckung $(U_i)_{i \in I}$ von $X = \bigcup_{i=1}^{n} X_i$ durch paarweise disjunkte Mengen aus $X \cap \mathfrak{B}$ sei so gewählt, daß jedes X_s in der Gestalt

$$X_s = \bigcup_{i \in I_s} U_i \qquad (s = 1, \ldots, n)$$

dargestellt werden kann. Für alle nichtnegativen ganzen $m_1, \ldots, m_n$ erhalten wir

$$P\big(\Phi(X_s)=m_s \text{ für } 1 \leqq s \leqq n, \quad \Phi(U_i) \leqq 1 \text{ für } i \in I\big)$$

$$= \sum_{\substack{J \subseteq I \\ \text{Anz}(J \cap I_s)=m_s \text{ für } 1 \leqq s \leqq n}} P\big(\Phi(U_i) \geqq 1 \text{ für } i \in J, \quad \Phi(U_i)=0 \text{ für } i \in I \setminus J\big)$$

$$- \sum_{\substack{J \subseteq I \\ \text{Anz}(J \cap I_s)=m_s \text{ für } 1 \leqq s \leqq n}} P\big(\Phi(U_i) \geqq 1 \text{ für } i \in J, \quad \Phi(U_i)=0 \text{ für } i \in I \setminus J,$$

$$\Phi(U_i) > 1 \text{ für mindestens ein } i \in J\big).$$

Nun sind aber die Mengen

$\{\Phi : \Phi(U_i) \geqq 1$ für $i \in J$, $\Phi(U_i)=0$ für $i \in I \setminus J$, $\Phi(U_i) > 1$ für mindestens ein $i \in J\}$

paarweise disjunkt und in

$$\{\Phi : \Phi(U_i) > 1 \text{ für mindestens ein } i \in I\}$$

enthalten. Somit ergibt sich

$$\Big| P\big(\Phi(X_s)=m_s \text{ für } 1 \leqq s \leqq n, \quad \Phi(U_i) \leqq 1 \text{ für } i \in I\big)$$

$$- \sum_{\substack{J \subseteq I \\ \text{Anz}(J \cap I_s)=m_s \text{ für } 1 \leqq s \leqq n}} P\big(\Phi(U_i) \geqq 1 \text{ für } i \in J, \quad \Phi(U_i)=0 \text{ für } i \in I \setminus J\big)\Big|$$

$$\leqq P\big(\Phi(U_i) > 1 \text{ für mindestens ein } i \in I\big)$$

und folglich

$$\Big| P\big(\Phi(X_s) = m_s \text{ für } 1 \leqq s \leqq n\big) - \sum_{\substack{J \leqq I \\ \text{Anz}(J \cap I_s)=m_s \text{ für } 1 \leqq s \leqq n}} P\big(\Phi(U_i) \geqq$$

$$\geqq 1 \text{ für } i \in J, \Phi\Big(\bigcup_{i \in I \setminus J} U_i\Big) = 0\Big)\Big|$$

$$\leqq 2P\big(\Phi(U_i) > 1 \text{ für mindestens ein } i \in I\big).$$

Eine entsprechende Ungleichung gilt auch für Q, so daß wir mit 3. auf

$$\big|P\big(\Phi(X_s)=m_s \text{ für } 1 \leqq s \leqq n\big) - Q\big(\Phi(X_s)=m_s \text{ für } 1 \leqq s \leqq n\big)\big|$$

$$\leqq 2\big[P\big(\Phi(U_i) > 1 \text{ für mindestens ein } i \in I\big)$$

$$+ Q\big(\Phi(U_i) > 1 \text{ für mindestens ein } i \in I\big)\big]$$

schließen können.

Die rechte Seite dieser Ungleichung kann vermöge 4. beliebig klein gemacht werden, indem $(U_i)_{i \in I}$ so ausgewählt wird, daß die Durchmesser der U_i sämtlich hinreichend klein sind. Es ist also

$$P\big(\Phi(X_s)=m_s \text{ für } 1 \leqq s \leqq n\big) = Q\big(\Phi(X_s)=m_s \text{ für } 1 \leqq s \leqq n\big),$$

d.h., es gilt

$$P_{X_1, \ldots, X_n} = Q_{X_1, \ldots, X_n}$$

für alle endlichen Folgen $X_1, \ldots, X_n$ paarweise disjunkter Mengen aus $\mathfrak{B}$.

6. Es bleibt nur noch nachzuweisen, daß die zusätzliche Voraussetzung aus 5. stets erfüllt ist. Jedem Φ aus M ordnen wir die einfache Punktfolge

$$\Phi^* = \sum_{\Phi(\{a\})>0} \delta_a$$

zu. Die Meßbarkeit der Abbildung

$$\Phi \curvearrowright \Phi^*$$

wurde unter 2. gezeigt. Ein Verteilungsgesetz Q auf $\mathfrak{M}$ wird durch diese Abbildung in ein Verteilungsgesetz Q^* auf $\mathfrak{M}$ mit der Eigenschaft

$$Q^*(\Phi \text{ ist einfach}) = 1$$

überführt. Für alle X aus $\mathfrak{B}$, alle nichtnegativen ganzen m und alle abzählbaren Überdeckungen $(U_i)_{i\in I}$ von X durch paarweise disjunkte Mengen aus $X \cap \mathfrak{B}$ erhalten wir mit Hilfe der Schlüsse aus 5.

$$\Big| Q^*(\Phi(X) = m) - \sum_{\substack{J \subseteq I \\ \text{Anz } J = m}} Q^*\Big(\Phi(U_i) \geqq 1 \quad \text{für} \quad i \in J, \Phi\Big(\bigcup_{i \in I \setminus J} U_i\Big) = 0\Big)\Big|$$
$$\leqq 2Q^*(\Phi(U_i) > 1 \text{ für mindestens ein } i \in I).$$

Nun ist aber stets

$$Q^*\Big(\Phi(U_i) \geqq 1 \quad \text{für} \quad i \in J, \Phi\Big(\bigcup_{i \in I \setminus J} U_i\Big) = 0\Big) =$$
$$= Q\Big(\Phi(U_i) \geqq 1 \quad \text{für} \quad i \in J, \Phi\Big(\bigcup_{i \in I \setminus J} U_i\Big) = 0\Big),$$

und es ergibt sich

$$\Big| Q^*(\Phi(X) = m) - \sum_{\substack{J \subseteq I \\ \text{Anz } J = m}} Q\Big(\Phi(U_i) \geqq 1 \quad \text{für} \quad i \in J, \Phi\Big(\bigcup_{i \in I \setminus J} U_i\Big) = 0\Big)\Big|$$
$$\leqq 2Q^*(\Phi(U_i) > 1 \text{ für mindestens ein } i \in I),$$

woraus wiederum wegen 3. auf

$$|P(\Phi(X) = m) - Q^*(\Phi(X) = m)| \leqq 2[P(\Phi(U_i) > 1 \text{ für mindestens ein } i \in I)$$
$$+ Q^*(\Phi(U_i) > 1 \text{ für mindestens ein } i \in I)]$$

geschlossen werden kann. Wegen 4. kann die rechte Seite beliebig klein gemacht werden, d.h., es ist

$$Q_X = P_X = Q_X^*$$

für alle X aus $\mathfrak{B}$.

Wäre

$$Q(\Phi \text{ ist einfach}) < 1,$$

so müßte ein X aus $\mathfrak{B}$ mit der Eigenschaft

$$Q(\Phi((\cdot) \cap X) \text{ ist einfach}) < 1$$

existieren. Dann ist

$$Q(\Phi^*(X) < \Phi(X)) > 0,$$

was der Gleichung

$$Q_X = Q_X^*$$

widerspricht.

Damit ist unser Beweis beendet.

Abschließend geben wir ein einfaches Beispiel dafür an, daß unser Satz nicht gilt, wenn nicht vorausgesetzt wird, daß $\Phi \in M$ fast sicher einfach bezüglich P ist. Dabei gehen wir von einem Gegenbeispiel aus, in dem P, Q Verteilungsgesetze zufälliger Vektoren mit nichtnegativen ganzen Koordinaten sind. Die n-dimensionalen zufälligen Vektoren $[\xi_1, \xi_2, \ldots, \xi_n]$ mit nichtnegativen ganzen Koordinaten ξ_i lassen sich nämlich als spezielle zufällige Elemente des meßbaren Raumes $[M, \mathfrak{M}]$ ansehen, denn im Spezialfall $A = \{1, \ldots, n\}$ erhalten wir in Gestalt von

$$\Phi \curvearrowright [\Phi(\{1\}), \ldots, \Phi(\{n\})] \qquad (\Phi \in M)$$

eine umkehrbar eindeutige Abbildung von M auf die Menge aller n-dimensionalen Vektoren mit nichtnegativen ganzen Koordinaten.

Es seien P, Q Verteilungsgesetze zufälliger Vektoren $[\xi_1, \xi_2]$ mit nichtnegativen ganzen Koordinaten. Die Verteilungsgesetze von $\sum_{i \in I} \xi_i$ mit $I \subseteq \{1, 2\}$ bezüglich P, Q sind gleich, d.h., es gilt

$$P_{\sum_{i \in I} \xi_i} = Q_{\sum_{i \in I} \xi_i}$$

für alle $I \subseteq \{1, 2\}$.

Hat nun P die Gestalt

$$P = \pi_1 \times \pi_1$$

wobei π_1 die Poissonverteilung mit dem Erwartungswert 1 ist, so kann man mindestens ein Q mit den genannten Eigenschaften und $Q \neq P$ angeben.

Ein Verteilungsgesetz P auf $\mathfrak{M}$ heißt kontinuierlich, wenn alle a aus A der Gleichung

$$P\big(\Phi(\{a\}) > 0\big) = 0$$

genügen.

Wenn wir nun in unserem Satz die Voraussetzung, daß Φ fast sicher einfach bezüglich P ist, ersetzen durch die Bedingung P auf $\mathfrak{M}$ ist kontinuierlich, so kann nicht auf $P = Q$ geschlossen werden, falls $P_X = Q_X$ für alle X aus $\mathfrak{B}$ gilt. D.h. also, wir können nicht die Bedingung

$$P(\Phi \text{ ist einfach}) = 1$$

durch

$$P\big(\Phi(\{a\}) > 0\big) = 0 \quad \text{für alle} \quad a \in A$$

ersetzen. Dazu geben wir ein Beispiel an und erhalten zugleich das bereits angekündigte Beispiel dafür, daß unser Satz ohne die Voraussetzung

$$P(\Phi \text{ ist einfach}) = 1$$

nicht gilt.

Es sei $[A, \varrho_A]$ speziell das Intervall $[-1, +1]$ mit der üblichen Metrik. Mit Hilfe eines zufälligen Vektors $\big[[\xi_1, \xi_2], \eta_1, \eta_2\big]$ mit den unabhängigen Komponenten $[\xi_1, \xi_2]$, η_1 und η_2 bilden wir die Menge M aller Punktfolgen der Gestalt

$$\Phi = \xi_1 \delta_{\eta_1} + \xi_2 \delta_{\eta_2}.$$

Die zufälligen Vektoren $[\xi_1, \xi_2]$ mit nichtnegativen ganzen Komponenten mögen die voneinander verschiedenen Verteilungen

$$P(\xi_1 = i, \xi_2 = j) \quad \text{und} \quad Q(\xi_1 = i, \xi_2 = j)$$

mit gleichen Rand- und Summenverteilungen haben. Verteilungen P und Q mit den genannten Eigenschaften existieren, wie das oben angegebene Gegenbeispiel zeigt. Weiter wird vorausgesetzt, daß die Zufallsgröße η_1 in $[-1, 0)$ gleichverteilt ist und die Zufallsgröße η_2 in $[0, +1]$. Mit Hilfe von P und Q erhalten wir dann auf der zu M gehörenden σ-Algebra $\mathfrak{M}$ zwei Verteilungsgesetze. Φ ist nicht fast sicher einfach bezüglich P, denn es ist $P(\xi_1>1$ oder $\xi_2>1)>0$. Die Verteilungsgesetze P und Q auf $\mathfrak{M}$ sind offensichtlich kontinuierlich, aber nicht gleich, weil für die Komponenten ξ_1 und ξ_2 der zufälligen Vektoren $[\xi_1, \xi_2]$ gilt $\xi_1=\Phi([1, 0))$ bzw. $\xi_2=\Phi([0, +1])$ und weil die zufälligen Vektoren $[\xi_1, \xi_2]$ die voneinander verschiedenen P- und Q-Verteilungen haben. Wir können jedoch zeigen, daß

$$P_X=Q_X$$

für alle X aus $\mathfrak{B}$ gilt.

Für $n>0$ und alle $X\in\mathfrak{B}$ gilt

$$P_X(\{n\}) = P(\Phi(X) = n) = \sum_{m_1+m_2=n} P(\Phi(X_1) = m_1, \Phi(X_2) = m_2),$$

wobei $X_1 = X\cap[-1, 0)$ und $X_2 = X\cap[0\ \ +1]$ ist.

Die einzelnen Summanden $P(\Phi(X_1)=m_1, \Phi(X_2)=m_2)$ lassen sich wegen der stochastischen Unabhängigkeit der Komponenten des Vektors $[[\xi_1, \xi_2], \eta_1, \eta_2]$ wie folgt berechnen:

1. Es sei $m_1>0$ und $m_2>0$. Dann ist

$$P(\Phi(X_1)=m_1, \Phi(X_2)=m_2)=P(\eta_1\in X_1)\cdot P(\eta_2\in X_2)\cdot P(\xi_1=m_1, \xi_2=m_2).$$

2. Es sei $m_1=0$ und $m_2>0$. Dann ist

$$P(\Phi(X_1)=0, \Phi(X_2)=m_2)=P(\eta_1\in X_1)\cdot P(\eta_2\in X_2)\cdot P(\xi_1=0, \xi_2=m_2)$$
$$+P(\eta_1\notin X_1)\cdot P(\eta_2\in X_2)\cdot P(\xi_2=m_2).$$

3. Es sei $m_1>0$ und $m_2=0$. Dann ist

$$P(\Phi(X_1)=m_1, \Phi(X_2)=0)=P(\eta_1\in X_1)\cdot P(\eta_2\in X_2)\cdot P(\xi_1=m_1, \xi_2=0)$$
$$+P(\eta_2\notin X_2)\cdot P(\eta_1\in X_1)\cdot P(\xi_1=m_1).$$

Somit erhalten wir für $n>0$ und alle $X\in\mathfrak{B}$:

$$P_X(\{n\})=P(\Phi(X)=n)=P(\eta_1\in X_1)P(\eta_2\in X_2) \sum_{m_1+m_2=n} P(\xi_1=m_1, \xi=m_2)+$$
$$+P(\eta_1\in X_1)P(\eta_2\notin X_2)P(\xi_1 = n)+P(\eta_1\notin X_1)P(\eta_2\in X_2)P(\xi_2=n).$$

Für $n=0$ und alle $X\in\mathfrak{B}$ gilt:

$$P_X(\{0\})=P(\Phi(X)=0)=P(\Phi(X_1)=0, \Phi(X_2)=0)$$
$$=P(\eta_1\notin X_1)\cdot P(\eta_2\notin X_2)+P(\eta_1\in X_1)\cdot P(\eta_2\notin X_2)\cdot P(\xi_1=0)$$
$$+P(\eta_1\notin X_1)\cdot P(\eta_2\in X_2)\cdot P(\xi_2=0)+P(\eta_1\in X_1)\cdot P(\eta_2\in X_2)\cdot P(\xi_1=0, \xi_2=0).$$

Entsprechende Darstellungen gelten auch für $Q_X(\{n\})$ mit $n>0$ und $n=0$. Nach Voraussetzung haben die Verteilungen P, Q der zufälligen Vektoren $[\xi_1, \xi_2]$ die gleichen Rand- und Summenverteilungen, d.h., für die Werte $i, j, k = 0, 1, 2, \ldots$ gilt

$$P(\xi_2=j)=\sum_i P(\xi_1=i, \xi_2=j)=\sum_i Q(\xi_1=i, \xi_2=j)=Q(\xi_2=j),$$

$$P(\xi_1=i)=\sum_j P(\xi_1=i, \xi_2=j)=\sum_j Q(\xi_1=i, \xi_2=j)=Q(\xi_1=i),$$

$$\sum_{i+j=k} P(\xi_1=i, \xi_2=j)=\sum_{i+j=k} Q(\xi_1=1, \xi_2=j).$$

Dann erkennen wir aber unmittelbar, daß

$$P_X = Q_X$$

für alle $X \in \mathfrak{B}$ gilt.

LITERATUR

[1] RÉNYI, A.: Remarks on the Poisson process, *Studia Scientiarum Mathematicarum Hungarica* **2** (1967), 119—123.

[2] HALMOS, P. R.: *Measure Theory*, New York 1950. (russische Übersetzung).

[3] PRÉKOPA, A.: On stochastic set functions, Acta Math. Acad Sci. Hung. 7 (1956), 215—263.

[4] PRÉKOPA, A.: On secondary processes generated by a random point distribution of Poisson type, Annales Univ. Sci. Budapest de R. Eötvös Nom. **1** (1958), 153—170.

[5] PRÉKOPA, A.: On Poisson and composed Poisson stochastic set functions, Studia Math. **16** (1957), 142—155,

VEB *Kombinat Zentronik*
50 *Erfurt*—DDR
Neuwerkstraße 45

(*Eingegangen:* 21. *August* 1969.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 91—97.

HERMITE-FEJÉR INTERPOLATION BASED ON THE ROOTS OF LAGUERRE POLYNOMIALS

by

P. O. H. VÉRTESI

1. In this paper we intend to continue our investigation begun in [2]. For a continuous function $f(x)$ on the interval $[0, \infty)$ we define the uniquely determined Hermite—Fejér interpolating polynomials of degree $\leqq 2n-1$ as follows.

$$(1.1)\qquad \begin{cases} H_n^{(\alpha)}(f; x) = \sum_{k=1}^{n} f(x_{kn}^{(\alpha)}) h_{kn}^{(\alpha)}(x) \; ^{*} \text{ or} \\ \overline{H}_n^{(\alpha)}(f; x) = H_n^{(\alpha)}(f; x) + \sum_{k=1}^{n} \beta_{kn} \mathfrak{h}_{kn}^{(\alpha)}(x), \end{cases}$$

where $L_n^{(\alpha)}(x)$ is the nth Laguerre polynomial defined by the well-known relation

$$(1.2)\qquad e^{-x} x^{\alpha} L_n^{(\alpha)}(x) = \frac{1}{n!} \frac{d^n}{dx^n} (e^{-x} x^{\alpha+n}),$$

$x_{kn}^{(\alpha)}$ are the roots of $L_n^{(\alpha)}(x)$ having the property

$$(1.3)\qquad 0 < x_{1n}^{(\alpha)} < x_{2n}^{(\alpha)} < \ldots < x_{nn}^{(\alpha)} < \infty,$$

$$(1.4)\qquad l_{kn}^{(\alpha)}(x) = \frac{L_n^{(\alpha)}(x)}{L_n'^{(\alpha)}(x_{kn}^{(\alpha)})(x - x_{kn}^{(\alpha)})},$$

$$(1.5)\qquad h_{kn}^{(\alpha)}(x) = \left[1 - \frac{L_n''^{(\alpha)}(x_{kn}^{(\alpha)})}{L_n'^{(\alpha)}(x_{kn}^{(\alpha)})}(x - x_{kn}^{(\alpha)})\right] [l_{kn}^{(\alpha)}(x)]^2 = v_{kn}^{(\alpha)}(x) [l_{kn}^{(\alpha)}(x)]^2,$$

$$(1.6)\qquad \mathfrak{h}_{kn}^{(\alpha)}(x) = (x - x_{kn}^{(\alpha)}) [l_{kn}^{(\alpha)}(x)]^2,$$

β_{kn} are prescribed (s. [1], 14. 1). It is well known (s. [1], 14. 1) that

$$(1.7)\qquad \sum_{k=1}^{n} h_{kn}^{(\alpha)}(x) \equiv 1 \qquad (n = 1, 2, \ldots, x \in [0, \infty)),$$

$$(1.8)\qquad H_n^{(\alpha)}(f; x_{kn}^{(\alpha)}) = \overline{H}_n^{(\alpha)}(f; x_{kn}^{(\alpha)}) = f(x_{kn}^{(\alpha)}) \qquad (k = 1, 2, \ldots, n;\; n = 1, 2, \ldots),$$

$$(1.9)\qquad H_n'^{(\alpha)}(f; x_{kn}^{(\alpha)}) = 0, \quad \overline{H}_n'^{(\alpha)}(f; x_{kn}^{(\alpha)}) = \beta_{kn} \qquad (k = 1, 2, \ldots, n;\; n = 1, 2, \ldots).$$

We have the following theorem (s. [1], Theorem 14. 7.).

* Throughout this paper let $\alpha > -1$.

THEOREM 1. 1. *Let $f(x)$ be continuous on $[0, \infty)$, $f(x)=O(x^m)$ for $x\to\infty$ ($m>0$ is fixed). The sequence $H_n^{(\alpha)}(f; x)$ uniformly converges to $f(x)$ in $[\delta, \Delta]$ ($0<\delta<\Delta<\infty$ are fixed). If $\alpha<0$, then the uniform convergence is true on $[0, \delta]$ as well. Further, if $\alpha\geqq 0$, then there exists a continuous function such that*

$$\overline{\lim_{n\to\infty}}\, |H_n^{(\alpha)}(f; 0)-f(0)| > 0. \tag{1.10}$$

2. The purpose of this paper is to give some further estimations for the difference $H_n^{(\alpha)}(f; x)-f(x)$. Let ε an arbitrarily small, but fixed positive number. Denote by $\omega_\Delta(f; t)$ the modulus of continuity of $f(x)$ on the interval $[0, \Delta+\varepsilon]$ further let $\omega_\Delta(t)$ be a modulus of continuity on $[0, \Delta+\varepsilon]$. Then we have

THEOREM 2. 1. *Let $f(x)$ be a continuous function on $[0, \infty)$, $f(x)=O(x^m)$ when $x\to\infty$ ($m>0$ is fixed). Suppose $\omega_\Delta(f; t)=O[\omega_\Delta(t)]$. For the interval $[\delta, \Delta]$ we have the relations*

$$|f(x)-H_n^{(\alpha)}(f; x)| = \begin{cases} O(1)\displaystyle\sum_{i=1}^{[\sqrt{n}]}\omega_\Delta\left(\frac{i}{\sqrt{n}}\right)\frac{1}{i^2} & \text{for } -\dfrac{1}{2}\leqq\alpha,\ 0<\delta\leqq x\leqq\Delta, \\ O(1)\left[n^{-\alpha-\frac{3}{2}}\displaystyle\sum_{i=1}^{[\sqrt{n}]}\omega_\Delta\left(\frac{i}{\sqrt{n}}\right)+\sum_{i=1}^{[\sqrt{n}]}\omega_\Delta\left(\frac{i}{\sqrt{n}}\right)\frac{1}{i^2}\right] & \text{for} \\ \qquad -1<\alpha<-\dfrac{1}{2},\quad 0<\delta\leqq x\leqq\Delta. \end{cases} \tag{2.1}$$

Before proving this we investigate some special cases.

2. 1. The case $-\frac{1}{2}\leqq\alpha$.

We can easily prove the convergence. Indeed, we have

$$|f(x)-H_n^{(\alpha)}(f; x)| = O(1)\left[\omega_\Delta\left(\frac{\log n}{\sqrt{n}}\right)\sum_{i=1}^{[\sqrt{n}]}\left(\frac{1}{i\log n}+\frac{1}{i^2}\right)\right] =$$

$$= O(1)\omega_\Delta\left(\frac{\log n}{\sqrt{n}}\right)\left(\frac{\log\sqrt{n}}{\log n}+\sum_{i=1}^{[\sqrt{n}]}\frac{1}{i^2}\right) = O(1)\omega_\Delta\left(\frac{\log n}{\sqrt{n}}\right) \qquad (\delta\leqq x\leqq\Delta).$$

If $f\in \mathrm{Lip}\,\varrho\,(0<\varrho\leqq 1)$, then we obtain

$$|f(x)-H_n^{(\alpha)}(f; x)| = O(1)\left(\frac{1}{n^{\varrho/2}}\sum_{i=1}^{[\sqrt{n}]} i^{\varrho-2}+\frac{1}{n^{\varrho/2}}\right). \tag{2.2}$$

I.e.,

$$|f(x)-H_n^{(\alpha)}(f; x)| = \begin{cases} O\left(n^{-\frac{\varrho}{2}}\right) & \text{for } 0<\varrho<1,\quad \delta\leqq x\leqq\Delta,\quad -\dfrac{1}{2}\leqq\alpha, \\ O\left(\dfrac{\log n}{\sqrt{n}}\right) & \text{for } \varrho=1,\quad \delta\leqq x\leqq\Delta,\quad -\dfrac{1}{2}\leqq\alpha. \end{cases}$$

2. 2. The case $-1<\alpha<-\frac{1}{2}$.

By (2. 2) we can see that the difference between 2. 1 and 2. 2 is the part $n^{-\alpha-\frac{3}{2}}\sum_{i=1}^{[\sqrt{n}]}\omega_\Delta\left(\frac{1}{\sqrt{n}}\right)$. So we have to estimate only this expression. Prove the convergence.

$$|f(x)-H_n^{(\alpha)}(f;x)| = O(1)\left[\omega_\Delta\left(\frac{\log n}{\sqrt{n}}\right)+n^{-\alpha-\frac{3}{2}}\omega_\Delta(n^{-1-\alpha})\sum_{i=1}^{[\sqrt{n}]}\left(n^{\alpha+\frac{1}{2}}i+1\right) =\right.$$

$$= O(1)\left[\omega_\Delta\left(\frac{\log n}{\sqrt{n}}\right)+\omega_\Delta\left(\frac{1}{n^{1+\alpha}}\right)\left(1+\frac{1}{n^{1+\alpha}}\right)\right] = O(1)\omega_\Delta\left(\frac{1}{n^{1+\alpha}}\right) \qquad (\delta \leqq x \leqq \Delta).$$

If $f\in \mathrm{Lip}\,\varrho$ $(0<\varrho\leqq 1)$, then we obtain

$$n^{-\alpha-\frac{3}{2}}\sum_{i=1}^{[\sqrt{n}]}\frac{i^\varrho}{n^{\varrho/2}} = O(1)n^{-\alpha-\frac{3}{2}-\frac{\varrho}{2}+\frac{\varrho}{2}+\frac{1}{2}} = O(n^{-\alpha-1}).$$

I.e.

$$|f(x)-H_n^{(\alpha)}(f;x)| = \begin{cases} O\left(n^{-\frac{\varrho}{2}}\right) & \text{for} \quad \alpha+1 \geqq \dfrac{\varrho}{2}, \quad \delta \leqq x \leqq \Delta, \quad -1<\alpha<-\dfrac{1}{2}, \\ O(n^{-\alpha-1}) & \text{for} \quad \alpha+1 \leqq \dfrac{\varrho}{2}, \quad \delta \leqq x \leqq \Delta, \quad -1<\alpha<-\dfrac{1}{2}. \end{cases}$$

2. 3. *Proof of Theorem 2. 1.*

We shall use some formulae of [1]. We have

$$(2.\ 3) \qquad v_{kn}^{(\alpha)}(x) = \frac{x_k(x_k-\alpha)+x(\alpha+1-x_k)}{x_k}{}^{*} \qquad (k=1,2,\ldots,n;\ n=1,2,\ldots),$$

$$(2.\ 4) \qquad 2\sqrt{x_{kn}^{(\alpha)}} = \frac{1}{\sqrt{n}}[k\pi+O(1)] \qquad (0<x_{kn}^{(\alpha)}\leqq\Omega,\ n=1,2,\ldots),$$

$$(2.\ 5) \qquad |L_n'^{(\alpha)}(x_k)| \sim k^{-\alpha-\frac{3}{2}}n^{\alpha+1} \qquad (0<x_k\leqq\Omega,\ n=1,2,\ldots)$$

(The notation $z_n\sim w_n$ means that $c_1\leqq|z_n|/|w_n|<c_2$ $(n\geqq N)$ where $0<c_1\leqq c_2<\infty$, $w_n\neq 0$. In (2. 4), (2. 5) and (2. 6) the O, c_1 and c_2 depend only on α, β and Ω.)

$$(2.\ 6) \qquad |L_n^{(\alpha)}(x)| = \begin{cases} x^{-\frac{\alpha}{2}-\frac{1}{4}}O\left(n^{\frac{\alpha}{2}-\frac{1}{4}}\right) & \text{for} \quad cn^{-1}\leqq x\leqq\Omega, \\ O(n^\alpha) & \text{for} \quad 0\leqq x\leqq cn^{-1} \end{cases}$$

where c and Ω are arbitrary fixed positive numbers (see [1], (14. 5. 5), (8. 9. 10), (8. 9. 11) and (7. 6. 8.)).

Now we estimate the difference $|f(x)-f(x_k)|$. Denote by $x_{jn}^{(\alpha)}$ the nearest roots to x (plainly $j=j(n)$).

---

* Sometimes we omit the unnecessary indices.

Then by (2. 4) we have

$$(2.7)\quad |x_k - x| = O(1)\left|\sqrt{x_k}-\sqrt{x_j}\right|\left|\sqrt{x_k}+\sqrt{x_j}\right| = O(1)\frac{ij+i^2}{n} \quad \text{if} \quad k \neq j, \qquad \begin{matrix} k = j+i \text{ for } k > j, \\ k = j-i \text{ for } k < j \quad (0 \leqq x,\ x_k \leqq \Omega). \end{matrix}$$

So

$$(2.8)\quad |f(x)-f(x_k)| = O(1)\left[\omega_\Delta\left(\frac{ij}{n}\right)+\omega_\Delta\left(\frac{i^2}{n}\right)\right] \quad (k\neq j,\ k=j\pm i,\ 0\leqq x,\ x_k\leqq \Delta+\varepsilon).$$

We prove the estimations

$$(2.9)\quad \begin{cases} j(n)=1 & \text{for } x=0, \\ j(n)\sim\sqrt{n}, \quad k\sim\sqrt{n} & \text{for } 0<\delta\leqq x,\ x_k\leqq\Omega. \end{cases}$$

The first formula is obvious. On the other hand, by (2. 4) we have that $x_{kn}\leqq x\leqq\Omega$ if and only if $k=O(\sqrt{n})$. Thus we can easily prove the second statement as well.

Further by (2. 4) $x_{k+1}-x_k \sim \dfrac{2k+1}{n}$, so

$$x_{k+1}-x_k \sim \begin{cases} \dfrac{1}{n} & \text{for } k = O(1), \\ \dfrac{1}{\sqrt{n}} & \text{for } \delta \leqq x_k, \quad x_{k+1} \leqq \Omega. \end{cases}$$

I.e.

$$(2.10)\quad |f(x)-f(x_j)| = \begin{cases} O(1)\,\omega_\Delta\left(\dfrac{1}{n}\right) & \text{for } x = 0, \\ O(1)\,\omega_\Delta\left(\dfrac{1}{\sqrt{n}}\right) & \text{for } \delta \leqq x,\ x_j \leqq \Delta+\varepsilon. \end{cases}$$

Consider the expression $v_k(x)$. By (2. 3) we can prove that there exists a $\xi(\alpha)$ such that

$$(2.11)\quad 0<m(\alpha)\leqq v_{kn}^{(\alpha)}(x)\leqq M(\alpha) \quad \text{if} \quad |x-x_{kn}| \leqq \xi(\alpha) \leqq \varepsilon$$
$$(k=1, 2, \dots, n; \quad n=1, 2, \dots)\,*$$

This means that

$$(2.12)\quad 0 \leqq m(\alpha)\,l_{jn}^{2(\alpha)}(x) \leqq h_{kn}^{(\alpha)}(x) \leqq M(\alpha)\,l_{jn}^{2(\alpha)}(x) \qquad (|x-x_{kn}| \leqq \xi(\alpha)).$$

By (2. 3) and (2. 4) we get

$$(2.13)\quad |v_{kn}^{(\alpha)}(x)| = \begin{cases} O(x_k^{-1}) = O\left(\dfrac{n}{k^2}\right) & \text{for} \quad 0 < x_k \leqq x \leqq \Omega, \\ O(x_k) = O\left(\dfrac{k^2}{n}\right) & \text{for} \quad \delta \leqq x < x_k \leqq \Omega. \end{cases}$$

* Clearly we can suppose that $0<\xi(\alpha)\leqq\varepsilon$.

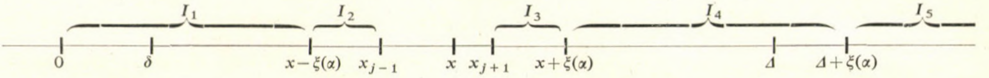
We shall use the following relations

$$(2.14)\qquad \sum_{k=1}^{n} x_k^{m+1}[L_n'^{(\alpha)}(x_k)]^{-2} = \frac{\Gamma(n+1)\Gamma(m+\alpha+3)}{\Gamma(n+\alpha+1)} = O(n^{-\alpha})$$
$$(m = 1, 2, \dots, 2n-3),$$

$$(2.15)\qquad \frac{d}{dx} L_n^{(\alpha)}(x) = -L_{n-1}^{(\alpha+1)}(x).$$

(s. [1], (14. 7. 5), (5. 1. 14)).

We divide the interval $[0, \infty)$ as follows:



By (1. 7); (2. 10); $A_i \equiv \omega_\Delta\left(\frac{ij}{n}\right)+\omega_\Delta\left(\frac{i^2}{n}\right)$; (1. 4), (2. 13), (2. 6), (2. 5); (2. 11), (2. 6), (2. 5), (2. 7); (2. 13), (2. 6), (2. 5), $f(x)=O(x^m)$; (2. 6) and (2. 15) we get

$$|f(x)-H_n(f;x)| = \left|\sum_{k=1}^{n}[f(x)-f(x_k)]h_k(x)\right| =$$
$$= \left|\sum_{k\neq j}[f(x)-f(x_k)]h_k(x)+[f(x)-f(x_j)]h_j(x)\right| =$$
$$= O(1)\left[\sum_{k\neq j}|f(x)-f(x_k)|\,|h_k(x)|+\omega_\Delta\left(\frac{1}{\sqrt{n}}\right)h_j(x)\right] =$$
$$= O(1)\left[\left(\sum_{x_k\in I_1}+\sum_{x_k\in I_2}+\sum_{x_k\in I_3}+\sum_{x_k\in I_4}\right)A_i|h_k(x)|+\sum_{x_k\in I_5}|f(x)-f(x_k)|\,|h_k(x)|+\right.$$
$$\left.+\omega_\Delta\left(\frac{1}{\sqrt{n}}\right)h_j(x)\right] = O(1)\left\{\sum_{x_k\in I_1}A_i nk^{-2}n^{\alpha-\frac{1}{2}}k^{2\alpha+3}n^{-2\alpha-2}\xi^{-2}+\right.$$
$$+\sum_{x_k\in I_2\cup I_3}A_i M(\alpha)n^{\alpha-\frac{1}{2}}k^{2\alpha+3}n^{-2\alpha-2}n^2(ij+i^2)^{-2}+$$
$$+\sum_{x_k\in I_4}A_i k^2 n^{-1}n^{\alpha-\frac{1}{2}}k^{2\alpha+3}n^{-2\alpha-2}\xi^{-2}+\sum_{x_k\in I_5}x_k^m x_k n^{\alpha-\frac{1}{2}}[L_n'^{(\alpha)}(x_k)]^{-2}\xi^{-2}+$$
$$\left.+\omega_\Delta\left(\frac{1}{\sqrt{n}}\right)M(\alpha)[n^{\alpha+\frac{1}{2}}j^{2\alpha+3}n^{-2\alpha-2}+1]\ ^{*}\right\}.$$

By (2. 9) and (2. 14) we obtain

$$= O(1)\left[n^{-\alpha-\frac{3}{2}}\sum_{x_k\in I_1}A_i k^{2\alpha+1}+\sum_{x_k\in I_2\cup I_3}\frac{A_i}{i^2}+\frac{1}{n}\sum_{x_k\in I_4}A_i+\frac{1}{\sqrt{n}}+\omega_\Delta\left(\frac{1}{\sqrt{n}}\right)\right].$$

* If $x=x_j$, then $h_j(x_j)=1$. On the other hand,

$$\frac{L_n^{(\alpha)}(x)-L_n^{(\alpha)}(x_j)}{x-x_j} = L_n'^{(\alpha)}(x^*) = -L_{n-1}^{(\alpha+1)}(x^*).$$

If $2\alpha+1\geqq 0$ $(-\frac{1}{2}\leqq\alpha)$, then by (2.9) we get

$$|f(x)-H_n(f;x)| = O(1)\left\{\sum_{i=1}^{[\sqrt{n}]}\left[\omega_\Delta\left(\frac{ij}{n}\right)+\omega_\Delta\left(\frac{i^2}{n}\right)\right]\left(\frac{1}{n}+\frac{1}{i^2}\right)+\frac{1}{\sqrt{n}}+\omega_\Delta\left(\frac{1}{\sqrt{n}}\right)\right\} =$$

$$= O(1)\left\{\frac{1}{n}\sum_{i=1}^{[\sqrt{n}]}\omega_\Delta(1)\left[\left(\frac{i\sqrt{n}}{n}+\frac{i^2}{n}+2\right)+\right.\right.$$

$$(2.15)\qquad +\sum_{i=1}^{[\sqrt{n}]}\left[\omega_\Delta\left(\frac{i}{\sqrt{n}}\right)+\omega_\Delta\left(\frac{i^2}{n}\right)\right]\frac{1}{i^2}+\frac{1}{\sqrt{n}}+\omega_\Delta\left(\frac{1}{\sqrt{n}}\right)\Bigg] =$$

$$= O(1)\left[\frac{1}{\sqrt{n}}+\omega_\Delta\left(\frac{1}{\sqrt{n}}\right)+\sum_{i=1}^{[\sqrt{n}]}\omega_\Delta\left(\frac{i}{\sqrt{n}}\right)\frac{1}{i^2}\right] = O(1)\sum_{i=1}^{[\sqrt{n}]}\omega_\Delta\left(\frac{i}{\sqrt{n}}\right)\frac{1}{i^2}.$$

For $-1<\alpha<-\frac{1}{2}$ we have by a similar computation

$$(2.16)\qquad |f(x)-H_n(f;x)| = O(1)\left[n^{-\alpha-\frac{3}{2}}\sum_{i=1}^{[\sqrt{n}]}\omega_\Delta\left(\frac{i}{\sqrt{n}}\right)+\sum_{i=1}^{[\sqrt{n}]}\omega_\Delta\left(\frac{i}{\sqrt{n}}\right)\frac{1}{i^2}\right]^{*},$$

as we stated. Q.e.d.

3. In this part we investigate the convergence at the point 0. It can be shown arguing as before the following

THEOREM 3. 1. *Let $f(x)$ be a continuous function on* $[0, \infty]$, $f(x)=O(x^m)$ *when* $x\to\infty$ $(m>0$, *fixed*). *For* $x=0$ *we have the relation*

$$(3.1)\qquad |f(0)-H_n^{(\alpha)}(f;0)| = O(n^\alpha).$$

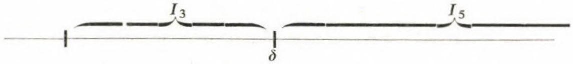
Before proving this let us consider the special case.

3. 1. By $\alpha<0$ we have the convergence for $x=0$.

3. 2. *Proof of Theorem 3. 1.* By (2. 3) we get

$$(3.2)\qquad v_k(0) = x_k-\alpha.$$

We divide the interval $[0, \infty)$ as follows.



Let $x=0$. Then by (1. 7); (2. 8), (3. 2), (2. 6), (2. 5), (2. 4); (2. 14)

$$|f(0)-H_n(f;0)| = \left|\sum_{k=1}^{n}[f(0)-f(x_k)]h_k(x)\right| =$$

$$(3.3)\qquad = O(1)\left\{\sum_{x_k\in I_3}\omega_\delta\left(\frac{i^2}{n}\right)\left|\frac{i^2}{n}-\alpha\right| n^{2\alpha}i^{2\alpha+3}n^{-2\alpha-2}n^2 i^{-4}+\right.$$

$$\left.+\sum_{x_k\in I_5}x_k^{m+1}n^{2\alpha}[L'^{(\alpha)}(x_k)\delta]^{-2}\right\} = O(1)\left[\sum_{i=1}^{[\sqrt{n}]}\omega_\delta\left(\frac{i^2}{n}\right)\left(\frac{i^{2\alpha+1}}{n}+i^{2\alpha-1}\right)+n^\alpha\right] =$$

$$= O(1)\left[\frac{\omega_\delta(1)}{n}\sum_{i=1}^{[\sqrt{n}]}\left(\frac{i^2}{n}+2\right)i^{2\alpha+1}+n^\alpha\right] = O(n^\alpha).$$

Q.e.d.

* If $x_k\in I_1$, then $k^{2\alpha+1}=O(1)$.

4. Now we wish to give a lower estimation for the above mentioned Hermite—Fejér procedure. We shall use Theorem 3. 2 from [3].

Let $\omega(t)$ be a modulus of continuity on $[0, \Delta]$, $\omega(t)\cdot t^{-1}\to\infty$ for $t\to+0$, $C_\Delta(\omega)=\{g(x); g(x)$ is continuous on $0\leqq x\leqq\Delta$, $g(x)\equiv 0$ for $x\geqq\Delta$, $\omega(g;t)=$ $=O[\omega(t)]\}$. With the notations of [3]

$$L_n(g;x) = H_n^{(\alpha)}(g;x) = \sum_{k=1}^{n} g(x_{kn})h_{kn}^{(\alpha)}(x).$$

Let $I_n=\{x_{1n}, x_{kn}\}$ where $x_{kn}\leqq\Delta$. Hence

$$\lambda_n(I_n;0)\equiv\sum_{\substack{x_{kn}<\Delta\\ k\neq 1}}|h_{kn}^{(\alpha)}(0)|\sim\sum_{\substack{x_{kn}<\Delta\\ k\neq 1}}\frac{|x_k-\alpha|}{x_k^2}n^{2\alpha}k^{2\alpha+3}n^{-2\alpha-2}\sim$$

$$\sim\frac{1}{n}\sum_{k=1}^{\sqrt{n}}(k^{2\alpha+1}+n\cdot k^{2\alpha-1})\sim n^{\alpha}.$$

We know that

$$d_n(I_n)=\min(x_{k+1,n}-x_{kn})\sim\frac{1}{n}$$

Hence by Theorem 3. 2 from [3] we have

THEOREM 4. 1. *If* $\overline{\lim_{n\to\infty}}\,\lambda_n(I_n;0)>1$ or $\underline{\lim}_{n\to\infty}\,\lambda_n(I_n;0)<1$, *then there exists an* $f(x)\in C_\Delta(\omega)$ *such that*

$$|H_n^{(\alpha)}(f;0)-f(0)|>n^{\alpha}\omega\left(\frac{1}{n}\right)\qquad(n=n_1,n_2,\ldots).\tag{4. 1}$$

(Here $0<n_1<n_2<\cdots$ are suitable integers.) (We can see (by remodelling of the proof) that the theorem mentioned above can be applied in our case, as well. We notice that $f(x)=Q\sum_{i=1}^{\infty}\omega(d_{n_i}(I_{n_i}))g_{n_i}(x)\in C_\Delta(\omega)$ because of $f(x)\equiv 0$ for $x\geqq\Delta$.)

5. *Notes*

5. 1. We can obtain similar results for $\overline{H}_n^{(\alpha)}(f;x)$ as well.

5. 2. If α is large enough then $H_n^{(\alpha)}(f;0)$ diverges very rapidly for a suitable $f(x)$ (s. (4. 1)).

REFERENCES

[1] SZEGŐ, G.: *Orthogonal polynomials*, Amer. Math. Soc. New York, 1959.

[2] VÉRTESI, P. O. H. On the convergence of Hermite—Fejér interpolation, *Acta Math. Acad. Sci. Hung.* **22** (1971).

[3] KIS, O. and VÉRTESI, P. O. H.: On certain linear operators, *Acta Math. Acad. Sci. Hung.* **22** (1971).

Mathematical Institute of the Hungarian Academy of Sciences

(Received September 24, 1969)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 99—111.

ON A PROOF OF JACKSON'S THEOREM THROUGH AN INTERPOLATION PROCESS

by

K. K. MATHUR

1. Jackson's theorem has been proved by various mathematicians using difterent methods. In a recent paper, Professor FREUD [1] has given a proof of this well-known theorem in the closed interval $[-\frac{1}{2}, \frac{1}{2}]$ directly by means of an interpolatory polynomial constructed on the roots of $T_n(x)$-Čebyšev polynomial of the first kind. SAXENA [5] improved FREUD's result by proving the theorem in the closed interval $[-1, 1]$ through an interpolation process having as abscissas the zeros of $U_n(x)$-Čebyšev polynomial of the second kind. VÉRTESI [7] proved the same result by choosing the interpolatory polynomial on the roots of Čebyšev polynomial of the first kind. Later, in a joint paper, FREUD and VÉRTESI [2] remarkably improved he result contained in [7], and gave a very elegant proof of TIMAN's well-known approximation theorem. Further KIS and VÉRTESI [3] proved the same theorem by choosing another interpolation process, constructed on the nodes $\cos\frac{2k\pi}{2n+1}$, $k=1, 2, \ldots, n$.

M. SALLAY [4] has proved Jackson's theorem in $(-1, +1)$ with the help of an interpolation process on the zeros of an orthogonal polynomial, whose weight function is positive on the segment $[-1, +1]$ and satisfies a Lipschitz condition of order 1. In the present paper, we prove Jackson's theorem in the interval $[-1, +1]$ through an interpolation process, built on the abscissas

$$x_{kn} = \cos\frac{2k-1}{2n+1}\pi, \qquad k = 1, 2, \ldots, n+1. \tag{1. 1}$$

The points x_{kn} $(k=1, 2, \ldots, n)$ stand for the zeros of Jacobi polynomial $P_n^{(-\frac{1}{2}, \frac{1}{2})}(x)$, where

$$P_n^{(-\frac{1}{2}, \frac{1}{2})}(x) = \frac{\Gamma(n+\frac{1}{2})}{\sqrt{\pi}\,\Gamma(n+1)} \frac{\cos(n+\frac{1}{2})\theta}{\cos\frac{\theta}{2}}, \qquad x = \cos\theta. \tag{1. 2}$$

2. Setting

$$\mu_{kn}(x) = \left(\frac{1+x}{1+x_{kn}}\right)^2 [v_{kn}(x) l_{kn}^4(x) + 2(x-x_{kn}) l_{kn}^3(x) X_{n-1}(x_{kn}, x)], \tag{2. 1}$$

where

$$v_{kn}(x) = 1 + \frac{1-2x_{kn}}{(1-x_{kn}^2)}(x-x_{kn}). \tag{2. 2}$$

$$X_{n-1}(x_{kn}, x) = \frac{2\pi}{2n+1} \sum_{j=1}^{n-1} \left\{\frac{\Gamma(j+1)}{\Gamma(j+\frac{1}{2})}\right\}^2 P_j'^{(-\frac{1}{2}, \frac{1}{2})}(x_{kn}) P_j^{(-\frac{1}{2}, \frac{1}{2})}(x) \tag{2. 3}$$

and

$$(2.4)\qquad l_{kn}(x) = \frac{P_n^{(-\frac{1}{2},\frac{1}{2})}(x)}{P_n'^{(-\frac{1}{2},\frac{1}{2})}(x_{kn})(x-x_{kn})} = \frac{(-1)^{k+1}\sin^2\theta_k\cos(n+\frac{1}{2})\theta}{(2n+1)\sin\frac{\theta_k}{2}\cos\frac{\theta}{2}(\cos\theta-\cos\theta_k)},$$

$$x = \cos\theta$$

is the fundamental polynomial of Lagrange interpolation, we have for an arbitrary function $f(x)$ defined in $-1 \leqq x \leqq 1$, the interpolation process:

$$(2.5)\qquad S_n(f; x) = -xf(-1) + \sum_{k=1}^{n} \big(f(x_{kn}) + xf(-1)\big)\mu_{kn}(x) \text{ [1]}$$

of degree $\leqq 4n-1$, such that

$$S_n(f; x_{kn}) = f(x_{kn}) \text{ [2]}, \qquad k = 1, 2, \dots, n+1.$$

We shall prove the following:

THEOREM. *Let $f(x)$ be a continuous function in the closed interval $[-1, +1]$, then for the sequence of interpolatory polynomials $\{S_n(f; x)\}$ given by (2. 5), we have*

$$|S_n(f; x) - f(x)| \leqq 738\omega\left(f, \frac{1}{2n+1}\right)$$

in $-1 \leqq x \leqq 1$, where $\omega(f, \delta)$ stands for the modulus of continuity of $f(x)$.

To prove this theorem we require the following lemmas.

3. LEMMA 3. 1. *Setting*

$$(3.1)\qquad Y_{n-1}(t, u) = \frac{2}{(2n+1)}\left[1 + \pi\sum_{j=1}^{n-1}\left\{\frac{\Gamma(j+1)}{\Gamma(j+\frac{1}{2})}\right\}^2 P_j^{(-\frac{1}{2},\frac{1}{2})}(t)P_j^{(-\frac{1}{2},\frac{1}{2})}(u)\right],$$

we have

$$(3.2)\qquad \sum_{k=1}^{n}\mu_k(x) = [(1+x)Y_{n-1}(x, x)]^2.$$

PROOF. By formula (4. 5. 7) on page 70 in [6], putting $\alpha = -\frac{1}{2}$, $\beta = \frac{1}{2}$ and $x = x_k$, we get

$$(3.3)\qquad \frac{(1-x_k^2)P_n'^{(-\frac{1}{2},\frac{1}{2})}(x_k)}{(n+1)} = -P_{n+1}^{(-\frac{1}{2},\frac{1}{2})}(x_k) = \frac{(-1)^{k+1}\Gamma(n+\frac{3}{2})}{\sqrt{\pi}\,\Gamma(n+2)}\sqrt{2(1-x_k)}$$

using (1. 1) and the fact that $\theta_k = \dfrac{2k-1}{2n+1}\pi$.

---

[1] From (2. 1) obviously $\mu_{kn}(x_{jn}) = \begin{cases} 0 & \text{for } j \neq k \\ 1 & j = k. \end{cases}$

[2] For simplicity we shall write x_k for x_{kn} and l_k for l_{kn} etc.

The well-known Christroffel-Darboux formula for $P_n^{(-\frac{1}{2},\frac{1}{2})}(x)$ with the help of (2. 4) and (3. 3) gives

$$(3.4)\qquad l_k(x) = \frac{2(1+x_k)}{(2n+1)}\left[1+\pi\sum_{j=1}^{n-1}\left\{\frac{\Gamma(j+1)}{\Gamma(j+\frac{1}{2})}\right\}^2 P_j^{(-\frac{1}{2},\frac{1}{2})}(x_k)P_j^{(-\frac{1}{2},\frac{1}{2})}(x)\right]$$

$$= (1+x_k)Y_{n-1}(x_k, x).$$

Now, let

$$\varphi_{2n}(x) \stackrel{\text{def}}{=} [(1+x)Y_{n-1}(x,\xi)]^2,$$

be a polynomial of degree $\leqq 2n$, which satisfies the properties

$$\varphi_{2n}(x_k) = l_k^2(\xi), \quad \varphi_{2n}(-1) = 0$$

and

$$\varphi'_{2n}(x_k) = \frac{2l_k^2(\xi)}{(1+x_k)} + 2(1+x_k)l_k(\xi)X_{n-1}(x_k,\xi), \quad \varphi'_{2n}(-1) = 0.$$

Then, constructing the well-known Hermite—Fejér interpolation process of degree $\leqq 2n+1$ on the abscissas (1. 1) for the polynomial $\varphi_{2n}(x)$, defined above, we have

$$\varphi_{2n}(x) = [(1+x)Y_{n-1}(x,\xi)]^2 = \sum_{k=1}^{n}\left(\frac{1+x}{1+x_k}\right)^2\left[\left(1-\frac{x-x_k}{1-x_k^2}\right)l_k^2(x)l_k^2(\xi)+\right.$$

$$\left.+\left\{\frac{2l_k^2(\xi)}{(1+x_k)}+2(1+x_k)l_k(\xi)X_{n-1}(x_k,\xi)\right\}(x-x_k)l_k^2(x)\right].$$

Putting $x=\xi$, we find that

$$[(1+x)Y_{n-1}(x,x)]^2 = \sum_{k=1}^{n}\left(\frac{1+x}{1+x_k}\right)^2\left[\left(1+\frac{1-2x_k}{1-x_k^2}(x-x_k)\right)l_k^4(x)+\right.$$

$$\left.+2(1+x_k)(x-x_k)l_k^3(x)X_{n-1}(x_k,x)\right] = \sum_{k=1}^{n}\mu_k(x)$$

by (2. 1), which proves the lemma.

LEMMA 3. 2. *For* $-1\leqq x\leqq 1$, *we have*

$$\left|1-\sum_{k=1}^{n}\mu_k(x)\right| < 3$$

and

$$(1-x^2)\left|1-\sum_{k=1}^{n}\mu_k(x)\right| < \frac{6}{2n+1}.$$

PROOF. From (3. 1) and (1. 2), we get

$$(1+x)Y_{n-1}(x,x) = 1+\frac{(n+\frac{1}{2})\sin\theta}{\sin(n-\frac{1}{2})\theta\cos(n+\frac{1}{2})\theta}, \qquad x=\cos\theta$$

which on simplification, gives

$$1-[(1+x)Y_{n-1}(x,x)]^2 = -\frac{\cos(n+\frac{1}{2})\theta\sin(n-\frac{1}{2})\theta}{(n+\frac{1}{2})\sin\theta}\times \tag{3. 5}$$
$$\times\left[2+\frac{\cos(n+\frac{1}{2})\theta\sin(n-\frac{1}{2})\theta}{(n+\frac{1}{2})\sin\theta}\right].$$

Now

$$\left|\sin\left(n-\frac{1}{2}\right)\theta\right| \leqq \left(n-\frac{1}{2}\right)|\sin\theta| \text{ or } (2n-1)\left|\sin\frac{\theta}{2}\right| \quad \text{and} \quad |\cos n\theta|\leqq 1. \tag{3. 6}$$

Thus with the help of (3. 2) and (3. 6), (3. 5) gives

$$\left|1-\sum_{k=1}^{n}\mu_k(x)\right| \leqq \frac{(n-\frac{1}{2})}{(n+\frac{1}{2})}\left[2+\frac{(n-\frac{1}{2})}{(n+\frac{1}{2})}\right] < 3$$

and

$$(1-x^2)\left|1-\sum_{k=1}^{n}\mu_k(x)\right| \leqq \frac{1}{(n+\frac{1}{2})}\left[2+\frac{1}{(n+\frac{1}{2})}\right] < \frac{6}{(2n+1)}.$$

LEMMA 3. 3.

$$\left|\sqrt{\frac{1+x}{2}}X_{n-1}(x_k,x)\right| \leqq \frac{1}{2\cos^3\frac{\theta_k}{2}}+\frac{1}{\cos^2\frac{\theta_k}{2}\left|\sin\frac{\theta_k}{2}\right|\left|\sin\frac{\theta_k-\theta}{2}\right|},$$

where $X_{n-1}(x_k, x)$ *is given by* (2. 3), $x=\cos\theta$ *and* $\theta_k = \dfrac{2k-1}{2n+1}\pi$, $k=1,2,\dots,n$.

PROOF. From (2. 3) and (1. 2)

$$X_{n-1}(x_k,x) = \frac{1}{\left(n+\frac{1}{2}\right)}\sum_{j=1}^{n-1}\frac{\cos\left(j+\frac{1}{2}\right)\theta}{\cos\frac{\theta}{2}\sin\theta_k}\times$$
$$\times\left\{\frac{\left(j+\frac{1}{2}\right)\sin\left(j+\frac{1}{2}\right)\theta_k\cos\frac{\theta_k}{2}-\frac{1}{2}\cos\left(j+\frac{1}{2}\right)\theta_k\sin\frac{\theta_k}{2}}{\cos^2\frac{\theta_k}{2}}\right\},$$

where $x=\cos\theta$. Therefore,

$$\left|\sqrt{\frac{1+x}{2}}X_{n-1}(x_k,x)\right| \leqq$$
$$\leqq \frac{1}{(2n+1)}\left|\sum_{j=1}^{n-1}\frac{(j+\frac{1}{2})\{\sin(j+\frac{1}{2})(\theta_k+\theta)+\sin(j+\frac{1}{2})(\theta_k-\theta)\}}{\sin\theta_k\cos\frac{\theta_k}{2}}\right| + \tag{3. 7}$$
$$+\frac{1}{4(2n+1)}\left|\sum_{j=1}^{n-1}\frac{\cos(j+\frac{1}{2})(\theta_k-\theta)+\cos(j+\frac{1}{2})(\theta_k+\theta)}{\cos^3\frac{\theta_k}{2}}\right|.$$

Further, we refer to the summation formulae

$$\sum_{j'=1}^{n-1} \cos\left(j'+\frac{1}{2}\right)t = \frac{\sin nt - \sin t}{2\sin\frac{t}{2}} \tag{3.8}$$

and

$$\sum_{j'=1}^{n-1} \left(j'+\frac{1}{2}\right)\sin\left(j'+\frac{1}{2}\right)t = -\frac{(n\cos nt - \cos t)\sin\frac{t}{2} - \frac{1}{2}\cos\frac{t}{2}(\sin nt - \sin t)}{2\sin^2\frac{t}{2}} \tag{3.9}$$

(the second we obtain by differentiating the first).

Next, from (3. 8) and the inequalities

$$|\sin\theta| \leqq 1, \quad |\sin n\theta| \leqq n|\sin\theta| \quad \text{and} \quad |\cos n\theta| \leqq 1, \tag{3.10}$$

we get

$$\left|\sum_{j=1}^{n-1} \{\cos(j+\tfrac{1}{2})(\theta_k-\theta) + \cos(j+\tfrac{1}{2})(\theta_k+\theta)\}\right| \leqq 2(n+1). \tag{3.11}$$

Similarly, from (3. 9) and (3. 10), we have

$$\begin{aligned} &\left|-\sum_{j=1}^{n-1}\left(j+\frac{1}{2}\right)\left[\sin\left(j+\frac{1}{2}\right)(\theta_k-\theta) + \sin\left(j+\frac{1}{2}\right)(\theta_k+\theta)\right]\right| \leqq \\ &\leqq \frac{(n+1)}{\left|\sin\frac{\theta_k-\theta}{2}\right|} + \frac{(n+1)}{\left|\sin\frac{\theta_k+\theta}{2}\right|}. \end{aligned} \tag{3.12}$$

Thus (3. 7) with the help of (3. 11) and (3. 12) completes the proof of the lemma.

4. In this section we shall estimate the sum

$$\sum_{k=1}^{n}\left(\frac{1+x}{1+x_k}\right)^2 l_k^4(x)|x-x_k|.$$

Let

$$f_k \overset{\text{def}}{=} \left(\frac{1+x}{1+x_k}\right)^2 l_k^4(x)|x-x_k| \tag{4.1}$$

and

$$x = \cos\theta, \quad \theta_i = \frac{2i-1}{2n+1}\pi, \qquad i = 1, 2, \ldots, (n+1),$$

then (2. 4) and (3. 3) reduce (4. 1) to

$$f_k = \frac{\sin^4\theta_k \cos^4(n+\frac{1}{2})\theta}{(n+\frac{1}{2})^4|(\cos\theta - \cos\theta_k)^3|}. \tag{4.2}$$

We now need the following

LEMMA 4. 1. *In the interval* $[x_{i+1}, x_i]$, *we have*

$$f_k \leqq \begin{cases} \dfrac{1}{(n+\frac{1}{2})(i-k)^3} & \text{for } 1 \leqq k < i \leqq n \\ \dfrac{1}{(n+\frac{1}{2})(k-i-1)^3} & \text{for } 1 \leqq i \leqq n-2, \quad i+2 \leqq k \leqq n \end{cases} \tag{4.3}$$

$$f_i \leqq \frac{1}{(n+\frac{1}{2})} \quad \text{for} \quad 1 \leqq i \leqq n \tag{4.4}$$

$$f_{i+1} \leqq \frac{8}{(n+\frac{1}{2})} \quad \text{for} \quad 1 \leqq i \leqq n-1 \tag{4.5}$$

and

$$\sum_{k=1}^{n} f_k < \frac{13}{(n+\frac{1}{2})} \quad \text{for} \quad -1 \leqq x \leqq 1. \tag{4.6}$$

PROOF. By the elementary inequality

$$\frac{\sin\theta_k}{|\cos\theta-\cos\theta_k|} \leqq \frac{1}{\left|\sin\dfrac{\theta_k-\theta}{2}\right|} \tag{4.7}$$

we have

$$f_k \leqq \frac{1}{\left(n+\dfrac{1}{2}\right)^4 \left|\sin^3\dfrac{\theta_k-\theta}{2}\right|} \quad \text{by (4.2).} \tag{4.8}$$

If $x_{i+1} \leqq x \leqq x_i$, then $\theta_i \leqq \theta \leqq \theta_{i+1}$. For $1 \leqq k < i \leqq n$, we have $\theta_i - \theta_k \leqq \theta - \theta_k \leqq \leqq \pi$ and

$$\sin\frac{\theta-\theta_k}{2} \geqq \sin\frac{\theta_i-\theta_k}{2} = \sin\left(\frac{i-k}{2n+1}\right)\pi > \frac{i-k}{(n+\frac{1}{2})}. \tag{4.9}$$

Also, if $1 \leqq i \leqq n-2$, $i+2 \leqq k \leqq n$, then $\theta_k - \theta_{i+1} \leqq \theta_k - \theta \leqq \pi$ and

$$\sin\frac{\theta_k-\theta}{2} \geqq \sin\frac{\theta_k-\theta_{i+1}}{2} = \sin\left(\frac{k-i-1}{2n+1}\right)\pi > \frac{k-i-1}{(n+\frac{1}{2})}. \tag{4.10}$$

From (4. 8), (4. 9) and (4. 10), we easily get (4. 3).

In order to prove (4. 4), we observe that

$$\begin{aligned} \left|\cos\left(n+\frac{1}{2}\right)\theta\right| &= \left|\cos\left(n+\frac{1}{2}\right)\theta - \cos\left(n+\frac{1}{2}\right)\theta_i\right| \\ &= 2\left|\sin\left(n+\frac{1}{2}\right)\left(\frac{\theta+\theta_i}{2}\right)\right| \left|\sin\left(n+\frac{1}{2}\right)\left(\frac{\theta_i-\theta}{2}\right)\right| \\ &\leqq 2\left|\sin\left(n+\frac{1}{2}\right)\left(\frac{\theta_i-\theta}{2}\right)\right| < 2\left(n+\frac{1}{2}\right)\left|\sin\frac{\theta_i-\theta}{2}\right| \end{aligned} \tag{4.11}$$

and

$$\sin\theta_i \leqq \sin\frac{\theta+\theta_i}{2} \quad \text{since} \quad \theta_i \leqq \theta. \tag{4.12}$$

Now (4. 2) for $k=i$, (4. 11) and (4. 12) easily give

$$f_i \leqq \frac{|\cos(n+\frac{1}{2})\theta|\,|\sin\theta_i|}{(n+\frac{1}{2})} \leqq \frac{1}{(n+\frac{1}{2})}.$$

Also (4. 2) and (4. 7) for $k = i+1$ and (4. 11) with i replaced by $i+1$ easily prove the inequality (4. 5).

Lastly, to prove (4. 6), we have

$$\sum_{k=1}^{n} f_k = \sum_{k=1}^{i-1} f_k + f_i + f_{i+1} + \sum_{k=i+2}^{n} f_k,$$

which on account of (4. 3), (4. 4) and (4. 5) gives (4. 6) for $-1 \leqq x < 1$ and by continuity it is valid for $x=1$.

5. In this section we shall estimate

$$\sum_{k=1}^{n}\left(\frac{1+x}{1+x_k}\right)^2 l_k^4(x)\,\frac{|1-2x_k|}{(1-x_k^2)}\,(x-x_k)^2.$$

Let

$$f_k^* \overset{\text{def}}{=} \left(\frac{1+x}{1+x_k}\right)^2 l_k^4(x)\,\frac{|1-2x_k|}{(1-x_k^2)}\,(x-x_k)^2, \qquad 1 \leqq k \leqq n \tag{5.1}$$

$$= \frac{\cos^4(n+\frac{1}{2})\theta\,\sin^2\theta_k\,|1-2\cos\theta_k|}{(n+\frac{1}{2})^4(\cos\theta-\cos\theta_k)^2}, \quad \text{when} \quad x=\cos\theta$$

$$\leqq \frac{3\cos^4(n+\frac{1}{2})\theta\,\sin^2\theta_k}{(n+\frac{1}{2})^4(\cos\theta-\cos\theta_k)^2}. \tag{5.2}$$

LEMMA 5. 1. *In the interval* $[x_{i+1}, x_i]$, *we have*

$$f_k^* \leqq \begin{cases} \dfrac{3}{(n+\frac{1}{2})^2(i-k)^2} & \textit{for} \quad 1 \leqq k < i \leqq n \\[2ex] \dfrac{3}{(n+\frac{1}{2})^2(k-i-1)^2} & \textit{for} \quad 1 \leqq i \leqq n-2, \quad i+2 \leqq k \leqq n \end{cases} \tag{5.3}$$

$$f_i^* \leqq \frac{3}{(n+\frac{1}{2})^2} \quad \textit{for} \quad 1 \leqq i \leqq n \tag{5.4}$$

$$f_{i+1}^* \leqq \frac{12}{(n+\frac{1}{2})^2} \quad \textit{for} \quad 1 \leqq i \leqq n-1 \tag{5.5}$$

and

$$\sum_{k=1}^{n} f_k^* \leqq \frac{25}{(n+\frac{1}{2})^2} \quad \textit{for} \quad -1 \leqq x \leqq 1. \tag{5.6}$$

As the proof of this lemma is exactly similar to that of lemma 4. 1, we omit the details.

6. In this section we shall estimate

$$\sum_{k=1}^{n}\left(\frac{1+x}{1+x_k}\right)^2 |(1+x_k)l_k^3(x)X_{n-1}(x_k,x)|(x-x_k)^2.$$

Let

$$f_k^{**}\overset{\text{def}}{=}\left(\frac{1+x}{1+x_k}\right)^2 |(1+x_k)l_k^3(x)X_{n-1}(x_k,x)|(x-x_k)^2, \qquad 1\leqq k\leqq n.$$

Then putting $x=\cos\theta$, we have

$$f_k^{**}=\frac{8\sin^2\frac{\theta_k}{2}\cos^3\frac{\theta_k}{2}\cos^3\left(n+\frac{1}{2}\right)\theta\,|\sin\theta_k|\left|\sqrt{\frac{1+x}{2}}X_{n-1}(x_k,x)\right|}{\left(n+\frac{1}{2}\right)^3|\cos\theta-\cos\theta_k|}$$

$$\text{(6. 1)}\qquad \leqq\frac{8\cos^3\left(n+\frac{1}{2}\right)\theta\,|\sin\theta_k|}{\left(n+\frac{1}{2}\right)^3|\cos\theta-\cos\theta_k|}\left[\frac{1}{2}+\frac{1}{\left|\sin\frac{\theta_k-\theta}{2}\right|}\right] \quad\text{by lemma 3. 3}$$

$$\leqq\frac{4\cos^3\left(n+\frac{1}{2}\right)\theta}{\left(n+\frac{1}{2}\right)^3}\left[\frac{1}{\left|\sin\frac{\theta_k-\theta}{2}\right|}+\frac{2}{\sin^2\frac{\theta_k-\theta}{2}}\right] \quad\text{by (4. 7)}$$

$$\leqq\frac{12\cos^3\left(n+\frac{1}{2}\right)\theta}{\left(n+\frac{1}{2}\right)^3\sin^2\frac{\theta_k-\theta}{2}}.$$

LEMMA 6. 1. *In the interval* $[x_{i+1}, x_i]$, *we have*

$$\text{(6. 2)}\qquad f_k^{**}\leqq\begin{cases}\dfrac{12}{(n+\frac{1}{2})(i-k)^2} & \textit{for}\quad 1\leqq k<i\leqq n\\[2ex] \dfrac{12}{(n+\frac{1}{2})(k-i-1)^2} & \textit{for}\quad 1\leqq i\leqq n-2,\quad i+2\leqq k\leqq n\end{cases}$$

$$\text{(6. 3)}\qquad f_i^{**}\leqq\frac{48}{(n+\frac{1}{2})}\quad \textit{for}\quad 1\leqq i\leqq n$$

$$\text{(6. 4)}\qquad f_{i+1}^{**}\leqq\frac{48}{(n+\frac{1}{2})}\quad \textit{for}\quad 1\leqq i\leqq n-1$$

and

$$\text{(6. 5)}\qquad \sum_{k=1}^{n}f_k^{**}<\frac{136}{(n+\frac{1}{2})}\quad \textit{for}\quad -1\leqq x\leqq 1.$$

The lemma easily follows on the same lines as lemma 4. 1 by using (6. 1).

7. LEMMA 7. 1. *For* $-1\leqq x\leqq 1$, *we have*

$$\sum_{k=1}^{n}|x-x_k|\,|\mu_k(x)|<\frac{174}{(n+\frac{1}{2})}.$$

Which obviously follows on account of (2. 1), (4. 6), (5. 6) and (6. 5).

8. Estimation of the sum

$$\sum_{k=1}^{n}\left(\frac{1+x}{1+x_k}\right)^2 l_k^4(x).$$

If we denote $\left(\frac{1+x}{1+x_k}\right)^2 l_k^4(x)$ by g_k, then putting $x=\cos\theta$ and using (2. 4), we have

$$g_k=\frac{\cos^4(n+\frac{1}{2})\theta\sin^4\theta_k}{(n+\frac{1}{2})^4(\cos\theta-\cos\theta_k)^4}. \tag{8. 1}$$

LEMMA 8. 1. *In the interval* $[x_{i+1}, x_i]$, *we have*

$$g_k\leqq\begin{cases}\dfrac{1}{(i-k)^4} & \textit{for}\quad 1\leqq k<i\leqq n\\[2mm] \dfrac{1}{(k-i-1)^4} & \textit{for}\quad 1\leqq i\leqq n-2,\quad i+2\leqq k\leqq n\end{cases} \tag{8. 2}$$

$$g_i\leqq 16 \quad \text{for} \quad 1\leqq i\leqq n \tag{8. 3}$$

$$g_{i+1}\leqq 16 \quad \text{for} \quad 1\leqq i\leqq n-1 \tag{8. 4}$$

and

$$\sum_{k=1}^{n} g_k<36 \quad \textit{for} \quad -1\leqq x\leqq 1. \tag{8. 5}$$

9. Estimation of the sum

$$\sum_{k=1}^{n}\left(\frac{1+x}{1+x_k}\right)^2 l_k^4(x)\,|1-2x_k|\,\frac{|x-x_k|}{(1-x_k^2)}.$$

Let

$$g_k^*=\left(\frac{1+x}{1+x_k}\right)^2 l_k^4(x)\,|1-2x_k|\,\frac{|x-x_k|}{(1-x_k^2)}.$$

Then $x=\cos\theta$ obviously gives

$$g_k^*\leqq\frac{3\cos^4(n+\frac{1}{2})\theta\sin^2\theta_k}{(n+\frac{1}{2})^4|(\cos\theta-\cos\theta_k)^3|} \tag{9. 1}$$

$$\leqq\frac{3\cos^4\left(n+\dfrac{1}{2}\right)\theta}{\left(n+\dfrac{1}{2}\right)^4|\cos\theta-\cos\theta_k|\sin^2\dfrac{\theta_k-\theta}{2}} \quad \text{by (4. 7).}$$

LEMMA 9. 1. *In the interval* $[x_{i+1}, x_i]$, *we have*

$$g_k^* \leq \begin{cases} \dfrac{3}{(i-k)^2} & \text{for } 1 \leq k < i \leq n \\ \dfrac{3}{(k-i-1)^2} & \text{for } 1 \leq i \leq n-2, \quad i+2 \leq k \leq n \end{cases} \tag{9.2}$$

$$g_i^* \leq 12 \quad \text{for} \quad 1 \leq i \leq n \tag{9.3}$$

$$g_{i+1}^{**} \leq 12 \quad \text{for} \quad 1 \leq i \leq n-1 \tag{9.4}$$

and in the interval $-1 \leq x \leq 1$, *we have*

$$\sum_{k=1}^{n} g_k^* < 34. \tag{9.5}$$

10. Estimation of the sum

$$\sum_{k=1}^{n} \left(\frac{1+x}{1+x_k}\right)^2 |(1+x_k)(x-x_k) l_k^3(x) X_{n-1}(x_k, x)|.$$

Let

$$g_k^{**} = \left(\frac{1+x}{1+x_k}\right)^2 |(1+x_k)(x-x_k) l_k^3(x) X_{n-1}(x_k, x)|, \qquad 1 \leq k \leq n.$$

Then owing to $x = \cos\theta$ and (2. 4), we get

$$g_k^{**} = \frac{\left|\cos^3\left(n+\frac{1}{2}\right)\theta\right| \sin^4\theta_k}{\left(n+\frac{1}{2}\right)^3 \left|\sin\frac{\theta_k}{2}\right| (\cos\theta-\cos\theta_k)^2} \left|\sqrt{\frac{1+x}{2}}\, X_{n-1}(x_k, x)\right|$$

$$\leq \frac{\left|\cos^3\left(n+\frac{1}{2}\right)\theta\right| \sin^4\theta_k}{\left(n+\frac{1}{2}\right)^3 (\cos\theta-\cos\theta_k)^2 \left|\sin\frac{\theta_k}{2}\right|} \left[\frac{1}{2\cos^3\frac{\theta_k}{2}} + \frac{1}{\cos^2\frac{\theta_k}{2}\left|\sin\frac{\theta_k}{2}\right| \left|\sin\frac{\theta_k-\theta}{2}\right|}\right]$$

by lemma 3. 3.

$$\leq \frac{\left|\cos^3\left(n+\frac{1}{2}\right)\theta\right|}{\left(n+\frac{1}{2}\right)^3 \sin^2\frac{\theta_k-\theta}{2}} \left[\frac{4\sin^2\frac{\theta_k}{2}}{|\sin\theta_k|} + \frac{4}{\left|\sin\frac{\theta_k-\theta}{2}\right|}\right], \quad \text{since} \quad |\sin\theta_k| > \left|\sin\frac{\theta_k-\theta}{2}\right|$$

$$\leq \frac{8\left|\cos^3\left(n+\frac{1}{2}\right)\theta\right|}{\left(n+\frac{1}{2}\right)^3 \left|\sin^3\frac{\theta_k-\theta}{2}\right|}. \tag{10.1}$$

LEMMA 10. 1. *In the interval* $[x_{i+1}, x_i]$, *we have*

$$(10.2)\qquad g_k^{**} \leqq \begin{cases} \dfrac{8}{(i-k)^3} & \text{for } 1 \leqq k < i \leqq n \\ \dfrac{8}{(k-i-1)^3} & \text{for } i+2 \leqq k < n, \quad 1 \leqq i \leqq n-2 \end{cases}$$

$$(10.3)\qquad g_i^{**} \leqq 64 \quad \text{for} \quad 1 \leqq i \leqq n$$

$$(10.4)\qquad g_{i+1}^{**} \leqq 64 \quad \text{for} \quad 1 \leqq i \leqq n-1$$

and

$$(10.5)\qquad \sum_{k=1}^{n} g_k^{**} < 154 \quad \text{for} \quad -1 \leqq x \leqq 1.$$

We omit the proofs of lemmas 8. 1, 9. 1 and 10. 1 as they run on the same lines as lemma 4. 1.

11. In this section we shall find an estimation of the sum $\sum_{k=1}^{n} |\mu_k(x)|$.

LEMMA 11. 1. *For* $-1 \leqq x \leqq 1$, *we have*

$$\sum_{k=1}^{n} |\mu_k(x)| < 378.$$

PROOF. From (2. 1) and (2. 2), we get

$$\sum_{k=1}^{n} |\mu_k(x)| = \sum_{k=1}^{n} g_k + \sum_{k=1}^{n} g_k^* + 2\sum_{k=1}^{n} g_k^{**}.$$

Now making use of (8. 5), (9. 5) and (10. 5), we get the required estimation.

12. *Proof of the theorem.*

Let $\omega(f, \delta)$ be the modulus of continuity of the function $f(x)$, then

$$(12.1)\qquad \omega(f, p\delta) \leqq (p+1)\omega(f, \delta), \quad \delta > 0.$$

Now

$$f(x) - S_n(f; x) = f(x) - \sum_{k=1}^{n} f(x)\mu_k(x) + \sum_{k=1}^{n} f(x)\mu_k(x) + xf(-1) -$$

$$-\sum_{k=1}^{n} (f(x_k) - xf(-1))\mu_k(x) = (f(x) + xf(-1))\left[1 - \sum_{k=1}^{n} \mu_k(x)\right] +$$

$$+\sum_{k=1}^{n} (f(x) - f(x_k))\mu_k(x) = \frac{1}{2}[(1+x)\{f(x)+f(-1)\} + (1-x)\{f(x)-f(-1)\}]\times$$

$$\times[1 - \mu_k(x)] + \sum_{k=1}^{n} (f(x) - f(x_k))\mu_k(x).$$

Further, since $\omega(f, \delta)$ is the modulus of continuity of $f(x)$, therefore,

$$\left|\frac{(1+x)}{2}(f(x)+f(-1))\left[1-\sum_{k=1}^{n}\mu_k(x)\right]\right| \leqq \frac{|1+x|}{2}\,\omega(f, |x-1|)\left|1-\sum_{k=1}^{n}\mu_k(x)\right|$$

$$\leqq \frac{|1+x|}{2}(1+(2n+1)|x-1|)\,\omega\left(f, \frac{1}{2n+1}\right)\left|1-\sum_{k=1}^{n}\mu_k(x)\right|$$

by the inequality (12. 1) replacing p by $(2n+1)|x-1|$ and δ by $\dfrac{1}{2n+1}$.

$$<\omega\left(f, \frac{1}{2n+1}\right)\left|1-\sum_{k=1}^{n}\mu_k(x)\right|+\left(n+\frac{1}{2}\right)\omega\left(f, \frac{1}{2n+1}\right)(1-x^2)\left|1-\sum_{k=1}^{n}\mu_k(x)\right|$$

(12. 3) $$<6\omega\left(f, \frac{1}{2n+1}\right) \text{ by lemma 3. 2.}$$

Similarly

(12. 4) $$\left|\frac{(1-x)}{2}(f(x)-f(-1))\left[1-\sum_{k=1}^{n}\mu_k(x)\right]\right| < 6\omega\left(f, \frac{1}{2n+1}\right).$$

Also,

$$\sum_{k=1}^{n}|\mu_k(x)|\,|f(x)-f(x_k)| \leqq \sum_{k=1}^{n}|\mu_k(x)|\,\omega(f, |x-x_k|)$$

$$\leqq \sum_{k=1}^{n}|\mu_k(x)|(1+(2n+1)|x-x_k|)\,\omega\left(f, \frac{1}{2n+1}\right)$$

by the inequality (12. 1) replacing p by $(2n+1)|x-x_k|$ and δ by $\dfrac{1}{2n+1}$

$$\leqq \omega\left(f, \frac{1}{2n+1}\right)\left[\sum_{k=1}^{n}|\mu_k(x)|+(2n+1)\sum_{k=1}^{n}|x-x_k|\,|\mu_k(x)|\right]$$

(12. 5) $$<\omega\left(f, \frac{1}{2n+1}\right)(378+348) = 726\omega\left(f, \frac{1}{2n+1}\right),$$

by lemma 7. 1. and lemma 11. 1.

Hence (12. 2), (12. 3), (12. 4) and (12. 5) give

$$|f(x)-S_n(f; x)| < 738\omega\left(f, \frac{1}{2n+1}\right),$$

which completes the proof of the theorem.

In a subsequent paper, I shall prove that the interpolation process given by (2. 5) is of Timan's type.

I am very much thankful to the referee* for his helpful suggestions.

* Professor G. Freud

REFERENCES

[1] Freud, G.: Egy Jackson-féle interpolációs eljárásról, *Math. Lapok* **15** 4 (1964), 330—336.
[2] Freud, G. and Vértesi, P.: A new proof of A. F. Timan's approximation theorem, *Studia Sci. Math. Hung.* **2** (1967), 403—409.
[3] Kis, O. and Vértesi, P.: On a new interpolation process, *Annales Univ. Sci. Budapest* **10** (1967), 117—128.
[4] Sallay, M.: Über ein Interpolationsverfahren, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **9** (1964), 607—615.
[5] Saxena, R. B.: Polynomial of interpolation, *Studia Sci. Math. Hung.* **2** (1967), 167—183.
[6] Szegő, G.: *Orthogonal polynomials*, Amer. Math. Coll. Publ., New York (1959).
[7] Vértesi, P.: Jackson tételének bizonyítása interpolációs úton, *Math. Lapok* **18** (1967), 83—92.

Department of Mathematics and Astronomy, Lucknow University, Lucknow (INDIA).

(Received September 6, 1969. Revised August 24, 1970.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 113—116.

A REMARK CONCERNING POLYNOMIAL MATRICES

by

P. FISCHER

I. A rectangular matrix whose elements are polynoms of λ is called polynomial matrix. We assume that the coefficients of the polynoms are arbitrary complex numbers. A polynomial matrix is called an elementary polynomial one if this matrix possesses either a right or a left polynomial inverse. The elementary polynomial matrices appear at the examination of several technical problems as useful tools [1], [2], [3].

Throughout matrices will be denoted by upper-case letters, and its elements by the same lower-case letters. I denotes the unity matrix, and I_r the $r \times r$ unity matrix, and $R(A)$ denotes the rank of the matrix A.

Let us consider the following square matrices S^i and T^i $(i=1, 2, 3)$ defined in the following way: S^1 and S^2 and S^3 are different from I_n in the following elements:

$$s_{ii}^1 = c$$

$$s_{ij}^2 = b(\lambda)$$

$$s_{ii}^3 = s_{jj}^3 = 0, \quad s_{ij}^3 = s_{ji}^3 = 1$$

where $b(\lambda)$ is a polynom and c is a constant different from zero. Let the matrices T^i be the conjugate of the matrices S^i $(i=1, 2, 3)$. The so defined S^i (resp. T^i) are called elementary left- (resp. right-) sided matrices both of which together are referred to as elementary; or elementary matrices of order n.

The following theorem is known [4]. If the determinant of a square polynomial matrix $P(\lambda)$ is independent of λ and is a constant different from zero, then this matrix can be written as a product of a finite elementary matrix.

It is easy to see that a square polynomial matrix is an elementary one if and only if its determinant is a constant differing from zero.

Therefore it follows from the previously mentioned results, that a square polynomial matrix is an elementary one if and only if it can be written as a product of a finite elementary matrix, where the order of each elementary matrices is equal to the one of the original square polynomial matrix.

The purpose of this paper is to generalise this theorem for the case of non-square matrices, with such an addition, that we add to the set of elementary matrices the following two matrices also.

$$U = (u_{ij}) \qquad i = 1, \ldots, n, \quad n > r$$

$$j = 1, \ldots, r,$$

where $$u_{ij}=\delta_{ij}$$

$$V=(v_{ij}) \quad i=1, \dots, n, \quad n<r$$
$$j=1, \dots, r,$$

where $$v_{ij}=\delta_{ij}$$

δ_{ij} is the Kronecker symbolum.

For this, first we shall prove the following theorem:

THEOREM 1. *The necessary and sufficient condition for that an $A(\lambda)$ polynomial matrix be an elementary one is that it has an inverse matrix and $R[A(\lambda)]$ is a constant function.*

Remark 1. If $A(\lambda)$ is a square polynomial matrix the Theorem 1. expresses the same as one of the previously stated theorem.

Remark 2. Theorem 1. is a generalization of the following well-known theorem:

Let us take the polynomials

$$x_1(\lambda), \dots, x_n(\lambda)$$

then 1 can be written as a polynomial linear combination of the polynomials $x_1(\lambda), \dots, x_n(\lambda)$ if and only if their greatest common divisor is equal to 1.

Remark 3. From this theorem it follows: in order that an $(r\times n)$ polynomial matrix was an elementary one, the sufficient condition is that it possess an $r\times r$ subminor which is an elementary polynomial matrix.

PROOF. The condition is necessary. In fact, let $A(\lambda)$ be an $r\times n$ polynomial matrix. Then we can state without the loss of the generality that $r\leqq n$. According to this statement there exists an $(n\times r)$ polynomial matrix $B(\lambda)$ such that

$$A(\lambda)B(\lambda)=I_r \tag{1}$$

If there would be such a $\lambda=\lambda_0$ point, where $R[A(\lambda_0)]<r$, then on the basis concerning to the theorem on the rank of the product of matrices, we find that the inequality

$$R[A(\lambda_0)B(\lambda_0)]=R[I_r]\leqq\min\{R[A(\lambda_0)], R[B(\lambda_0)]\}<r \tag{2}$$

also holds, which is a contradiction.

The condition is sufficient. From the condition, that the rank of matrix $A(\lambda)$ at every point, is equal to r, it follows that the greatest common divisor of $r\times r$ subdeterminant of $A(\lambda)$ is equal to 1. From this we find that $i_r, \dots, i_1$ invariant divisors of the matrix $A(\lambda)$

$$i_r=\frac{D_r}{D_{r-1}}, \dots, i_{r-1}=\frac{D_{r-1}}{D_{r-2}}, \dots, i_1=\frac{D_1}{1} \tag{3}$$

are equal to 1 where D_k $(k=2, \dots, r)$ is the greatest common divisor of all $k\times k$ subdeterminants of $A(\lambda)$, because, as it is known, D_k is divisable by D_{k-1} $(k=2, \dots, r)$.

Let us consider the canonical form of the matrix $A(\lambda)$

$$A(\lambda) = \begin{bmatrix} A_1(\lambda) \end{bmatrix} \begin{bmatrix} i_1(\lambda) & & & \\ & \cdot & & \\ & & \cdot & \\ & & & i_r(\lambda) \end{bmatrix} \begin{bmatrix} A_2(\lambda) \end{bmatrix} \tag{4}$$

where the $A_1(\lambda)$ (resp. $A_2(\lambda)$) is an elementary polynomial matrix of the order r (resp. n).

So we have also

$$A(\lambda) = \begin{bmatrix} A_1(\lambda) \end{bmatrix} \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & 1 \end{bmatrix} \begin{bmatrix} A_2(\lambda) \end{bmatrix} \tag{5}$$

which proves our Theorem.

From the demonstration of the previous theorem we get immediately the following theorem.

THEOREM 2. *The $A(\lambda)(r\times n)$ $(r\leqq n)$ polynomial matrix is an elementary one if and only if it can be written in the form of a product of finite elementary matrices of the order r is multiplied on the left side by the matrix* $\begin{bmatrix} 1 & & & & \\ & 1 & & & 0 \\ & & \cdot & & \\ & & & \cdot & \\ & & & & 1 \end{bmatrix}$ *and their product is multiplied on the left side by the product of finite elementary matrices of the order n.*

II. The square elementary polynomial matrices can be characterized by the following Theorem.

Let us consider the following differential equations' system:

$$A(\lambda)x(t) = \begin{matrix} a_{11}(\lambda)x_1(t)+\cdots+a_{1n}(\lambda)x_n(t) = 0 \\ \vdots \\ a_{n1}(\lambda)x_1(t)+\cdots+a_{nn}(\lambda)x_n(t) = 0 \end{matrix} \tag{6}$$

where the polynomials $a_{ij}(\lambda)$ $(i,j=1,\ldots,n)$ are differential operators with constant coefficients with the initial conditions $x_1(0)=\cdots=x_n(0)=0$.

THEOREM 3. *The equation* (6) *has only a trivial solution if and only if the $A(\lambda)= =(a_{ij}(\lambda))$ matrix is an elementary polynomial matrix with the condition that the solution is differentiable up to a degree large enough.*

PROOF. This theorem is a consequence of the fact that if the functions $x_1(t), \ldots$ $\ldots, x_n(t)$ are differentiable up to a degree large enough then the equation (6) is equivalent to the equation

$$S^i A(\lambda) x(t) = 0 \tag{7}$$

for $i = 1, 2, 3$.

So if $A(\lambda)$ is an elementary polynomial matrix, then the equation (6) has only a trivial solution. If the $A(\lambda)$ is not an elementary polynomial matrix then it can be shown by induction that (6) has a non-trivial solution too.

REFERENCES

[1] YOULA, D. C.: On the Factorization of Rational Matrices, *IRE Trans. on Information Theory*, **15** (1961); Part 7, No. 3. 172—189.

[2] DAVIS, M. C.: Factoring the Spectral Matrix, *IEEE Trans. on Automatic Control* (1963), 296—305.

[3] CSÁKI, F.—FISCHER, P.: On the Spectrum Factorization, *Acta Technica Acad. Sci. Hungaricae*, **58** (1—2) (1967), 145—168.

[4] Гантмахер, Б. Р.: *Теория матриц*. Гос. Изд. Техникотеоретической Литературы, г. Москва, 1954.

(Received September 24, 1969.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 117—122.

A CLASS OF SPACES WITH IDENTICAL REMAINDERS

by

G. R. BLAKLEY[1], J. GERLITS and K. D. MAGILL, Jr.

1. Introduction. It is assumed that all topological spaces discussed in this paper are completely regular and Hausdorff. By a remainder of a space X, we mean any space of the form $\alpha X - X$ where αX is any compactification of X. We denote the family of all remainders of X by $\mathscr{R}(X)$. If one does not distinguish between homeomorphic spaces, and we shall not, then $\mathscr{R}(X)$ is actually a set whose cardinality does not exceed 2^{2^m} where m denotes the cardinality of X. This follows from the fact that 2^{2^m} is an upper bound for the number of different compactifications of X (see [6, p. 231]).

The third author has devoted several previous papers to the problem of determining when $\mathscr{R}(X)$ contains spaces with certain prescribed properties. For example, in [3, Theorem (2. 1), p. 1076], characterizations were given for those X which have the property that $\mathscr{R}(X)$ contains the space consisting of N points where N is a positive integer. It was shown in [4, Theorem (2. 1), p. 617] that for locally compact X, the family $\mathscr{R}(X)$ contains a countable space if and only if it contains all finite spaces. In [5, Theorem (2. 2), p. 323], it was shown that there exists a rather extensive class of spaces with the property that each such space includes all Peano continua among its remainders. These are the locally compact normal spaces which contain infinite, discrete, closed subsets. J. W. ROGERS [7] has recently extended this result by showing that any locally compact nonpseudocompact space includes all metric continua among its remainders.

This paper had its genesis in an attempt to answer the following question: Do the remainders of all Euclidean N-spaces coincide for $N>1$? It is easily seen that the answer is negative without the restriction $N>1$, for the real line E^1 has a two-point compactification while no E^N does whenever $N>1$ (this follows from [3, Theorem (2. 6), p. 1079]), i.e., $\mathscr{R}(E^1)$ contains the two-point space while $\mathscr{R}(E^N)$ does not. We have been able to show, however, that all the Euclidean spaces whose dimensions exceed one do have identical remainders. In fact, we show that if both X and Y are noncompact, locally compact, connected, locally connected metric spaces and neither have two-point compactifications, then the remainders of X coincide with those of Y.

2. The results. It will simplify our discussion considerably if we provide those spaces mentioned in the last sentence of section 1 with a name.

DEFINITION (2. 1). A space which is noncompact, locally compact, connected, locally connected and metric and has no two-point compactification will be referred to as a ringed space.

---

[1] Work supported, in part, by NSF Grant GP 12042.

As we noted previously, it follows from Theorem (2. 6) of [3, p. 1079] that if $N>1$, the Euclidean N-space E^N has no two-point compactification. It follows from that same theorem that the half-open interval [0, 1) whose topology is the usual topology induced by the real line has no two-point compactification. Thus, we have the following

PROPOSITION (2. 2). *The class of all ringed spaces includes all Euclidean N-spaces for* $N>1$ *as well as the half-open interval* [0, 1).

The half-open interval [0, 1) plays a rather central role in our subsequent deliberations and, because of this, we find it convenient to denote it hereafter by the symbol J. The family of all remainders of J will be denoted by $\mathcal{R}^*$. Our next result gives a sufficient condition in order that all of the remainders of one locally compact space be remainders of another. We then apply this to show that the family of remainders of many spaces includes all of the spaces in $\mathcal{R}^*$. As usual, the Stone—Čech compactification of a space X is denoted by βX.

THEOREM (2. 3). *Suppose that X and Y are locally compact and there exists a continuous function f mapping X onto Y with the property that for each point* $q\in Y$, *there exists an open subset G of Y containing q and a compact subset K of X such that* $f[X-K]\cap G=\emptyset$. *Then* $\mathcal{R}(Y)\subset\mathcal{R}(X)$.

PROOF. Let g denote the Stone—Čech extension of f which necessarily maps βX onto βY. We want to show that g maps $\beta X-X$ onto $\beta Y-Y$. Since g maps no points of X into $\beta Y-Y$, we have only to show that $g(p)\in\beta Y-Y$ for each $p\in\beta X-X$. Suppose, on the contrary, that $g(p)\in Y$ for some $p\in\beta X-X$. Then there exists an open subset G of Y containing $g(p)$ (which is also open in βY since Y is open in βY) and a compact subset K of X such that

$$g[X-K]\cap G=f[X-K]\cap G=\emptyset. \qquad (2.\ 3.\ 1)$$

By continuity of g at p, there exists an open subset V of βX containing p such that $g[V]\subset G$. Then $W=V\cap[\beta X-K]$ is also an open subset of βX which contains p and is mapped into G by the function g. Since X is dense in βX, there exists a point $t\in W\cap X$ and it follows from the previous remark that

$$g(t)=f(t)\in G.$$

But this contradicts (2. 3. 1) since $t\in X-K$. Consequently, $g(p)$ must belong to $\beta Y-Y$ and it follows that $\beta Y-Y$ is a continuous image of $\beta X-X$. Now Theorem (2. 1) of [5, p. 322] states that a space H is a remainder of a locally compact space T if and only if H is a continuous image of $\beta T-T$. Therefore, it follows that every remainder of Y is also a remainder of X.

Before stating our next result, we recall that a space is said to be σ-compact if it is the countable union of compact spaces.

THEOREM (2. 4). *If X is a locally compact, noncompact connected* σ*-compact space, then* $\mathcal{R}^*\subset\mathcal{R}(X)$.

PROOF. Let $X^*=X\cup\{\omega\}$ be the one-point compactification of X. Since X is σ-compact, the point ω is a G_δ set in X^* and, consequently, is a zero set. Hence, there exists a continuous function f mapping X^* into [0, 1] such that $f^{-1}(1)=\{\omega\}$. In fact, since X is connected, there is no real loss in generality in assuming that f

maps X^* onto [0, 1]. Thus, f maps X onto $J=[0, 1)$ and we show that it satisfies the conditions of Theorem (2. 3). Let y be any point in J and choose any point p such that $y<p<1$. Then $f^{-1}([0, p])$ is closed in X^* and is therefore compact. Thus $G=[0, p)$ is an open subset of J and $K=f^{-1}([0, p])$ is a compact subset of X with the property that $f[X-K]\cap G=\varnothing$. It now follows from Theorem (2. 3) that $\mathscr{R}^*\subset\mathscr{R}(X)$.

COROLLARY (2. 5). *If X is any ringed space, then $\mathscr{R}^*\subset\mathscr{R}(X)$.*

PROOF. A ringed space is locally compact, paracompact and connected. Consequently, it follows from Theorem 7. 3 of [1, p. 241] that X is σ-compact and we have only to apply Theorem (2. 4).

In view of the previous corollary, we must yet show that $\mathscr{R}(X)\subset\mathscr{R}^*$ for any ringed space X in order to get the main result of this paper. Our first step in this direction is to obtain a characterization of ringed spaces. First, some definitions and a lemma:

DEFINITION (2. 6). Let $\{K_n\}_{n=1}^{\infty}$ be a countable cover of an arcwise connected space X. The order (with respect to the given cover) of a pair of points p and q of X will be denoted by $O(p, q)$ and is defined as follows: if every arc joining p to q intersects K_1, then $O(p, q)=1$. Otherwise $O(p, q)$ is the largest integer N such that there exists an arc A joining p to q with the property that $A\cap K_j=\varnothing$ for each $j<N$.

There is no difficulty about the existence of the order of a pair of points p and q. Specifically, since $\{K_n\}_{n=1}^{\infty}$ is a cover, $p\in K_i$ for some i and it follows that $O(p, q)\leqq i$.

DEFINITION (2. 7). Again, let $\{K_n\}_{n=1}^{\infty}$ be a countable cover of an arcwise connected space X. When $K_i\neq\varnothing\neq K_{i+1}$, we define the complexity c_i of the pair (K_i, K_{i+1}) to be the least of all orders $O(p, q)$ where $p\in K_i$ and $q\in K_{i+1}$. Otherwise, define $c_i=i$.

We recall that any collection of subsets of a space is said to be locally finite provided that each point of the space belongs to an open subset which intersects only finitely many members of the collection.

DEFINITION (2. 8). A set of bands for an arcwise connected space is any locally finite countable cover $\{K_n\}_{n=1}^{\infty}$ where each K_j is a Peano continuum and $\lim c_i=\infty$.

LEMMA (2. 9). *Suppose X is noncompact, locally compact and locally connected. Then X has no two-point compactification if and only if for each compact subset C of X, there exists a compact subset K of X such that $C\subset K$ and $X-K$ is connected.*

PROOF. Sufficiency follows immediately from Theorem (2. 6) of [3, p. 1079] so we need only prove necessity here. Therefore, suppose X has no two-point compactification and let C be any compact subset of X. Denote the components of $X-C$ by $\{G_\alpha:\alpha\in\Lambda\}$. Since X is locally connected, each G_α is open. We assert that

(2. 9. 1) for some $\alpha_0\in\Lambda$, $C\cup G_{\alpha_0}$ is not compact.

We prove this by contradiction. If (2. 9. 1) does not hold, then $C\cup G_\alpha$ is compact for each $\alpha\in\Lambda$. This and the fact that X is not compact implies the existence of an open subset G of X such that $C\subset G$ and $G_\alpha\not\subset G$ for each α belonging to some infinite subset Λ_0 of Λ. Now let Λ_1 and Λ_2 be any two infinite subsets of Λ_0 such that

$\Lambda_1 \cap \Lambda_2 = \emptyset$ and $\Lambda_1 \cup \Lambda_2 = \Lambda_0$ and define

$$U_1 = \cup \{G_\alpha : \alpha \in \Lambda_1\},$$
$$U_2 = \cup \{G_\alpha : \alpha \in \Lambda - \Lambda_1\}.$$

Then U_1 and U_2 are nonintersecting open subsets of X whose union is $X-C$. Furthermore, $C \cup U_i$ $(i=1, 2)$ is not compact. Indeed $\{G\} \cup \{G_\alpha : \alpha \in \Lambda_i\}$ is an infinite cover of $C \cup U_i$ which has no proper subcover much less a finite one. All this, together with Theorem (2. 1) of [3, p. 1075] implies that X has a two-point compactification. This, of course, is the contradiction we seek and so we conclude that statement (2. 9. 1) is valid, i.e., $C \cup G_{\alpha_0}$ is not compact for some $\alpha_0 \in \Lambda$. Define

$$K = \cup \{G_\alpha : \alpha \neq \alpha_0\} \cup C.$$

Then $X - K = G_{\alpha_0}$ is connected and it follows from Theorem (2. 1) of [3, p. 1075] and the fact that X has no two-point compactification that K is compact.

Now we are in a position to characterize ringed spaces.

THEOREM (2. 10). *A topological space is a ringed space if and only if it is arcwise connected and has a set of bands.*

PROOF. First, suppose X is arcwise connected and has a set of bands $\{K_n\}_{n=1}^{\infty}$. Since the family $\{K_n\}_{n=1}^{\infty}$ is locally finite, there exists for each x in X an open neighborhood G of x such that only finitely many of the sets K_n, say $\{K_n\}_{n=1}^{N}$ intersect G. We may assume that $x \in K_n$ for $1 \leqq n \leqq M$ and $x \notin K_n$ for $M < n \leqq N$. Then

$$x \in G - [K_{M+1} \cup \ldots \cup K_N] \subset K_1 \cup \ldots \cup K_M.$$

That is, x belongs to the interior of $K_1 \cup \ldots \cup K_M$ which is a Peano continuum since $x \in \cap \{K_n\}_{n=1}^{M}$ and each K_n is a Peano continuum. It follows from this that X is locally compact and locally connected. Furthermore, X is metrizable since any (T_0) space is metrizable if it is the union of a locally finite family of closed metrizable subspaces [1, p. 207, section 9, problem 4].

The space X is not compact because a locally finite system in a compact space is necessarily finite. In order to conclude that X is a ringed space, we need only to show it has no two-point compactification. Let $C \subset X$ be a compact set. Since the family $\{K_n\}_{n=1}^{\infty}$ is locally finite, there exists an integer N such that $C \subset \cup \{K_n\}_{n=1}^{N}$ We choose an integer M such that for $m \geqq M$, $c_m > N$ holds. Now for each $m \geqq M$ there exists an arc A_m joining a point of K_m to a point of K_{m+1} and $A_m \cap K_i = \emptyset$ $(i=1, 2, \ldots, N)$. The set $H = \cup \{K_m \cup A_m\}_{m=M}^{\infty}$ is connected and $H \cap C = \emptyset$, hence $H \subset G$ where G is a suitable component of $(X-C)$. Since the subset G is open, it follows that the set $K = X - G \subset \cup \{K_i\}_{i=1}^{M-1}$ is compact, $C \subset K$ and $X-K$ is connected. Thus the Lemma (2. 9) implies that x has no two-point compactification.

Now suppose that X is a ringed space. Since X is locally compact, connected, locally connected and metric, it follows from [8, 5. 2, p. 38] that X is arcwise connected. Now, we want to construct a set of bands $\{K_n\}_{n=1}^{\infty}$ for X. It follows from Theorem 7. 3 of [1, p. 241] that X is σ-compact. Consequently, by Theorem 7. 2 of [1, p. 241], $X = \cup \{G_n\}_{n=1}^{\infty}$ where for each positive integer n, G_n is nonempty and open, cl G_n is compact and cl $G_n \subset G_{n+1}$. Let n be any integer greater than 2. Then since X is locally connected, each component of $G_{n+1} - \mathrm{cl}\, G_{n-2}$ is open and the collection of all such components is an open cover of cl $G_n - G_{n-1}$. Since the latter is compact some finite subfamily, say $\{C_i: i\in I_n\}$ is also a cover. Now each C_i is a closed subset of the space $\bigcup\{C_i: i\in I_n\}$ and thus $A_i = [\mathrm{cl}\, G_n - G_{n-1}]\cap C_i$ is compact. It follows from [8, p. 38, 5. 3] and [2, p. 183, § 45. I. 5] that there exists a Peano continuum L_i such that $A_i\subset L_i\subset C_i$ for each $i\in I_n$. Thus, we have

(2. 10. 1) $$\mathrm{cl}\, G_n - G_{n-1}\subset \bigcup\{L_i : i\in I_n\}\subset G_{n+1} - \mathrm{cl}\, G_{n-2}.$$

We need one more observation before we construct the set of bands $\{K_n\}_{n=1}^{\infty}$. An argument similar to that given above will lead one to the conclusion that

(2. 10. 2) $$\mathrm{cl}\, G_2\subset \bigcup\{L_i: i\in I_2\}\subset G_3$$

where each L_i is a Peano continuum. Now we are ready to define what will turn out to be a set of bands for X. The first I_2 terms of the sequence $\{K_n\}_{n=1}^{\infty}$ will be the members of the family $\{L_i: i\in I_2\}$. The next I_3 terms will be the members of the family $\{L_i: i\in I_3\}$ and so on. One readily shows that $\{K_n\}_{n=1}^{\infty}$ defined in this manner is a locally finite cover for X. We need only show that $\lim c_i=\infty$ where c_i denotes the complexity of the pair (K_i, K_{i+1}). Let any natural number n be given and let $C=\bigcup\{K_i\}_{i=1}^{n}$. By Lemma (2. 9), there exists a compact set W such that $C\subset W$ and $X-W=G$ is connected. Since the family $\{K_i\}_{i=1}^{\infty}$ is locally finite, the set $\{i: K_i\cap W\neq\emptyset\}$ is finite and we take M to be the largest integer in that set. Then for $m>M$, $K_m\subset G$ and since G is arcwise connected, it follows that $c_m>n$.

With this characterization, we are ready to prove the main result of the paper.

THEOREM (2. 11). *The remainders of all ringed spaces are identical.*

PROOF. We show that if X is any ringed space, then $\mathscr{R}(X)=\mathscr{R}^*$. In view of Corollary (2. 5), we need only show that $\mathscr{R}(X)\subset\mathscr{R}^*$ and for this, we appeal to Theorems (2. 3) and (2. 10). By the latter, X is arcwise connected and there exists a set of bands $\{K_n\}_{n=1}^{\infty}$ for X. For each positive integer n, define $I_n = [(n-1)/n, n/(n+1)]$. Since each K_n is a Peano continuum, there exists a continuous function f_{2n-1} mapping I_{2n-1} onto K_n for each positive integer n. Furthermore, it follows from the definition of a set of bands that for each n and each pair of points p and q with $p\in K_n$ and $q\in K_{n+1}$, there exists an arc A joining p to q such that $A\cap K_j=\emptyset$ for $j<c_n$ where c_n denotes the complexity of the pair (K_n, K_{n+1}). Consequently, there exists a homeomorphism f_{2n} from I_{2n} into X such that

$$f_{2n}((2n-1)/2n) = f_{2n-1}((2n-1)/2n),$$
$$f_{2n}(2n/(2n+1)) = f_{2n+1}(2n/(2n+1))$$

and

(2. 11. 1) $$f_{2n}[I_{2n}]\cap K_j=\emptyset \quad \text{for} \quad j<c_n.$$

Let f be the function which is defined by $f(x)=f_{2n-1}(x)$ for $x\in I_{2n-1}$ and $f(x)=f_{2n}(x)$ for $x\in I_{2n}$. It is immediate that f maps J continuously onto X. Now let p be any point of X. Since $\{K_n\}_{n=1}^{\infty}$ is locally finite, there exists an open subset G of X containing p and a positive integer N such that $G\cap K_j=\emptyset$ for $j>N$. Since $\lim c_n=\infty$, there exists a positive integer M such that $c_n>N$ for $n>M$. Let $T=\max\{N, M\}$ and define

$$H=\bigcup\{I_j\}_{j=1}^{2T}=[0, 2T/(2T+1)].$$

Then H is a compact subset of J and we assert that

$$(2.\ 11.\ 2) \qquad f[J-H]\cap G=\emptyset.$$

Take any point $x\in J-H$ and consider two cases: (1) $x\in I_{2n-1}$ and (2) $x\in I_{2n}$. Suppose (1) holds. It is immediate that $T<n$. Hence $N<n$ and

$$f_{2n-1}[I_{2n-1}]\cap G=K_n\cap G=\emptyset.$$

Thus $f(x)\notin G$ when $x\in I_{2n-1}$. Now consider the case where $x\in I_{2n}$. Again, $T<n$. Hence, $n>M$ and it follows that $c_n>N$. This implies that

$$f_{2n}[I_{2n}]\cap K_j=\emptyset \quad \text{for} \quad j\leqq N.$$

Since $G\subset\bigcup\{K_j\}_{j=1}^N$, it is true in this case also that $f(x)\notin G$. Thus (2. 11. 2) is valid and it now follows from Theorem (2. 3) that each remainder of X is also remainder of J.

Perhaps it is worthwhile to mention that, among other things, Theorem (2. 11) implies that in order to study the family of remainders of any ringed space, it is sufficient to devote one's attention to $\mathscr{R}^*$, the family of remainders of J. In conclusion, we make the observation that the class of all ringed spaces is by no means maximal with respect to the property of having identical remainders. Let X be the discrete union of any ringed space Y and any compact space. It is a routine matter to check that $\beta X-X=\beta Y-Y$ and it follows immediately from Theorem (2. 1) of [5, p. 322] that $\mathscr{R}(X)=\mathscr{R}(Y)$. However, X is certainly not a ringed space since it is not even connected.

REFERENCES

[1] DUGUNDJI, J.: *Topology*, Allyn and Bacon, Inc., Boston, 1966.
[2] KURATOWSKI, C.: *Topologie*, Vol. II, Warszava, 1961.
[3] MAGILL, K. D. JR.: *N*-point compactifications, *Amer. Math. Monthly*, **72**, No. 10 (1965), 1075—1081.
[4] MAGILL, K. D. JR.: Countable compactifications, *Canadian Journ. of Math.* **18** (1966), 616—620.
[5] MAGILL, K. D. JR.: A note on compactifications, *Math. Zeitschrift* **94** (1966), 322—325.
[6] MAGILL, K. D. JR.: The lattice of compactifications of a locally compact space, *Proc. London Math. Soc.*, Third Series XVIII (1968) 213—244.
[7] ROGERS, J. W.: On compactifications with continua as remainders, *Fund. Math* **70** (1971), 7—11.
[8] WHYBURN, G. T.: *Analytic topology*, Amer. Math. Soc., Colloquium Publications, Vol. XXVII, 1963.

Texas A and M University
Mathematical Institute of the Hungarian Academy of Sciences
State University of New York at Buffalo

(Received November 24, 1969. Revised March 1, 1970.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 123—125.

ИЗУЧЕНИЕ ОДНОЙ СИСТЕМЫ ГРУППОВОГО ОБСЛУЖИВАНИЯ

С. Х. СИРАЖДИНОВ, Т. А. АЗЛАРОВ, В. БАЗАРОВ

Рассматривается следующая система обслуживания m — размерным прибором. Требования поступают к месту ожидания по закону Пуассона с параметром λ. Прибор приступает к обслуживанию только в случайные моменты времени $t_1, t_2, \dots, t_k \dots$. Требования обслуживаются в порядке их поступления. Интервалы времени $\tau_k = t_{k+1} - t_k$, $k \geqq 1$ между $k-m$ и $k+1-m$ моментами начала обслуживания, будем предполагать независимыми и одинаково распределенными случайными величинами с общей функцией распределения $B(x) = \mathsf{P}(\tau_k < x)$ и с конечным математическим ожиданием

$$b = \mathsf{M}\tau_k = \int_0^\infty x\, dB(x) < \infty$$

Количество требований, имеющихся в приборе к началу k-го обслуживания' обозначим через η_k и будем считать, что $\eta_1, \eta_2, \dots$ — последовательность независимых, одинаково распределенных случайных величин:

$$q_i = \mathsf{P}(\eta_k = i), \quad 0 \leqq i \leqq m; \qquad a = \mathsf{M}\eta_k = \sum_{i=1}^{m} i q_i.$$

Следовательно, $m - \eta_k$ — число свободных мест в приборе к началу k-го обслуживания.

Число требований, прибывающих к месту ожидания в интервале (t_k, t_{k+1}), $k \geqq 1$ обозначим через $\zeta_k = \zeta(\tau_k)$. Тогда очевидно, что ζ_k имеет распределение

$$\pi_j = \mathsf{P}(\zeta_k = j) = \frac{1}{j!} \int_0^\infty e^{-\lambda x} (\lambda x)^j\, dB(x), \qquad j = 0, 1, 2, \dots$$

с математическим ожиданием

$$\mathsf{M}\zeta_k = \lambda b.$$

Необходимо заметить, что такая система часто встречается в жизни (городской транспорт, лифт, железнодорожная сортировочная станция и т. д.).

В предлагаемой статье изучаются стационарные распределения следующих основных характеристик:

а) числа требований ξ, находящихся в очереди в момент начала обслуживания; *б)* время ожидания w начала обслуживания произвольного требования.

Как и в работе [1], можно показать, что эти стационарные распределения существуют и не зависят от начального состояния системы в том и только в том случае, когда

$$\delta = \mathsf{M}(m-\eta_k)-\mathsf{M}\zeta_k = m-a-\lambda b > 0.$$

Пусть

$$P_n = \mathsf{P}(\xi=n), \quad n=0,1,\ldots; \qquad F(x)=\mathsf{P}(w<x);$$

$$P(z) = \sum_{n=0}^{\infty} P_n z^n; \quad q(z) = \sum_{i=0}^{m} q_i z^i;$$

$$\beta(s) = \int_0^\infty e^{-sx}\,dB(x); \quad f(s) = \int_0^\infty e^{-sx}\,dF(x).$$

Теорема 1. *В круге* $|z| \geqq 1$ *имеет место соотношение*

$$P(z) = \frac{\delta(z-1)\beta(\lambda(1-z))}{z^m-\beta(\lambda(1-z))q(z)} \prod_{r=1}^{m-1}\left(\frac{z-z_r}{1-z_r}\right), \tag{1}$$

где $z_1, z_2, \ldots, z_{m-1}$ — *корни уравнения* $z^m-\beta(\lambda(1-z))q(z)=0$ *находящиеся внутри единичного круга.*

Теорема 2. *В полуплоскости* $\operatorname{Re} s \geqq 0$ *справедливо соотношение*

$$f(s) = \frac{P\left(1-\frac{s}{\lambda}\right)}{b\cdot s}\left\{\frac{1}{\beta(s)}-1\right\}. \tag{2}$$

Математические ожидания случайных величин ξ *и* w *определяются следующими формулами:*

$$L = \mathsf{M}\xi = \frac{A^2+\lambda^2B^2+\lambda b}{2\delta} - \frac{2a+\delta-1}{2} + \sum_{r=1}^{m-1}\frac{1}{1-z_r}; \tag{3}$$

$$T = \mathsf{M}w = \frac{A^2+\lambda^2B^2+\lambda b}{2\lambda\delta} - \frac{(m+a)b-\lambda(1+B^2)}{2\lambda} + \sum_{r=1}^{m-1}\frac{1}{1-z_r}, \tag{4}$$

где

$$A^2 = \mathsf{D}\eta_k, \quad B^2 = \mathsf{D}\tau_k.$$

Обе теоремы являются обобщением соответствующих результатов Бейли [2] и Доунтона [3], полученных в предположении, что прибор в моменты начала обслуживания всегда бывает свободным.
В формулах (1)—(4) это соответствует случаю $q_0=1$, $q_i=0$, $i\neq 0$. Если $m=1$, $q_0=1$, $q_1=0$, то из формулы (1) следует [4] классический результат

$$P(z) = \frac{(1-\lambda b)(z-1)\beta(\lambda(1-z))}{z-\beta(\lambda(1-z))}.$$

полученный для системы $M|G|1$.

Однако преобразование Лапласа—Стильтьеса для времени ожидания,

$$f(s) = \frac{1-\dfrac{1}{\lambda b}}{1-\dfrac{s}{\lambda[1-\beta(s)]}}.$$

вычисленное при тех же условиях отличается от соответствующего результата

$$f(s) = \frac{1-\lambda b}{1-\dfrac{\lambda[1-\beta(s)]}{s}}$$

полученного для системы $M|G|1$.

Это объясняется тем, что если в системе $M|G|1$ требование, заставшее прибор свободным начинает немедленно обслуживаться, то в рассматриваемой системе поступившее требование ожидает очередного момента начала обслуживания.

Как показывают формулы (3) и (4), в условиях «большой загрузки», те. когда $\delta\downarrow 0$ средняя длина очереди L и среднее время ожидания T неограниченно возрастают. Поэтому выяснение асимптотического поведения распределений соответствующим образом нормированных случайных величин ξ и w представляет значительный интерес.

Теорема 3. *Пусть* $\mathsf{D}\tau_k = B^2 < \infty$. *Тогда при* $\delta\downarrow 0$ *имеют место предельные соотношения:*

$$\lim_{\delta\downarrow 0} \mathsf{P}(\delta\xi < x) = \begin{cases} 1-e^{-\frac{2b^2 x}{(m-a+A^2)b^2+(m-a)^2 B^2}}, & x>0 \\ 0, & x<0, \end{cases}$$

$$\lim_{\delta\downarrow 0} \mathsf{P}(\delta w < x) = \begin{cases} 1-e^{-\frac{2bx}{(m-a+A^2)b^2+(m-a)^2 B^2}}, & x>0 \\ 0, & x<0. \end{cases}$$

Две первые теоремы доказываются тем же способом, что и соответствующие результаты работ [2] и [3]. Однако при этом приходится усовершенствовать этот метод, чтобы его можно было применить в нашем более общем случае. А теорема 3 доказывается методом характеристических функций.

В заключение отметим, что аналогичные результаты можно получить и в том случае, когда допускается лишь конечный объем очереди.

ЛИТЕРАТУРА

[1] Finch, P. D.: A probability limit theorem with application to a generalization of queneing theory, *Acta Mat. Acad. Sci. Hungar.* **10** (1959), 317—325.

[2] Bailey, N. T. I.; On queueing processes with bulk service, *Journ. Royal Stat. Soc.*, s. B. **16** №1 (1954), 80—87.

[3] Downton, F.: Waiting time in bulk service queues, *Journ. Royal Stat. Soc.* s. B. **17** №2. 256—261.

[4] Кендалл, Д.: Стохастические процессы, встречающиеся в теории очередей, и их анализ методом вложенных цепей Маркова, *Сб. переводов «Математика», 3*; *6,* (1959), 97—111.

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 127—132.

SOME FURTHER CONSTRUCTIONS FOR $G_2(d)$ GRAPHS

by

R. C. BOSE and S. S. SHRIKHANDE

Abstract. A $G_2(d)$ graph is a finite, undirected graph without loops or multiple edges in which each pair of vertices is adjacent to exactly d other vertices, $d \geqq 2$. An infinite family of such graphs was given in [2]. The present paper gives some further constructions for these graphs.

1. Known results. We use the notation and terminology of [2] and quote some results contained in this paper.

THEOREM A. *A $G_2(d)$ graph, $d \geqq 2$ is regular of valence n_1 such that $v-1 = n_1(n_1-1)/d$ where v is the number of vertices and there exists a positive integer m such that*

(i) $n_1 = d+m^2$, *and*
(ii) d/m *is an integer with the same parity as* $v-1-m$.

We note that a $G_2(d)$ graph with parameters (v, n_1, d), $d \geqq 2$ is essentially a strongly regular graph with parameters $(v, n_1, p_{11}^1, p_{11}^2)$ where $p_{11}^1 = p_{11}^2 = d$. By a pseudo $L_r(k)$ graph we will mean a pseudo net graph $L_r(k)$ and by an $NL_r(k)$ graph a negative Latin Square $NL_r(k)$ graph.

THEOREM B. *The existence of a pseudo $L_{r_1}(2r_1)$ and a pseudo $L_{r_2}(2r_2)$ graph implies the existence of a pseudo $L_r(2r)$ graph with $r=2r_1r_2$.*

THEOREM C. *The existence of a pseudo $L_{r_1}(2r_1)$ and a $NL_{r_2}(2r_2)$ graph implies the existence of a $NL_r(2r)$ graph with $r=2r_1r_2$.*

THEOREM D. *Pseudo $L_r(2r)$ and $NL_r(2r)$ graphs exist for $r=3^m2^{m+n-1}$, where m, n are nonnegative integers $(m,n)\neq(0,0)$.*

Noting that a pseudo $L_r(k)$ graph is a strongly regular graph with parameters $(k^2, r(k-1), k-2+(r-1)(r-2), r(r-1))$ and a $NL_r(k)$ graph is strongly regular with parameters $(k^2, r(k+1), -k-2+(r+1)(r+2), r(r+1))$ we have

THEOREM E. *A $G_2(d)$ graph with parameters*

(i) $v = 4r^2$, $n_1 = r(2r-1)$, $d = r(r-1)$,
(ii) $v = 4r^2$, $n_1 = r(2r+1)$, $d = r(r+1)$,
(iii) $v = 4r^2-1$, $n_1 = 2r^2$, $d = r^2$,

exists for all $r=3^m2^{m+n-1}$, where m, n are non negative integers $(m,n)\neq(0,0)$.

2. New constructions. We first prove some preliminary results.

LEMMA 2. 1. *Let v, b, k, r be non negative integers, $k<v$ and $vr=bk$. Let there exist an incomplete block design D with v symbols (treatments) in b subsets (blocks) of size k such that any pair of treatments occurs together in at most one block. Then a necessary and sufficient condition that each block in D intersects precisely $k(r-1)$ other blocks in D is that each treatment occurs exactly r times in D.*

PROOF. It is obvious that any two blocks in D intersect in at most one treatment. If each treatment occurs r times in D, then any block containing, say, treatments $t_1, t_2, \ldots, t_k$ intersects $(r-1)$ other blocks containing t_i; $i=1, 2, \ldots, k$. The sets of $(r-1)$ blocks containing t_i and t_j are obviously disjoint, $i\neq j$. Hence this block intersects precisely $k(r-1)$ other blocks necessarily in one treatment. This proves the sufficiency part of the theorem.

Now suppose that each block intersects precisely $k(r-1)$ other blocks. Let r_i be the number of times the treatment i occurs in D. Then obviously

$$\bar{r} = \sum_1^v \frac{r_i}{v} = r.$$

Let $N=(n_{ij})$ be the usual (0, 1) incidence matrix of D with v rows and b columus where $n_{ij}=1$ or 0 according as treatment i occurs in block j or not. Then from our hypothesis

$$N'N = kI_b + A$$

where A is an adjacency matrix of order b which is regular and of valence $k(r-1)$. Also

$$NN' = \operatorname{diag}(r_1, \ldots, r_v) + B$$

where B is also an adjacency matrix of order v with i-th row sum $r_i(k-1)$. Hence

$$\begin{aligned}\operatorname{tr}\left((NN')(NN')\right) &= \sum_1^v r_i^2 + (k-1)\sum_1^v r_i \\ &= \sum_1^v r_i^2 + (k-1)vr.\end{aligned}$$

But

$$\begin{aligned}\operatorname{tr}\left((NN')(NN')\right) &= \operatorname{tr}\left(N(N'N)N'\right) \\ &= \operatorname{tr} N(kI_b + A)N' \\ &= k \operatorname{tr} NN' + \operatorname{tr} NAN' \\ &= k \operatorname{tr} NN' + \operatorname{tr} AN'N \\ &= k \operatorname{tr} NN' + \operatorname{tr}\left(A(kI_b + A)\right) \\ &= k \sum_1^v r_i + \operatorname{tr} A^2 \\ &= kvr + bk(r-1) \\ &= vr(k+r-1).\end{aligned}$$

Hence equating the two values of the trace

$$\sum_1^v (r_i-\bar r)^2 = \sum_1^v r_i^2 - v\bar r^2$$

$$= \sum_1^v r_i^2 - vr^2$$

$$= 0$$

which implies that each treatment occurs r times in D. This completes the proof of the lemma.

We now define the concept of an ascendant graph G^* of a strongly regular graph G with parameters $(v, n_1, p_{11}^1, p_{11}^2)$. Let (V_1, V_2) be a partition of the vertex set V of G where V_1 and V_2 respectively contain n_1^* and $v-n_1^*$ vertices. Let ∞ be a vertex not in V and let G^* be a graph with vertex set $(\infty \cup V)$. We define adjacency in G^* as follows: The vertex ∞ is adjacent only to vertices of V_1 (and to all vertices of V_1). If x, y are in V, then they are adjacent in G^* if and only if they are adjacent in G and belong both to V_1 or both to V_2, or if they are nonadjacent in G and belong one to V_1 and the other to V_2. If the graph G^* is strongly regular with parameters $(v^*, n_1^*, p_{11}^{1*}, p_{11}^{2*})$ where v^* is necessarily $v+1$, then G^* is said to be an ascendant of G.

We derive the conditions under which a graph G with parameters $(v, n_1, p_{11}^1, p_{11}^2)$ has an ascendant G^* with parameters $(v^*, n_1^*, p_{11}^{1*}, p_{11}^{2*})$. We will assume that G is neither a void graph nor a complete graph i.e. $n_1 \neq 0$ and $n_2 = v-1-n_1 \neq 0$.

If G^* is an ascendant of G, then G is a descendant of G^* with respect to the vertex ∞ and hence from [2]

$$p_{11}^{1*}+p_{11}^{2*} = 2n_1^* - \frac{v^*}{2}$$

$$= 2n_1^* - \frac{v+1}{2}. \tag{2.1}$$

$$n_1 = 2n_1^* - 2p_{11}^{2*}. \tag{2.2}$$

$$p_{11}^1 = n_1 - n_1^* + p_{11}^{1*}. \tag{2.3}$$

$$p_{11}^2 = n_1 - n_1^* + p_{11}^{2*}. \tag{2.4}$$

From (2. 2.), (2. 4)

$$2p_{12}^2 = 2(n_1 - p_{11}^2)$$

$$= 2(n_1^* - p_{11}^{2*})$$

$$= n_1.$$

Hence from the usual parametric relations

$$n_1/2 = p_{11}^2 = p_{12}^2 \tag{2.5}$$

which implies that

$$n_2/2 = p_{12}^1 = p_{22}^1 \tag{2.6}$$

Also from (2. 1), (2. 3) and (2. 4)

$$v+1 = 4n_1 - 2p^1_{11} - 2p^2_{11}$$

or

$$v = 6p^2_{11} - 2p^1_{11} - 1. \tag{2. 7}$$

Thus (2. 7) is a necessary condition for G to have an ascendant. It is easy to see that for a graph G with parameters $(v, n_1, p^1_{11}, p^2_{11})$ each of (2. 5), (2. 6), (2. 7) implies the other two. It also follows from [2] that the vertex set V of G can be partitioned into (V_1, V_2) with n^*_1 vertices in V_1 and $v-n^*_1$ vertices in V_2, where the set V_1 is the set of vertices in G^* which are adjacent to ∞. Further, each vertex in V_1 is adjacent to $p^{1*}_{11} = n^*_1 - n_1 + p^1_{11}$ vertices of V_1 in G and each vertex in V_2 is adjacent to

$$\begin{aligned} p^{2*}_{12} &= n^*_1 - p^{2*}_{11} \\ &= n_1 - p^2_{11} \\ &= p^2_{11} \end{aligned}$$

vertices of V_2 in G.

Conversely, suppose (2. 7) is satisfied for G and further there exists a partition (V_1, V_2) of the vertex set V of G with n^*_1 vertices in V_1 and $v-n^*_1$ vertices in V_2, such that each vertex in V_1 is adjacent to $n^*_1 - n_1 + p^1_{11}$ vertices in V_1 and each vertex in V_2 is adjacent to p^2_{11} vertices in V_2. Then by using arguments similar to those in [2], it can be shown that G^* is a strongly regular graph with parameters (2. 2), (2. 3), (2. 4) provided p^{1*}_{11}, p^{2*}_{11} are non negative integers. From the relation

$$n^*_1 p^{1*}_{12} = n^*_2 p^{2*}_{11}$$

it follows that n^*_1 satisfies the equation

$$f(x) \equiv x^2 + x(p^1_{11} - 5p^2_{11}) + vp^2_{11} = 0.$$

The above equation is easily seen to have real positive roots. Since n^*_1 is necessarily an integer a further necessary condition for G to have an ascendant G^* is that the above equation has an integral solution.

We easily verify that

$$\begin{aligned} f'(n_1 - p^1_{11}) &= -(p^1_{11} + p^2_{11}) < 0. \\ f'(n_1 - p^2_{11}) &= -(1 + p^1_{12} + p^2_{11}) < 0. \\ f(n_1 - p^2_{11}) &= p^2_{11} p^1_{12} > 0. \\ f(n_1 - p^1_{11}) &= p^2_{11}(p^1_{11} - 1). \end{aligned}$$

Further, it is easily seen that if $p^1_{11} = 0$, then $f(x) = 0$ has no integral solution. Hence, if the equation has an integral solution then we can assume that $p^1_{11} \geqq 1$ and then the above relations imply that

$$\begin{aligned} p^{1*}_{11} &= n^*_1 - n_1 + p^1_{11} \geqq 0. \\ p^{2*}_{11} &= n^*_1 - n_1 + p^2_{11} > 0. \end{aligned}$$

Thus the condition that $f(x) = 0$ has an integral solution n^*_1 is necessary and sufficient for non negativeness of p^{1*}_{11}, p^{2*}_{11}. We can, therefore, state the following theorem.

THEOREM 2. 1. *Let G be a strongly regular graph with parameters* $(v, n_1, p_{11}^1, p_{11}^2)$. *Then G has an ascendant* G^* *with parameters* $(v^*, n_1^*, p_{11}^{1*}, p_{11}^{2*})$ *if and only if the following parametric and structural conditions* (P) *and* (S) *are satisfied in G.*

(P) $v = 6p_{11}^2 - 2p_{11}^1 - 1$.

(S) *The equation*

$$x^2 + x(p_{11}^1 - 5p_{11}^2) + vp_{11}^2 = 0$$

has an integral solution n_1^* *and there exists a partition* (V_1, V_2) *of the vertex set V of G with* n_1^* *vertices in* V_1 *and* $v - n_1^*$ *vertices in* V_2 *such that every vertex in* V_1 *has* $n_1^* - n_1 + p_{11}^1$ *adjacent vertices in* V_1 *and every vertex in* V_2 *has* p_{11}^2 *adjacent vertices in* V_2.

The parameters of G^* are then given by

$$v^* = v+1, \quad n_1^*, \quad p_{11}^{1*} = n_1^* - n_1 + p_{11}^1, \quad p_{11}^{2*} = n_1^* - n_1 + p_{11}^2.$$

It is obvious that any two blocks of a BIBD with $\lambda = 1$ have at most one treatment in common. Consider a BIBD with $r = 2k+1$, $\lambda = 1$. Then the values of v and b are given by $v = 2k^2 - k$ and $b = 4k^2 - 1$. Consider the blocks as vertices of a graph G and define two blocks as adjacent or nonadjacent according as they have a treatment in common or not. Then [2] G is strongly regular with parameters $(4k^2 - 1, 2k^2, k^2, k^2)$ and satisfies the condition (P) of the above theorem. Also the equation $f(x) = 0$ has integral solutions $k(2k-1)$ and $k(2k+1)$. Take $n_1^* = k(2k-1)$. If the $4k^2 - 1$ blocks can be partitioned into sets V_1 and V_2 of $k(2k-1)$ and $(k+1)(2k-1)$ blocks respectively such that each block in V_1 is adjacent to $k^2 - k$ blocks in V_1 and each block in V_2 is adjacent to k^2 blocks in V_2, then the condition (S) is also satisfied. From Lemma 2. 1 this means that the set V_1 (respectively V_2) contains each of the $2k^2 - k$ treatments exactly k (respectively $k+1$) times.

We note that a BIBD with $r = 2k+1$, $\lambda = 1$ is a partial geometry $(r, k, t) = (2k+1, k, k)$. The graph G is then [1] the graph of the dual configuration and is also a partial geometry $(k, 2k+1, k)$. We can, therefore, state the following theorem.

THEOREM 2. 2. *Let G be the graph of the dual of a BIBD with* $r = 2k+1$, $\lambda = 1$. *Then G has an ascendant* G^* *which is a pseudo* $L_k(2k)$ *graph if and only if the* $4k^2 - 1$ *blocks of the BIBD can be partitioned into sets* V_1 *and* V_2 *of* $k(2k-1)$ *and* $(k+1)(2k-1)$ *blocks respectively such that each of the* $2k^2 - k$ *treatments of the BIBD occur k times in* V_1 *and* $k+1$ *times in* V_2.

BIBD's having the structure of the above theorem exist for $k = 5$ and 7. See for example Appendix I in [4]. Hence we have the following result.

COROLLARY. *Pseudo* $L_5(10)$ *and pseudo* $L_7(14)$ *graphs exist.*

GOETHALS and SEIDEL [3] have constructed a pseudo $L_5(10)$ graph in precisely the same manner.

Using the Corollary and Theorems B, C and D we have the following theorem

THEOREM 2. 3. (I) *pseudo* $L_r(2r)$ *graphs exist for all* $r = 3^m 5^a 7^c 2^{m+a+c+n-1}$ *where m, n, a, c are non negative integers* $(m, n, a, c) \neq (0, 0, 0, 0)$.

(II) *$NL_r(2r)$ graphs exist for $r=5^a 7^c 2^{a+c}$ where, a, c are non negative integers and for $r=3^m 5^a 7^c 2^{m+a+c+n-1}$ where m, n, a, c are non negative integers and $(m, n) \neq (0, 0)$.*

The proof is similar to that of Theorem 9. 3 and 9. 5 in [2] and is omitted.

Finally, noting that pseudo $L_r(2r)$ and $NL_r(2r)$ graphs are $G_2(d)$ graphs we have

THEOREM 2. 4. *$G_2(d)$ graphs with the following parameters exist*

(i) $v = 4r^2$, $n_1 = r(2r-1)$, $d = r(r-1)$;

(ii) $v = 4r^2-1$, $n_1 = 2r^2$, $d = r^2$;

for all $r=3^m 5^a 7^c 2^{m+a+c+n-1}$ where m, n, a, c are non negative integers $(m, n, a, c) \neq (0, 0, 0, 0)$.

(iii) $v=4r^2$, $n_1 = r(2r+1)$, $d = r(r+1)$;

with $r=5^a 7^c 2^{a+c}$, where a, c are non negative integers and with $r=3^m 5^a 7^c 2^{m+a+c+n-1}$ where m, n, a, c are non negative integers and $(m, n) \neq (0, 0)$.

We remark that since our construction is essentially by a composition method, any new $G_2(d)$ graph with parameters as in Theorem E can be utilised in conjunction with the above theorem to enlarge such a family considerably.

REFERENCES

[1] BOSE, R. C.: Strongly regular graphs, partial geometries and partially balanced designs, *Pacific J. Math.* 13 (1963), pp. 389—418.

[2] BOSE, R. C. and SHRIKHANDE, S. S.: Graphs in which each pair of vertices is adjacent to the same number d of other vertices, *Studia Sci. Math. Hungar.* **5** (1970), 181—196.

[3] GOETHALS, J. M. and SEIDEL, J. J.: Strongly regular graphs derived from combinatorial designs, *Canad. J. Math.* **22** (1970), 449—471

[4] HALL MARSHALL: *Combinatorial Theory*, Blaisdell (1967).

Colorado State University, Fort Collins, CO, USA. University of Bombay, Bombay, India

(Received October 1, 1969.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 133—135.

ÜBER PARKETTIERUNGEN KONSTANTER NACHBARNZAHL

von

G. FEJES TÓTH

Eine Menge konvexer Polygone, die die Ebene schlicht und lückenlos überdecken wird eine *Parkettierung* genannt. Zwei Polygone, die gemeinsame Randpunkte haben, nennen wir *Nachbarn*. Hat in einer Parkettierung jedes Polygon dieselbe Anzahl n von Nachbarn, so sprechen wir von einer *n-Nachbarnparkettierung*.

Für $n=6$, 7, 8, 9, 10, 12, 14, 16 und 21 sind einfache Beispiele für n-Nachbarnparkettierungen mit kongruenten Flächen bekannt [1]. Andererseits lässt sich zu jedem Wert von $n>5$ eine n-Nachbarnparkettierung konstruieren, indem man von einem n-Eck ausgehend zu den freien Seiten immer neue n-Ecke hinzufügt, und zwar so, dass sich die n-Ecke immer entlang ganzer Seiten aneinanderschliessen [2]. Die so entstehenden Parkettierungen sind aber im Sinne der hierfolgenden Definition keine Normalparkettierungen. Eine Parkettierung wird *normal* genannt [1], wenn die Inkreisradien der Flächen eine positive untere, und die Umkreisradien eine endliche obere Schranke haben.

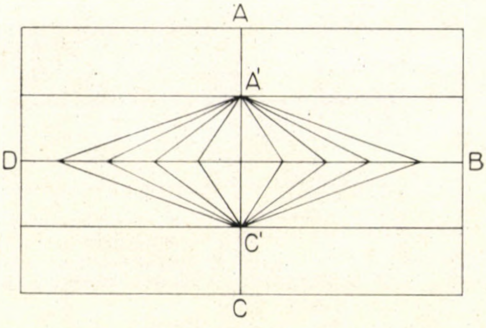


Abb. 1

Es erhebt sich die Frage ob es für jeden Wert von $n>5$ eine normale n-Nachbarnparkettierung existiert. L. FEJES TÓTH [1] sprach die Vermutung aus, dass die Antwort bejahend ist. Bisher war aber ausser den oben genannten Fällen von Parkettierungen mit kongruenten Flächen nur ein Beispiel, nämlich eine normale 11-Nachbarnparkettierung bekannt. In diesem Aufsatz beweisen wir folgenden Satz, der die obige Vermutung bestätigt.

Es existiert zu jedem ganzzahligen Wert von $n>5$ eine normale n-Nachbarnparkettierung.

Es sei R ein Rechteck. Wir bezeichnen die Seitenmittelpunkte von R in ihrer zyklischen Reihenfolge mit A, B, C und D. Wir betrachten auf der Srecke AC zwei Punkte A' und C' so, dass $AA'=C'C=AC/4$ sei. Ferner zerlegen wir die Strecke BD durch $n-5$ Punkte in $n-4$ kongruente Teilstrecken. Wir verbinden diese Punkte mit A' und C', und zeichnen noch die durch A' und C' hindurchgehende, zu BD parallele Sehnen von R. Die so erhaltenen Strecken zerlegen R, zusammen mit den Srecken AA', $C'C$ und BD in 4 Rechtecke, 4 Trapezen und 2 $(n-6)$ Dreiecke (Abb. 1).

Wir sagen, dass diese Vielecke einen *n-Block* bilden. Innerhalb von einem n-Block haben die Rechtecke $n-3$, die Trapezen $n-1$ und die Dreiecke n Nachbarn. Wir nennen AC die Achse des Blocks.

Wir verschieben den obigen Block in der Richtung AC um $\overline{AC}, \ldots, (m-1)\overline{AC}$. Die verschobenen Blöcke bilden, zusammen mit dem ursprünglichen ein *m-faches n-Block*. Bemerken wir, dass innerhalb von einem mehrfachen n-Block die „inneren" Rechtecke $n-1$ Nachbarn haben.

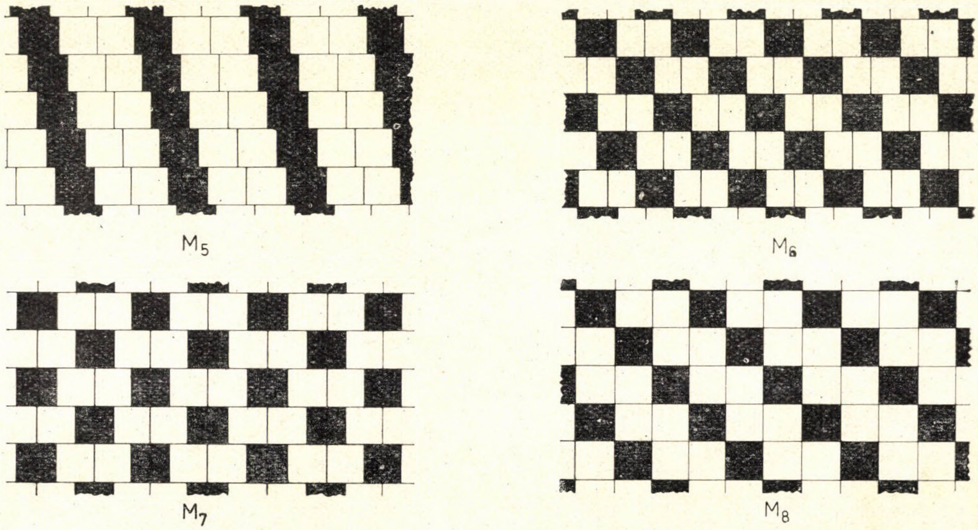


Abb. 2

Wir konstruieren jetzt für $n > 8$ die in unserem Satz angedeuteten Parkettierungen. Wir unterscheiden vier Fälle, je nachdem ob n durch 4 dividiert 0, 1, 2 oder 3 als Rest ergibt.

Wir betrachten die in Abb. 2 dargestellten, aus schwarzen Quadraten und weissen Quadraten bzw. Rechtecken bestehende Mosaike, die wir mit M_5, M_6, M_7 und M_8 bezeichnen. Die schwarzen Quadrate bedeuten k-fache n-Blöcke mit vertikalen Achsen, je nachdem ob $n = 4k+5$, $4k+6$, $4k+7$ oder $4k+8$ ist.

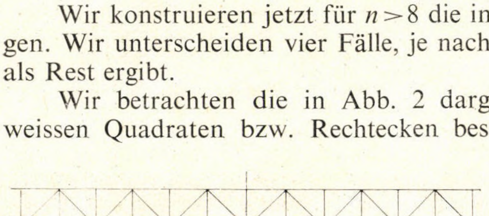


Abb. 3

In M_5 berühren sich die schwarzen Quadrate entlang grösseren Strecken als eine halbe Seite. In M_6 ist die kürzere Seite eines weissen Rechtecks grösser als die halbe Seite eines schwarzen Quadrats. In M_7 liegen die Ecken der Quadrate in je einem Seitenmittelpunkt eines anderen Quadrats. M_8 bedarf keiner Erklärung.

Wir erinnern an die Tatsache, dass innerhalb von einem k-fachen n-Block die Dreiecke n, die Trapezen und die inneren Rechtecke $n-1$ und die übrigen vier Rechtecke $n-3$ Nachbarn haben. Man überzeugt sich leicht, dass in M_j ($j=5, 6, 7, 8$) diese Vielecke noch 0, 1 bzw. 3 weitere Nachbarn bekommen. Jedes weisse Viereck bekommt von dem entlang einer vertikalen Seite anstossenden k-fachen n-Block $4k$, und von den übrigen Polygonen genau j Nachbarn. Folglich hat in den konstruierten Parkettierungen jede Fläche genau n Nachbarn. Da in diesen Parkettierungen nur eine endliche Anzahl von verschiedenen, d.h. nicht kongruenten Polygonen vorkommt, sind sie normal. Damit ist der Satz bewiesen.

Es sei $K=K(n)$ die kleinste ganze Zahl mit der Eigenschaft, dass sich aus kongruenten Exemplaren von K konvexen Polygonen eine n-Nachbarnparkettierung konstruieren lässt. Wir haben erwähnt, dass $K(6)=K(7)=K(8)=K(9)=K(10)=$ $=K(12)=K(14)=K(16)=K(21)=1$ gilt. Die in Abb. 3 dargestellte 11-Nachbarnparkettierung zeigt, dass $K(11)\leqq 2$ ist, und wir haben vermutlich $K(11)=2$. Man zählt leicht nach, dass in unseren obigen Beispielen $\left[\frac{n+1}{2}\right]$ Sorten von Polygonen auftreten, so dass $K(n)\leqq\frac{n+1}{2}$ gilt. Es lässt sich vermuten, dass $\lim_{n\to\infty} K(n)=\infty$ ist. Dies steht aber noch nicht fest.

LITERATURVERZEICHNIS

[1] Fejes Tóth, L.: Scheibenpackungen konstanter Nachbarnzahl, *Acta Math. Acad. Sci. Hung.* **20** (1969) 375—381.

[2] Steinhaus, H.: *Kaleidoskop der Mathematik* (Berlin, 1959).

Ungarische Akademie der Wissenschaften, Budapest

(Eingegangen: 27. October, 1969.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 137—145.

A GENERALIZATION OF BELLMAN'S INEQUALITY FOR STIELTJES INTEGRALS AND A UNIQUENESS THEOREM

by

I. GYŐRI

1. Introduction

In the study of uniqueness and stability of solutions of differential equations some integral inequalities play a very important role. An inequality referred often to of this kind was published by BELLMAN [1] in 1943; several extensions and generalizations of this inequality have been established by a number of authors (see e.g. BIHARI [2], VISWANATHAM [5], BRAUER [3]).

In this paper we discuss a non-linear inequality of HINTON's type involving Lebesgue—Stieltjes integrals, and with the aid of this shall obtain a uniqueness theorem for the solution of the equation

$$y' = f(t, y), \quad y(t_0) = y_0,$$

as general as implying the theorems of NAGUMO, OSGOOD [2], YANG EN-HO [6].

2. Generalization of Bellman's inequality

THEOREM 2. 1. *Assume*

(a) $u(t)$ *is a non-negative and bounded function in the interval* $a \leqq t \leqq b$;
(b) $v(t)$ *is increasing and continuous in* (a, b);
(c) $g(t, u) \geqq 0$ *is continuous for*

$$a \leqq t \leqq b, \qquad 0 \leqq u \leqq k + \sup_{a \leqq x \leqq b} u(x) = u_0,$$

where k *is an arbitrary positive constant. For fixed* t, $g(t, u)$ *is a non-decreasing function of* u, *and, moreover* $g(t, u(t))$ *is integrable.*

Then the inequality

$$u(t) \leqq k + \int_a^t g(s, u(s))\, dv(s) \qquad (a \leqq t \leqq b) \tag{2. 1}$$

implies for some $a < b' \leqq b$ *that,*

$$u(t) \leqq z(t) \qquad (a \leqq t \leqq b') \tag{2. 2}$$

where $z(t)$ *is the maximal solution of the equation*

$$z(t) = k + \int_a^t g(s, z(s))\, dv(s) \tag{2. 3}$$

on the interval $a \leqq t \leqq b'$, *where this solution exists.*[1]

PROOF. It is enough to show that (2. 3) has a solution satisfying (2. 2), and this can be done by successive approximations as follows. Put

$$z_0(t) = u(t)$$

$$z_{n+1}(t) = k + \int_a^t g(s, z_n(s))\, dv(s) \qquad (n = 0, 1, 2, \dots); \tag{2. 4}$$

the functions $z_n(t)$ $(n=1, 2, \dots)$ are obviously continuous. Denote by $M>0$ the maximum of $g(t, u)$ while $0 \leqq u \leqq u_0$, $a \leqq t \leqq b$. We are going to show that for a suitable $a<b' \leqq b$ we have

$$0 \leqq z_n(t) \leqq u_0 \tag{2. 5}$$

for all positive integer n, whenever $a \leqq t \leqq b'$.

In fact, assume that $z_n(t)$ $(n \geqq 1$; the case $n=0$ is obvious) already satisfies (2. 5).

Then by (2. 4) we infer

$$0 \leqq z_{n+1}(t) \leqq k + M[v(t) - v(a)] \leqq u_0,$$

whenever $t \geqq a$ is as small as satisfying the inequality

$$v(t) \leqq \frac{u_0 - k}{M} + v(a) = \frac{\sup\limits_{a \leqq x \leqq b} u(x)}{M} + v(a)$$

The continuity of $v(t)$ implies the existence of such a t. This shows that the sequence $\{z_n(t)\}$ is bounded in the interval $a \leqq t \leqq b'$, for some $b'>a$; moreover, for every t and n the point $(t, z_n(t))$ belongs to the domain of $g(t, u)$. By induction we easily obtain that $z_n(t) \leqq z_{n+1}(t)$ for every $a \leqq t \leqq b'$, $n \geqq 0$. Namely, if $z_{n-1}(t) \leqq z_n(t)$ $(n \geqq 1$, the case $n=0$ being obvious), then by the monotonity of $g(t, u)$ and $v(t)$ we have

$$z_n(t) = k + \int_a^t g(s, z_{n-1}(s))\, dv(s) \leqq k + \int_a^t g(s, z_n(s))\, dv(s) = z_{n+1}(t)$$
$$(a \leqq t \leqq b').$$

The definition of the functions $\{z_n(t)\}$ implies that for any $\varepsilon>0$ there exists a $\delta>0$ such that for every $a \leqq t_1, t_2 \leqq b'$ and $n \geqq 1$ the inequality

$$|z_n(t_2) - z_n(t_1)| \leqq M|v(t_2) - v(t_1)| < \varepsilon$$

[1] The notion of maximal solution is analogous to that in the case of differential equations, and also its existence can be proved in the same way.

holds, provided $|t_2-t_1| < \delta$. Therefore ARZELA's well-known theorem implies that this sequence of functions converges uniformly. Put

$$\bar{z}(t) = \lim_{n\to\infty} z_n(t) \qquad (a \leqq t \leqq b').$$

It is easily seen that this function is a solution of equation (2. 3). Inequality (2. 2) is also satisfied, because we have

$$u(t) = z_0(t) \leqq z_1(t) \leqq \ldots \leqq \bar{z}(t) \leqq z(t).$$

Remark. By the substitution $v(s) = s$ we obtain the generalization given by VISWANATHAM [6] of BELLMAN's lemma.

As a consequence of the above theorem we now derive

THEOREM 2. 2. *Assume*

(a) *$u(t)$ is a non-negative and bounded function on the interval $[a, b]$;*

(b) *$v(t)$ is a continuous non-decreasing function on $[a, b]$;*

(c) *$g(u)$ is a continuous non-decreasing function on $0 \leqq u < \infty$ and there exists an $u_0 > 0$ such that $g(u) > 0$ for $u \geqq u_0$, and moreover $g(u(t))$ is integrable.*

Then, for an arbitrary positive constant k the inequality

$$u(t) \leqq k + \int_a^t g(u(s))\, dv(s) \tag{2. 6}$$

implies

$$u(t) \leqq G^{-1}(G(k) + v(t) - v(a)) \qquad (a \leqq t \leqq b' \leqq b), \tag{2. 7}$$

where $b' > a$ and

$$G(u) = \int_{u_0}^{u} \frac{1}{g(t)}\, dt \qquad (u \geqq 0), \tag{2. 8}$$

and G^{-1} is the inverse of G; moreover b' is chosen so that $G(k) + v(t) - v(a)$ belongs to the domain of G^{-1} whenever $a \leqq t \leqq b'$.

PROOF. In view of the theorem just proved we have

$$u(t) \leqq z(t)$$

where $z(t)$ is the maximal solution of the equation

$$z(t) = k + \int_a^t g(z(s))\, dv(s) \tag{2. 9}$$

Thus it is enough to show that any solution of equation (2. 9) satisfies the inequality

$$z(t) < G^{-1}(G(k+\varepsilon) + v(t) - v(a)) = y(t) \tag{2. 10}$$

This, however, will obviously follow if we show that the right hand side of (2. 10) is a solution of the equation

$$(2.11)\qquad y(t) = k+\varepsilon+\int_a^t g(y(s))\,dv(s),$$ [2]

i.e.

$$(2.12)\qquad G^{-1}(G(k+\varepsilon)+v(t)-v(a)) = k+\varepsilon+\int_a^t g(G^{-1}(G(k+\varepsilon)+v(\tau)-v(a)))\,dv(\tau).$$

Here for the right-hand side we obtain

$$\int_a^t g(G^{-1}(G(k+\varepsilon)+v(\tau)-v(a)))\,dv(\tau) = \int_{v(a)}^{v(t)} g(G^{-1}(G(k+\varepsilon)+s-v(a)))\,ds,$$

i.e., taking into account the definition of G, or rather its direct consequence, the equality

$$\frac{dG^{-1}(t)}{dt} = g(G^{-1}(t)),$$

we have

$$\int_a^t g(G^{-1}(G(k+\varepsilon)+v(\tau)-v(a)))\,dv(\tau) =$$

$$= G^{-1}(G(k+\varepsilon)+v(t)-v(a))-G^{-1}(G(k+\varepsilon)+v(a)-v(a)) =$$

$$= -k-\varepsilon+G^{-1}(G(k+\varepsilon)+v(t)-v(a)).$$

This proves (2. 12) and therewith the theorem.

Remark. If $v_1(t)\geqq 0$ is continuous and

$$v(t) = \int_a^t v_1(\tau)\,d\tau$$

then our theorem yields the assertion of BIHARI's lemma (see [2]).

3. A uniqueness theorem

In the last section of our paper we apply our Theorem 2. 2 to the derivation of a uniqueness theorem of a general type, and, as a conclusion of these notes, we list some of its various implications.

THEOREM 3. 1. *Assume*

(a) $f(t, y)$ *is a function continuous in the domain*

$$G=\{(t, y)\colon\ t_0\leqq t\leqq t_0+a,\ |y-y_0|\leqq b\}$$

for some positive a and b;

[2] An easy argument, similar to classical ones in the theory of differential equations, shows that each solution of (2. 9) is dominated by any solution of (2. 11).

(b) *for every sufficiently small* δ, $0<\delta<a$, *the function* $v(t;\delta)$ *is continuous and non-decreasing in* t, *provided* $t_0+\delta \leqq t \leqq t_0+a$

(c) $g(u)$ *is continuous and non-decreasing if* $0 \leqq u < \infty$;

(d) *There is no positive constant* k *such that the inequality*

$$\int_{\varepsilon+\delta\psi(\delta)}^{\varepsilon+k} \frac{1}{g(\tau)}\, d\tau \leqq v(t_0+a;\delta)-v(t_0+\delta;\delta), \tag{3. 1}$$

holds for any ε, $\delta>0$ *and for any function* $\psi(\delta)$ *tending to* 0 *if* $\delta\to 0$.

If, apart from the above conditions, $f(t, y)$ *satisfies the inequality*

$$\begin{gathered} |f(t,y_2)-f(t,y_1)| \leqq v'(t;\delta)g(|y_2-y_1|)\,{}^{3} \\ (t_0+\delta \leqq t \leqq t_0+a;\ y_0-b \leqq y_1,\ y_2 \leqq y_0+b) \end{gathered} \tag{3. 2}$$

then, for some positive η, *the solution passing through the point* (t_0, y_0) *of the differential equation*

$$y'(t)=f(t,y) \tag{3. 3}$$

is uniquely determined in the interval $(t_0, t_0+\eta)$.

PROOF. Suppose that $y(t)$ and $z(t)$ are two solutions of equation (3. 3) such that $y(t_0)=z(t_0)=y_0$. Then in a sufficiently small neighbourhood of (t_0, y_0) we have

$$z(t)-y(t)=\int_{t_0}^{t} [f(\tau, z(\tau))-f(\tau, y(\tau))]\, d\tau. \tag{3. 4}$$

Let $t>t_0$ and fix $\delta>0$ so that $t_0+\delta<t$. Now by (3. 2) we have

$$\int_{t_0+\delta}^{t} |f(\tau,z(\tau))-f(\tau,y(\tau))|\, d\tau \leqq \int_{t_0+\delta}^{t} g(|z(\tau)-y(\tau)|)v'(\tau;\delta)\, d\tau \leqq$$

$$\leqq \int_{t_0+\delta}^{t} g(|z(\tau)-y(\tau)|)\, dv(\tau;\delta);\ {}^{4}$$

therefore (3. 4) implies

$$|z(t)-y(t)| \leqq \int_{t_0}^{t} |f(\tau,z(\tau))-f(\tau,y(\tau))|\, d\tau \leqq \delta \max_{t_0 \leqq \tau \leqq t_0+\delta} |f(\tau,z(\tau))-f(\tau,y(\tau))| +$$

$$+\int_{t_0+\delta}^{t} g(|z(\tau)-y(\tau)|)\, dv(\tau;\delta).$$

[3] Namely, $v'(t;\delta)$ may possibly not exist for every t.

[4] This last inequality follows since by omitting the singular part of $v(t;\delta)$ in its canonical decomposition as a continuous and monotonic function we have equality.

Choosing the function $\psi(\delta)$ featuring in (3. 1) as

$$\max_{t_0\leqq\tau\leqq t_0+\delta} |f(\tau, z(\tau))-f(\tau, y(\tau))|, \tag{3. 5}$$

we obtain

$$|z(t)-y(t)| \leqq \delta\psi(\delta)+\int_{t_0+\delta}^{t} g(|z(\tau)-y(\tau)|)\,dv(\tau;\delta) \qquad (t_0+\delta\leqq t).$$

If $\varepsilon>0$ is arbitrary but fixed, the inequality

$$|z(t)-y(t)|+\varepsilon \leqq \delta\psi(\delta)+\varepsilon+\int_{t_0+\delta}^{t} g(|z(\tau)-y(\tau)|+\varepsilon)\,dv(\tau;\delta)$$

holds, to which we may apply Theorem 2. 2 to obtain

$$G(|z(t)-y(t)|+\varepsilon) \leqq G(\delta\psi(\delta)+\varepsilon)+v(t;\delta)-v(t_0+\delta;\delta),$$

where $G(u)$ is the same as in (2. 8):

$$G(u)=\int_{u_0}^{u}\frac{1}{g(t)}\,dt \qquad (u_0>0,\ u\geqq 0).$$

By the definition of G here we obtain

$$\int_{\varepsilon+\delta\psi(\delta)}^{|z(t)-y(t)|+\varepsilon}\frac{1}{g(\tau)}\,d\tau \leqq v(t;\delta)-v(t_0+\delta;\delta)\leqq v(t_0+a;\delta)-v(t_0+\delta;\delta) \tag{3. 8}$$

$$(t_0+\delta\leqq t)$$

for every ε, $\delta>0$ which, by virtue of condition (d), implies

$$|z(t)-y(t)|=0.$$

$$(t_0\leqq t)$$

Since $t>t_0$ was arbitrary the assertion of the theorem follows.

Remarks 1. (NAGUMO's uniqueness theorem.) Assume $g(u)=u$ and

$$v(t;\delta)-v(t_0+\delta;\delta)=\int_{t_0+\delta}^{t}\frac{1}{\tau-t_0}\,d\tau \qquad (t_0+\delta<t).$$

Then $g(u)$ and $v(t,\delta)$ satisfy the requirements of our theorem, and so the inequality

$$|f(t,y_2)-f(t,y_1)|\leqq\frac{1}{t-t_0}|y_2-y_1|$$

implies the uniqueness of the solution of (3. 3).

2. (OSGOOD's theorem.) Put

$$v(t;\delta)-v(t_0+\delta;\delta)=t-t_0-\delta$$

and assume $g(0)=0$ and

$$\lim_{0<u_0\to 0}\int_{u_0}^{u}\frac{1}{g(t)}\,dt=\infty \qquad (u\neq 0).$$

Then, according to our theorem, the inequality

$$|f(t, y_2)-f(t, y_1)| \leqq g(|y_2-y_1|)$$

implies uniqueness.

3. It is not necessary to require

$$\lim_{0<u_0\to 0} \int_{u_0}^{u} \frac{1}{g(t)}\, dt = \infty$$

as we did in the previous remark.

In fact, if $t > t_0+\delta\,(\delta > 0)$ and

$$v(t;\delta)-v(t_0+\delta;\delta) = (t-t_0)^\delta-\delta^\delta$$

holds, and moreover $g(u)=\sqrt{u}$, then $g(u)$ and $v(t;\delta)$ satisfy the requirements of our Theorem 2. 2; namely the inequality

$$a^\delta-\delta^\delta \geqq \int_{\varepsilon+\delta\psi(\delta)}^{k+\varepsilon} \frac{1}{\sqrt{u}}\, du = 2\sqrt{\varepsilon+k}-2\sqrt{\varepsilon+\delta\psi(\delta)}$$

for sufficiently small ε, $\delta>0$ implies $k=0$. Thus

$$|f(t, y_2)-f(t, y_1)| \leqq \delta(t-t_0)^{\delta-1}\sqrt{|y_2-y_1|}$$

$$(\delta \leqq t-t_0)$$

also implies uniqueness.

4. (Theorem of YANG EN-HO [7].) Put $g(u)=u^\beta$ and

$$v(t;\delta)-v(t_0+\delta;\delta) = L\int_{t_0+\delta}^{t} \frac{1}{(\tau-t_0)^\alpha}\, d\tau, \qquad (t_0+\delta \leqq t \leqq t_0+a)$$

where $\beta>1$, $0\leqq\alpha\leqq\beta$, $\alpha\neq 1$ and $L>0$ are arbitrary constants. Then conditions (b)—(d) of Theorem 3. 1 are satisfied. This is obvious for (b) and (c). The satisfaction of conditions (d) can be verified as follows.

For a start assume the contrary, i.e. that for some positive constant k the inequality

$$\int_{\varepsilon+\delta\psi(\delta)}^{\varepsilon+k} \frac{1}{g(\tau)}\, d\tau \leqq |v(t_0+a;\delta)-v(t_0+\delta;\delta)| \qquad (\delta \leqq t-t_0)$$

holds for every ε, $\delta>0$. This implies, provided $t > t_0+\delta$, by the definition of $g(u)$ and $v(t;\delta)$ that

$$\frac{1}{(1-\beta)(\varepsilon+k)^{\beta-1}}-\frac{1}{(1-\beta)(\varepsilon+\delta\psi(\delta))^{\beta-1}} \leqq \frac{L}{1-\alpha}\left[\frac{1}{a^{\alpha-1}}-\frac{1}{\delta^{\alpha-1}}\right],$$

i.e.

$$\frac{1}{(1-\beta)(\varepsilon+k)^{\beta-1}}-\frac{L}{(1-\alpha)a^{\alpha-1}} \leqq \frac{1}{(1-\beta)(\varepsilon+\delta\psi(\delta))^{\beta-1}}-\frac{L}{(1-\alpha)\delta^{\alpha-1}}$$

for all $\delta, \varepsilon > 0$. Put $\varepsilon = \delta - \delta\psi(\delta)$; then the above inequality implies

$$(3.9) \qquad \frac{1}{(1-\beta)(\delta-\delta\psi(\delta)+k)^{\beta-1}} - \frac{L}{(1-\alpha)a^{\alpha-1}} \leqq \frac{1}{(1-\beta)\delta^{\beta-1}} - \frac{L}{(1-\alpha)\delta^{\alpha-1}}$$

By the assumptions for the constants involved we can infer

$$\lim_{\delta\to 0}\left[\frac{1}{(1-\beta)\delta^{\beta-1}} - \frac{L}{(1-\alpha)\delta^{\alpha-1}}\right] = \lim_{\delta\to 0}\frac{(1-\alpha)-L(1-\beta)\delta^{\beta-\alpha}}{(1-\alpha)(1-\beta)\delta^{\beta-1}} = -\infty,$$

thus, if $\delta \to 0$, (3. 9) implies

$$\frac{1}{(1-\beta)k^{\beta-1}} - \frac{L}{(1-\alpha)a^{\alpha-1}} \leqq -\infty,$$

and this is an obvious contradiction. The case $t < t_0$ can be handled similarly.

Since, as we have shown, conditions (b)—(d) of our theorem are satisfied, and

$$v'(t;\delta) = \frac{L}{(t-t_0)^\alpha},$$

the continuity (as required in condition (a)) of $f(t;y)$ in the domain $t_0 \leqq t \leqq t_0 + a$ $|y-y_0| \leqq b$, and the inequality

$$|f(t,y_2)-f(t,y_1)| \leqq \frac{L}{(t-t_0)^\alpha}|y_2-y_1|^\beta,$$

where, we recall, $1 < \beta$, $0 \leqq \alpha \leqq \beta$, $\alpha \neq 1$ and $L > 0$, imply the uniqueness of the solution of the equation (3. 3) at the point (t_0, y_0).

5. Define

$$v(t;\delta) = \int_{t_0+\delta}^{t} \frac{1}{g(\tau-t_0)}\,d\tau, \qquad (t_0+\delta \leqq t \leqq t_0+a)$$

where $g(u)$ is the function defined in Theorem 3. 1. Assume that condition (d) of Theorem 3. 1 is satisfied, i.e. that the inequality

$$(3.10) \qquad \int_{\varepsilon+\delta\psi(\delta)}^{\varepsilon+k} \frac{1}{g(t)}\,dt \leqq \int_{\delta}^{a} \frac{1}{g(t)}\,dt$$

does not hold with any positive k for every $\varepsilon, \delta > 0$.

Then, by Theorem 3. 1 if $f(t, y)$ is continuous in a neighbourhood of (t_0, y_0) and satisfies the inequality

$$(3.11) \qquad |f(t,y_2)-f(t,y_1)| \leqq \frac{1}{g(t-t_0)}\,g(|y_2-y_1|), \qquad (t_0 \leqq t \leqq t_0+a)$$

then at most one solution of the equation (3. 3) passes through the point (t_0, y_0).

If in particular $g(u) = e^u - 1$, then (3. 10) cannot be satisfied and (3. 11) becomes

$$|f(t,y_2)-f(t,y_1)| \leqq \frac{1}{e^{t-t_0}-1}\left(e^{|y_2-y_1|}-1\right) \qquad (t_0 \leqq t).$$

REFERENCES

[1] BELLMAN, R.: The stability of solutions of linear differential equations. *Duke Math. Journal*, **10** (1943), 643—647.

[2] BIHARI, I.: A generalization of a lemma of Bellman and its application to uniqueness problems of differential equations. *Acta Math. Acad. Sci. Hung.* **7** (1956), 81—94.

[3] BRAUER, F.: Bounds for solutions of ordinary differential equations. *Proc. Am. Math. Soc.* **14**, (1963).

[4] HINTON, D. B.: A Stieltjes—Volterra integral equations theory. *Canadian Journal of Math.* **18**, (1966), 314—331.

[5] WISWANATHAM, B.: A generalization of Bellman's lemma, *Proc. Am. Math. Soc.*, **14**, (1963), 15—18.

[6] EN-HO YANG: Note on the uniqueness of solutions of differential equations (Chinese) *Shuxue Jinzham* **8**, (1965) 183—186. See Mathematical Reviews (1969) Vol. **37**, 3079.

(Received October 24, 1969.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 147—150.

DURCHLEUCHTUNG GITTERFÖRMIGER KUGELPACKUNGEN MIT LICHTBÜNDELN

von

I. HORTOBÁGYI

Herrn Professor J. MOLNÁR zu seinem 50. Geburstag gewidmet

Eine Menge kongruenter nicht übereinandergreifender Kugeln wird eine Kugelpackung genannt. Wir sagen, dass die Packung in einer Richtung undurchsichtig ist, wenn jede zu der gegebenen Richtung parallele Gerade mindestens eine Kugel trifft. Liegt die Packung zwischen zwei parallelen Ebenen und ist sie in der zu diesen Ebenen senkrechte Richtung undurchsichtig, so sprechen wir von einer Kugelwolke. Den Abstand der begrenzenden Ebenen nennen wir die Breite der Kugelwolke.

FEJES TÓTH [1] hat bewiesen, dass die Breite einer Einheitskugelwolke stets $\geqq 2+\sqrt{2}$ ist, und Gleichheit nur dann gilt, wenn die Wolke aus zwei quadratischen Kugelschichten besteht, so dass jede Kugel vier Kugeln seiner eigenen Schicht und vier Kugeln der anderen Schicht berührt.

Er hat das analoge Problem für Dunkelwolken aufgeworfen, d.h. für Wolken die in jeder solchen Richtung undurchsichtig sind, die zu den begrenzenden Ebenen nicht parallel ist. Da aber die Bestimmung der minimalen Breite einer Dunkelwolke hoffnungslos schwierig zu sein scheint, müssen wir uns hier mit Abschätzungen begnügen. Eine bessere untere Schranke als $2+\sqrt{2}$ ist auch bei Dunkelwolken nicht bekannt. Eine obere Schranke hat BÖRÖCZKY [2] angegeben, der aus 7 Schichten eine verhältnismässig dünne Dunkelwolke konstruiert hat.

Die Schwierigkeit dieser Frage zeigt folgender Satz von HEPPES [3]: Eine gitterförmige Kugelpackung ist stets in drei linear unabhängigen Richtungen durchsichtig (nicht undurchsichtig). Der folgende Satz beantwortet eine diesbezüglich eine Frage von HEPPES.

SATZ. *Eine gitterförmige Packung von Einheitskugeln lässt sich stets in drei linear unabhängigen Richtungen mit einem zylinderförmigen Lichtbündel vom Radius* $\frac{3\sqrt{2}}{4}-1 \approx 0{,}0606$ *durchleuchten.*

Diese Konstante lässt sich nicht durch eine grössere ersetzen, wie es das Beispiel des dichtesten Kugelgitters zeigt. Vor dem Beweis des Satzes machen wir einige Bemerkungen:

1. Die ortogonale Projektion eines regulären Tetraeders mit der Kantenlänge 2 auf eine zu einer Kante des Tetraeders senkrechte Ebene ist ein Dreieck mit den Seiten $\sqrt{3}$, $\sqrt{3}$, 2. Der Umkreisradius dieses Dreiecks ist $\frac{3\sqrt{2}}{4}$.

2. Gilt für die Seiten a, b, c eines Dreiecks $a \geqq \sqrt{3}$, $b \geqq \sqrt{3}$ und $c \geqq 2$, dann ist der Umkreisradius des Dreiecks $\geqq \frac{3\sqrt{2}}{4}$.

3. Gelten für die Seiten a, b, c eines Dreiecks die Ungleichungen $\sqrt{3} \leqq a < 2$, $\sqrt{3} \leqq b < 2$ und $\sqrt{3} \leqq c < 2$, dann ist das Dreieck spitzwinklig.

4. Das Spiegelbild einer Ecke eines spitzwinkligen Dreiecks bezüglich des Mittelpunktes der gegenüberliegenden Seite liegt ausserhalb des Umkreises des Dreiecks.

Wir wenden uns jetzt dem Beweis unseres Satzes zu. In einer gitterförmigen Kugelpackung seien $\overrightarrow{OB}$ und $\overrightarrow{OC}$ zwei nicht parallele Gittervektoren, und $\overrightarrow{OA}$ ein von $\overrightarrow{OB}$ und $\overrightarrow{OC}$ linear unabhängiger kürzerster Gittervektor. Wir werden zeigen, dass sich die Packung in der Richtung $\overrightarrow{OA}$ mit einem zylinderförmigen Lichtbündel vom Radius $\varepsilon = \frac{3\sqrt{2}}{4} - 1$ durchleuchten lässt.

Wir projizieren die Packung auf die zu $\overrightarrow{OA}$ senkrechte Ebene π. Die Projektion der Kugelmittelpunkte bildet ein Punktgitter, und die Projektion der Kugeln besteht aus Einheitskreisen um diese Punkte. Wir werden zeigen, dass diese Kreise die Ebene π nicht vollständig überdecken.

Wir bezeichnen die Projektionen von A, B, C mit A', B', C', den Umkreis des Dreiecks $A'B'C'$ mit k und den durch k hindurchgehenden geraden Kreiszylinder mit H (Abb. 1.). Wir können voraussetzen, dass das Dreieck $A'B'C'$ spitzwinklig ist. Wir werden zeigen, dass für den Radius R von k $R \geqq \frac{3\sqrt{2}}{4}$ gilt.

In [3] beweist HEPPES, dass jede Seite von $A'B'C' \geqq \sqrt{3}$ ist. Nach der Bemerkung **2** können wir voraussetzen, dass die Seiten von $A'B'C' < 2$ sind. Wegen der Wahl des Vektors $\overrightarrow{OA}$ ist $AB \geqq OA$ und $AC \geqq OA$, also $\sphericalangle AOC \geqq \sphericalangle ACO$ und $\sphericalangle AOB \geqq \sphericalangle ABO$. Deshalb können wir voraussetzen, dass $OA = 2$ ist, weil wir im Falle $OA > 2$ den Punkt A durch einen Punkt A^* ersetzen können, für den $OA^* = 2$ und $\overrightarrow{OA^*} \| \overrightarrow{OA}$ ist. Weiterhin können wir ohne Beschränkung der Allgemeinenheit voraussetzen, dass die senkrechte Projektionen von B und C auf die Gerade OA auf der offenen Strecke OA liegen.

Wir bezeichnen die durch den Mittelpunkt F der Strecke OA gehende und zu der Ebene π parallele Ebene mit π_1. Wir untersuchen die folgenden Möglichkeiten:

I. Die Ebene π_1 trennt die Punkte B und C nicht.

II. Die Ebene π_1 trennt die Punkte B und C.

I. Wir können voraussetzen, dass die Punkte B, C und O auf derselben Seite von π_1 sind. Weiterhin setzen wir voraus, dass B näher bei π_1 liegt als C. Liegt B nicht in π_1, dann bewegen wir B auf der zu OA parallelen Geraden in π_1. Inzwischen verändert sich der Radius von k nicht, und die Entfernungen werden nicht kleiner als 2.

Ist die Strecke $FB > \sqrt{3}$, dann bewegen wir B auf dem Durchschnitt k_1 von H und π_1, so dass $FB = \sqrt{3}$ wird. In diesem Fall ist $AB = OB = 2$ und inzwischen hat sich der Radius von k nicht verändert. Wir werden zeigen, dass diese Bewegung möglich ist, d.h. dass während der Bewegung der Abstand BC nicht < 2 wird. Wir bezeichnen die senkrechte Projektion von C auf π_1 mit C_1, das Spiegelbild von C_1 bezüglich des Mittelpunktes von k_1 mit $\bar{C}_1$. Da $FC_1 \geqq \sqrt{3}$ und der Radius von

$k_1 < \sqrt{3}$ ist, erhalten wir $F\bar{C}_1 \leqq \sqrt{3}$. Deshalb nimmt der Abstand BC während der Bewegung von B monoton zu.

Wir bewegen den Punkt C parallel zu der Geraden OA auf π_1 zu, so dass $CB=2$ sei. Dann bewegen wir C auf der Durchdringungskurve von H und der um B mit dem Radius 2 geschlagenen Kugel gegen π_1. Nach der Bemerkung **2** können wir doraussetzen, dass wir den Punkt C nicht bis π_1 bewegen können. Während der Bewegungen verändert sich der Radius von k nicht. Nach den Bewegungen sind vie Kantenlängen des Tetraeders $OABC$ mit Ausnahme von AC gleich 2.

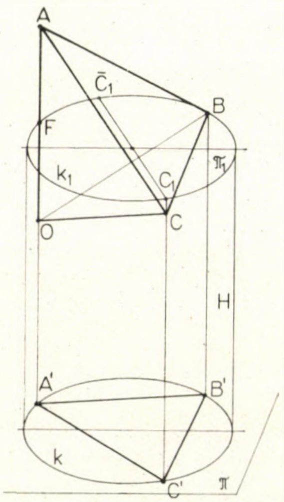


Abb. 1

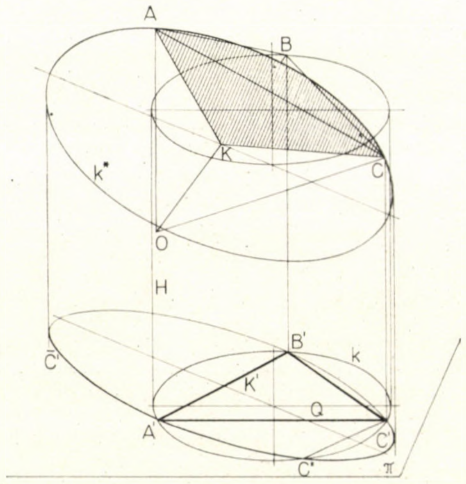


Abb. 2

Wir drehen den Punkt C um die Achse OB so, dass $AC=2$ sei. Wir werden zeigen, dass in diesem Fall der Radius von k abnimmt. Wir bezeichnen den Mittel-Punkt der Strecke OB mit K, die senkrechte Projektion von K auf π mit K'. Der Punkt C bewegt sich auf einem Kreis k^* mit dem Mittelpunkt K. Die senkrechte Projektion von k^* auf π ist eine Ellipse. Der Mittelpunkt der Ellipse ist K', seine kleine Achse ist $A'B'$; weiterhin geht sie durch den Punkt C' und durch den Spiegelpunkt $\bar{C}'$ von C' bezüglich K', sowie durch den Spiegelpunkt C'' von C' bezüglich der Mittelsenkrechte der Strecke $A'B'$ (Abb. 2.).

Nach der Bemerkung **4** ist $\bar{C}'$ ein äusserer Punkt von k. Während der Bewegung bewegt sich die Projektion von C auf dem Ellipsenbogen $A'C'$.

Wenn $A'C' \leqq C'B'$ ist dann liegt der Ellipsenbogen im Inneren von k; deshalb nimmt der Umkreisradius des Dreiecks $A'B'C'$ ab. Wenn $A'C' = \sqrt{3}$ ist, gelangt der Punkt C in die Ebene π_1 und es gilt $AC=2$.

Wenn $C'A' > C'B'$ ist, dann bewegen wir den Punkt C so, dass seine Projektion C'' sei (Abb. 2.). Wegen $A'C'' < C''B'$ haben wir wiederum mit dem ersten Fall zu tun. Nach der Drehung ist der Radius von k gleich $\frac{3\sqrt{2}}{4}$.

II. Die Punkte C und O bzw. B und A liegen auf zwei verschiedenen Seiten von π_1. Wir können die Punkte C und B parallel zu OA so bewegen, dass $OC=CB=2$ sei. Auf Grund von I. können wir voraussetzen, dass π_1 die Punkte B und C trennt. Bewegen wir den Punkt B auf der Durchdringungskurve von H und der um C mit dem Radius 2 geschlagenen Kugel gegen π_1. Während dieser Bewegung kommt der Punkt C entweder in π_1 oder, wird AB gleich 2. Auf Grund von I. können wir voraussetzen, dass $AB=2$ ist. Gleichzeitig gilt für die übrigen Kanten des Tetraeders $OABC$: $OA=OC=CB=2$, $AC>2$, $OB>2$. Wir drehen die Punkte C und B um die Achsen OB bzw. AC so, dass $AC=OB=2$ sei. Es ist leicht einzusehen, dass für die Projektionen B' und C' diejenigen Bedingungen des Falles I., die bei der Drehung benutzt wurden, erfüllt sind. Also nimmt der Umkreisradius des Dreiecks $A'B'C'$ ab, und nach den Drehungen wird $R=\frac{3\sqrt{2}}{4}$. Damit haben wir bewiesen, dass der Radius von k grösser oder gleich als $\frac{3\sqrt{2}}{4}$ ist.

Auf Grund der Bemerkungen **3** und **4** enthält der Kreis k ausser A', B' und C' keinen anderen Gitterpunkt. Q sei der Mittelpunkt von k. Der Kreis $\bar{k}$ mit dem Mittelpunkt Q und dem Radius $\frac{3\sqrt{2}}{4}-1$ wird die Projektionskreise der Kugeln höchstens berühren, deshalb wird der Zylinder mit dem Grundkreis $\bar{k}$ und der zu OA parallelen Achse die Kugeln höchstens berühren.

LITERATUR

[1] FEJES TÓTH, L.: Verdeckung einer Kugel durch Kugeln, *Publ. Mat. Debrecen*, **6** (1959), 234—240

[2] BÖRÖCZKY, K.: Über Dunkelwolken, *Proc. Coll. Convexity, Koppenhagen*, 1965 (1967), 13—17

[3] HEPPES, A.: Ein Satz über gitterförmige Kugelpackungen, *Annales Univ. Sci. Budapest, Sect. Math.*, **3—4**, (1960—61), 89—90.

Eötvös L. Universität, Budapest

(Eingegangen: 20 Oktober, 1969.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 151—160.

A HOMOMORPHISM THEOREM WITH AN APPLICATION TO THE CONJECTURE OF HADWIGER

by

I. T. JAKOBSEN

1. Definitions and Terminology. In this paper a graph is an undirected graph without loops and without multiple edges.

Let Γ be a graph. Then $V(\Gamma)$ denotes the set of vertices and $E(\Gamma)$ the set of edges. Γ is said to be *finite*, if $V(\Gamma)$ is finite. $n(\Gamma)$ denotes the number of vertices in Γ and $e(\Gamma)$ the number of edges. If S is a set, $|S|$ denotes the number of elements of S. An edge joining two vertices x and y is denoted by (x, y) or (y, x). A graph with v vertices in which each pair of distinct vertices are joined by an edge is called a *complete v-graph* and denoted by $\langle v\rangle$. A $\langle v\rangle$ with just one edge deleted is denoted by $\langle v-\rangle$ and a $\langle v\rangle$ with exactly two edges deleted is denoted by $\langle v=\rangle$.

If the graph Γ' is contained in the graph Γ as a subgraph (i.e. $V(\Gamma')\subseteqq V(\Gamma)$ and $E(\Gamma')\subseteqq E(\Gamma)$), $\Gamma=\Gamma'$ possibly, we write $\Gamma'\subseteqq\Gamma$. If $W\subseteqq V(\Gamma)$ then $\Gamma-W$ denotes the graph obtained from Γ by deleting all vertices belonging to W and all edges incident with at least one vertex of W. $\Gamma(W)$ denotes the *subgraph of Γ spanned by the set W* defined as the subgraph of Γ whose set of vertices is W and whose set of edges is the set of all edges of Γ having both end-vertices in W. Any such subgraph is called a *spanned subgraph*. If $\Gamma'\subseteqq\Gamma$, then $\Gamma-V(\Gamma')$ is also written $\Gamma-\Gamma'$. If $x, y\in V(\Gamma)$, $(x, y)\notin E(\Gamma)$ (resp. $(x, y)\in E(\Gamma)$), then $\Gamma\cup(x, y)$ (resp. $\Gamma-(x, y)$) denotes the graph obtained from Γ by adding (resp. deleting) the edge (x, y). For convenience if $(x, y)\notin E(\Gamma')$ then $\Gamma'-(x, y):=\Gamma'$. If $x\notin V(\Gamma')$, $\Gamma'\cup x$ denotes $\Gamma(V(\Gamma')\cup\{x\})$. $v(x, \Gamma)$ denotes the *valency* of x in Γ, i.e. the number of vertices of Γ joined to x by edges in Γ.

A *path* is a graph with vertices $x_1, x_2, \dots, x_\mu$, $\mu\geqq 2$ and edges (x_1, x_2), (x_2, x_3), … …, $(x_{\mu-1}, x_\mu)$, where $x_1, \dots, x_\mu$ are all distinct. x_1 and x_μ are called the end-vertices of the path and are said to be joined by the path. Let Π be a path and $x, y\in V(\Pi)$. Then $\Pi[x, y]$ denotes that subgraph of Π which is a path and has x and y as its two end-vertices. $\Pi[x, x]$ is defined as the graph consisting of the vertex x.

Let Γ_1 and Γ_2 be two mutually disjoint subgraphs of Γ. A *$(\Gamma_1)(\Gamma_2)$-path* is a path with one end-vertex belonging to Γ_1 and the other one to Γ_2 and which has nothing else in common with $\Gamma_1\cup\Gamma_2$.

A set of vertices of Γ is called *independent* if no two of them are joined by an edge, and a set of edges of Γ is called *independent* if no two of them have a vertex in common.

Γ is said to be *λ-fold connected*, $\lambda\geqq 1$, if $n(\Gamma)\geqq\lambda+1$ and whenever $\leqq\lambda-1$ vertices are deleted the remaining graph is connected.

If Γ is connected, a *cut-set* of Γ is a set S of vertices such that $\Gamma-S$ is disconnected.

Let Δ be another graph. A *contraction* is a mapping m from $V(\Gamma)$ onto $V(\Delta)$

such that (1) $\forall x \in V(\Delta)$: $\Gamma(m^{-1}(x))$ is connected and (2) $\forall x, y \in V(\Delta)$: $(x, y) \in E(\Delta)$ if and only if Γ contains at least one edge joining a vertex of $m^{-1}(x)$ and a vertex of $m^{-1}(y)$. Δ is said to be obtained from Γ through a *contraction* and Γ is said to be *contracted* into Δ if such a mapping exists and is applied on Γ. Γ is said to be *homomorphic* to Δ, written $\Gamma \succ \Delta$, if Γ can be contracted into a graph containing Δ as a subgraph.

Let Γ be a connected graph. Let Λ be a spanned proper subgraph of Γ. Let C denote a connected component of $\Gamma - \Lambda$ and let x be any vertex of Λ joined to C. Any contraction P from $\Gamma(\Lambda \cup C)$ onto a graph Δ defined by contracting $C \cup x$ into one vertex and keeping the other vertices of Λ fixed is called a *simple projection from C onto* Λ and Δ is denoted by $P\Lambda$. $V(\Delta) = V(\Lambda)$ and $\Delta \supseteqq \Lambda$. (A similar concept has been introduced by W. MADER in [6]).

Let $x, y \in V(\Lambda)$, $(x, y) \notin E(\Lambda)$. If by the simple projection P a graph $P\Lambda \supseteqq \Lambda \cup (x, y)$ is obtained, the new edge (x, y) is said to be *provided by P* for Λ from C.

For convenience the identical mapping on $V(\Lambda)$ is considered as a simple projection onto Λ. If P is a simple projection, then $P\Lambda = \Lambda$ possibly. Clearly $\Gamma \succ P\Lambda$.

For $k \geqq 2$: Γ is called *k-colourable* if there exists a partition of $V(\Gamma)$ into k disjoint classes each consisting of independent vertices. For $k \geqq 3$: Γ is called *k-chromatic* if it is k-colourable but not $(k-1)$-colourable. Γ is called *contraction-critical k-chromatic* if it is k-chromatic and connected and every graph with fewer vertices to which Γ can be contracted is $(k-1)$-colourable.

A $\langle v \rangle$*-cockade of strength* μ is any graph constructed by the following procedure: $\Phi_1, \ldots, \Phi_\varkappa$, are $\varkappa$ mutually disjoint $\langle v \rangle - s$, $\varkappa \geqq 1$, $v \geqq 2$. Let $1 \leqq \mu \leqq v-1$; if $\varkappa \geqq 2$ a $\langle \mu \rangle$ contained in Φ_1 and a $\langle \mu \rangle$ contained on Φ_2 are selected and identified with each other. If $\varkappa \geqq 3$ a new $\langle \mu \rangle$ is selected from the graph constructed in this way and identified with a $\langle \mu \rangle$ in Φ_3 and so on with $\Phi_4, \ldots, \Phi_\varkappa$, if $\varkappa \geqq 4$. The class of all $\langle v \rangle$-cockades of strength μ is denoted by $\mathscr{K}_v^\mu$. (These graphs were first described by G. A. DIRAC in [2]).

2. Summary. The object of the present paper s to give a proof of an extension of a certain kind of homomorphism-theorem and a new partial result concerning the Conjecture of HADWIGER [4]) for $k = 7$.

In [2] G. A. DIRAC proved for $v = 5$ and 6 that if Γ is a finite graph $n(\Gamma) \geqq v$, $e(\Gamma) \geqq (v-3)n - \frac{1}{2}(v-1)(v-4)$ and $\Gamma \notin \mathscr{K}_{v-1}^{v-4}$ then $\Gamma \succ \langle v = \rangle$. He also asked whether this is true for all v.

The present author has proved that the above holds for $v = 7$ and 8 as well, but in this paper only the result for $v = 7$ is included as Theorem 1 (for $v = 8$ the result is yet unpublished). For $v = 9$ it ceases to be true, as a complete 10-graph with five independent edges deleted shows.

In [7] K. WAGNER proved that every 5-chromatic graph is homomorphic to a $\langle 5 - \rangle$ and in [3] G. A. DIRAC proved that every 6-chromatic graph is homomorphic to a $\langle 6 - \rangle$. These results are partial results concerning the Conjecture of Hadwiger for $k = 5$ and 6. The Conjecture of Hadwiger states that every k-chromatic graph is homomorphic to a $\langle k \rangle$. Theorem 2 of the present paper gives a new partial result for $k = 7$, namely that every 7-chromatic graph is homomorphic to a $\langle 7 = \rangle$. This implies the five-colour theorem for planar graphs.

I want to thank G. A. DIRAC for his encouragement to take up this subject and B. TOFT for valuable discussions. As far as the proof-ideas are concerned I owe much to the papers [2] and [3] by G. A. DIRAC.

3. Results to be used

(A) A member of $\mathcal{K}_{\nu-1}^{\nu-4}$ with n vertices has exactly $(\nu-3)n-\frac{1}{2}(\nu-1)(\nu-4)$ edges, hence for $\nu=7$: a member of $\mathcal{K}_6^3$ with n vertices has exactly $4n-9$ edges. ([2].)
(B) If Γ is a finite graph with $n\geqq 6$ vertices and e edges such that $e\geqq 3n-5$ and $\Gamma\notin\mathcal{K}_5^2$, then $\Gamma\succ\langle 6=\rangle$. ([2].)
(C) An extension of the theorem of MENGER. ([1].)
(D) Let Γ be a contraction-critical k-chromatic graph different from a $\langle k\rangle$. Then
1. No three of the vertices joined to a vertex of valency k are independent.
2. If $k\geqq 7$ then Γ is 7-fold connected. (For D1 see [3]; for D2 see [5].)

4. A Homomorhism Theorem

Remark. Let $K\in\mathcal{K}_\nu^\mu$, K composed of $\Phi_1, \Phi_2, \ldots, \Phi_\varkappa$. Then every complete subgraph of K is a subgraph of some Φ_i.

LEMMA 1. *Let* $K\in\mathcal{K}_{\nu-1}^{\nu-4}$, $\nu\geqq 5$. *Then*
(A) If $K\neq\langle\nu-1\rangle$, *then* K *contains at least six vertices of valency* $\nu-2$, *among them there are six which in* K *span a graph consisting of two disjoint triangles.*
(B) $K\not\succ\langle\nu=\rangle$, *but if* $K\neq\langle\nu-1\rangle$ *and an edge joining any two vertices not already joined by an edge is added to* K *and any edge of* K *deleted, then the resulting graph is homomorphic to a* $\langle\nu=\rangle$.

PROOF. (A) and (B) are proved by induction over the number $\varkappa$ of $\langle\nu-1\rangle-s$ of which K is composed.

If $K=\langle\nu-1\rangle$, $K\not\succ\langle\nu=\rangle$. Hence assume $K\neq\langle\nu-1\rangle$, then $\varkappa\geqq 2$.

For $\varkappa=2$ (A) and (B) are easily verified.

Suppose then that the conclusions hold for cockades composed of fewer than $\varkappa$ $(\varkappa\geqq 3)$ $\langle\nu-1\rangle-s$. Let K be composed of $\varkappa\langle\nu-1\rangle-s$, $\Phi_1, \Phi_2, \ldots, \Phi_\varkappa$, successively. The cockade K' composed of $\Phi_1, \Phi_2, \ldots, \Phi_{\varkappa-1}$, successively, is a member of $\mathcal{K}_{\nu-1}^{\nu-4}$ composed of $\varkappa-1(\geqq 2)$ $\langle\nu-1\rangle-s$, hence $K'\neq\langle\nu-1\rangle$.

(A): By the induction hypothesis K' contains at least six vertices of valency $\nu-2$ in K', six of which span a graph consisting of two disjoint triangles. At most one of these triangles can have anything in common with $\Phi_\varkappa$, hence K contains three vertices of K' of valency $\nu-2$ in K, spanning a triangle and having nothing in common with $\Phi_\varkappa$. Furthermore three of the vertices of $\Phi_\varkappa$ have valency $\nu-2$ in K, span a triangle and have nothing in common with K'. Hence K contains at least six vertices of valency $\nu-2$ in K, six of which span a graph consisting of two disjoint triangles. Hence B holds.

(B): $K'\not\succ\langle\nu=\rangle$ by the induction hypothesis and $K'\cap\Phi_\varkappa=\langle\nu-4\rangle$, hence $K\not\succ\langle\nu=\rangle$.

Let $s, t\in V(K)$ such that $(s, t)\notin E(K)$ and let $k_1, k_2, \ldots, k_{\nu-1}$ be the vertices of $\Phi_\varkappa$, the notation being chosen so that $k_{\nu-1}, k_{\nu-2}, k_{\nu-3}$ have valency $\nu-2$ in K. Two alternative cases are considered:

i) $s, t\neq k_{\nu-1}, k_{\nu-2}, k_{\nu-3}$. Then (s, t) is added to K'. If any edge is deleted from K' or from $\Phi_\varkappa$ it follows by the induction hypothesis that the resulting graph is homomorphic to a $\langle\nu=\rangle$.

ii) *s or t is one of $k_{\nu-1}, k_{\nu-2}, k_{\nu-3}$.* Say $s=k_{\nu-1}$. If t is joined by an edge to every vertex of $\Phi_\varkappa - k_{\nu-1} - k_{\nu-2} - k_{\nu-3}$ then (cf. Remark) $V(\Phi_\varkappa - k_{\nu-1} - k_{\nu-2} - k_{\nu-3})$ and t are all vertices of the same $\langle \nu-1 \rangle$, Φ_i, say. $i \neq \varkappa$, because $t \notin \Phi_\varkappa$. Let ε denote an arbitrary edge of K. $\Phi_i \cup \Phi_\varkappa \cup (s, t) - \varepsilon \succ \langle \nu = \rangle$ by the induction hypothesis because $\Phi_i \cup \Phi_\varkappa$ is a $\langle \nu-1 \rangle$-cockade of strength $\nu-4$ composed of only two $\langle \nu-1 \rangle - s$. If t is not joined by edges to all the vertices of $\Phi_\varkappa - k_{\nu-1} - k_{\nu-2} - k_{\nu-3}$, say $(t, k_1) \notin E(K)$, then, if ε again denotes an arbitrary edge of K: $K \cup (s, t) - \varepsilon \succ$ $\succ K' \cup (t, k_1) - \varepsilon \succ \langle \nu = \rangle$ by the induction hypothesis.

This completes the proof of Lemma 1.

THEOREM 1. *Let Γ be a finite graph with n vertices and e edges. If $n \geqq 7$, $e \geqq 4n-9$ and $\Gamma \notin \mathscr{K}_6^3$, then $\Gamma \succ \langle 7 = \rangle$.*

PROOF. By induction over n. The theorem is trivially true for $n=7$.

Induction hypothesis: Assume the theorem is true for all graphs with m vertices satisfying the conditions, where $7 \leqq m \leqq n-1$.

Let Γ be a finite graph with n vertices and e edges satisfying the conditions of the theorem.

It is sufficient to consider the case $e = 4n-9$ in the rest of the proof. For assume that $e > 4n-9$ and that the theorem holds for all graphs with n vertices having exactly $4n-9$ edges. By deleting edges from Γ a graph Γ^* may be obtained such that $e(\Gamma^*) = 4n-9$. If $\Gamma^* \notin \mathscr{K}_6^3$ then $\Gamma^* \succ \langle 7 = \rangle$ by the last assumption, and therefore $\Gamma \succ \langle 7 = \rangle$. If $\Gamma^* \in \mathscr{K}_6^3$ then, because $n > 7$ and by Lemma 1B again $\Gamma \succ \langle 7 = \rangle$. This proves the assertion.

(1) If $\exists x \in V(\Gamma)$: $v(x, \Gamma) \leqq 3$, then $\Gamma \succ \langle 7 = \rangle$.

Proof of (1): $n(\Gamma - x) = n-1 \geqq 7$, $e(\Gamma - x) \geqq e-3 = 4(n-1)-8$. By (A) and the induction hypothesis $\Gamma - x \succ \langle 7 = \rangle$.

(2) Let Γ' be a graph with n' vertices, $n' < n$, and e' edges. If $n' \geqq 5$, $e' \geqq 4n'-9$, then either $\Gamma' \succ \langle 7 = \rangle$ or $\Gamma' \in \mathscr{K}_6^3$.

Proof of (2): If $n'=5$, then $e' \geqq 11$ impossible. If $n'=6$, then $e' \geqq 15$, hence $\Gamma' = \langle 6 \rangle \in \mathscr{K}_6^3$. If $n' \geqq 7$, then by the induction hypothesis $\Gamma' \succ \langle 7 = \rangle$ or $\Gamma' \in \mathscr{K}_6^3$. This proves (2).

(3) If Γ is disconnected or has a cut-set S such that $|S| \leqq 3$ or such that $|S| = 4$ and $\Gamma(S) = \langle 4 \rangle, \langle 4 - \rangle$, or $\langle 4 = \rangle$, then $\Gamma \succ \langle 7 = \rangle$.

Proof of (3): If Γ has a cut-set, then it has a minimal cut-set. In the sequel let S denote a minimal cut-set of Γ, when Γ has a cut-set.

Let $\Gamma = \Gamma_1 \cup \Gamma_2$, Γ_1 and Γ_2 being spanned subgraphs of Γ, $V(\Gamma_1 \cap \Gamma_2) = S$, where $S = \emptyset$ possibly. If $S \neq \emptyset$, S is a minimal cut-set of Γ. $\Gamma_1 - S$, $\Gamma_2 - S \neq \emptyset$. Let $|S| = \sigma$ and the vertices of S be denoted by $s_1, \ldots, s_\sigma$, $\sigma = 1$ possibly. Let $|E(\Gamma(S))| = p$ and $|V(\Gamma_i)| = n_i$, $|E(\Gamma_i)| = e_i$ for $i = 1, 2$. Then $n = n_1 + n_2 - \sigma$ and $e = e_1 + e_2 - p$. If $S \neq \emptyset$, then let Γ_i' be a connected component of $\Gamma - S$ contained in $\Gamma_i - S$ for $i = 1, 2$; Γ_i' is joined by edges to every vertex of S, because S is a minimal cut-set.

Let P_1 denote the simple projection from Γ_1' onto Γ_2 obtained by contracting $\Gamma_1' \cup s_1$ into one vertex, and let P_2 denote the simple projection from Γ_2' onto Γ_1 obtained by contracting $\Gamma_2' \cup s_1$ into one vertex.

If for $i=1$ or 2 $n_i \leqq 4$, then every vertex of $\Gamma_i - S$ has valency $\leqq 3$ in Γ, therefore by (1) $\Gamma \succ \langle 7= \rangle$ in this case. Hence it may be assumed from now on that

$$(3.1) \qquad n_i \geqq 5, \qquad i=1,2.$$

$$e_1+e_2 = e+p = 4n-9+p = 4(n_1+n_2)-4\sigma-9+p.$$

By the symmetry between Γ_1 and Γ_2 it may be assumed that $e_1 \geqq 4n_1 - \frac{1}{2}(4\sigma+9-p)$.
i) $\sigma=4$.

1) $\Gamma(S)=\langle 4 \rangle$.

Then $p=6$ and $e_1 \geqq 4n_1 - \dfrac{19}{2}$, hence $e_1 \geqq 4n_1-9$. By (3. 1) and (2) $\Gamma_1 \succ \langle 7= \rangle$ except when $\Gamma_1 \in \mathscr{K}_6^3$. If $\Gamma_1 \in \mathscr{K}_6^3$ then $\Gamma(S)$ is contained in a $\langle 6 \rangle \subseteqq \Gamma_1$. By contracting Γ_2' into one vertex Γ is contracted into a graph containing a $\langle 7= \rangle$ as a subgraph.

2) $\Gamma(S)=\langle 4- \rangle$.

Then $p=5$ and $e_1 \geqq 4n_1-10$. Assume w.l.g. $(s_1, s_2) \notin E(\Gamma)$. Consider $P_2\Gamma_1 = \Gamma_1 \cup (s_1, s_2)$. $n(P_2\Gamma_1)=n_1 \geqq 5$ by (3. 1), $e(P_2\Gamma_1) = e_1+1 \geqq 4n_1-9$. By (2) $P_2\Gamma_1 \succ \langle 7= \rangle$ except when $P_2\Gamma_1 \in \mathscr{K}_6^3$. In this case by (A) $e_1 = 4n_1-10$ and consequently $e_2 \geqq 4n_2-10$. Consider $P_1\Gamma_2 = \Gamma_2 \cup (s_1, s_2)$. $n(P_1\Gamma_2)=n_2 \geqq 5$ by (3. 1), $e(P_1\Gamma_2) = e_2+1 \geqq 4n_2-9$. By (2) $P_1\Gamma_2 \succ \langle 7= \rangle$ except when $P_1\Gamma_2 \in \mathscr{K}_6^3$. Suppose that $P_1\Gamma_2 \in \mathscr{K}_6^3$. $P_2\Gamma_1(S)$ is contained in a $\langle 6 \rangle \subseteqq P_2\Gamma_1$ and $P_1\Gamma_2(S)$ is contained in a $\langle 6 \rangle \subseteqq P_1\Gamma_2$. Let the former $\langle 6 \rangle$ be denoted by Λ' and the latter by Λ''. Then $\Gamma \supseteqq \Lambda' \cup \Lambda'' - (s_1, s_2) \succ \langle 7= \rangle$.

3) $\Gamma(S)=\langle 4= \rangle$.

Then $p=4$ and $e_1 \geqq 4n_1 - \dfrac{21}{2}$, hence $e_1 \geqq 4n_1-10$. Assume w.l.g. $(s_1, s_2) \notin E(\Gamma)$ and s_3 is incident with the other missing edge, denoted by ε, and s_1 is not incident with ε. Consider $P_2\Gamma_1 = \Gamma_1 \cup (s_1, s_2)$. $n(P_2\Gamma_1)=n_1 \geqq 5$ by (3. 1). $e(P_2\Gamma_1) = e_1+1 \geqq 4n_1-9$. By (2) $P_2\Gamma_1 \succ \langle 7= \rangle$ except when $P_2\Gamma_1 \in \mathscr{K}_6^3$. $P_2\Gamma_1 \neq \langle 6 \rangle$ because $\varepsilon \notin E(P_2\Gamma_1)$. Assume then that $P_2\Gamma_1 \in \mathscr{K}_6^3$. P' denotes the projection from Γ_2' onto Γ_1 obtained by contracting $\Gamma_2' \cup s_3$ into one vertex. $P'\Gamma_1 = P_2\Gamma_1 - (s_1, s_2) \cup \varepsilon$ and $\varepsilon \notin E(P_2\Gamma_1)$, hence by Lemma 1B $P'\Gamma_1 \succ \langle 7= \rangle$.
ii) $\sigma=3$.

1) $p=3$.

Then $e_1 \geqq 4n_1-9$. By (3. 1) and (2) $\Gamma \succ \langle 7= \rangle$ except when $\Gamma_1 \in \mathscr{K}_6^3$. In the latter case $e_1 = 4n_1-9$ by (A) and consequently $e_2 \geqq 4n_2-9$; by (3. 1) and (2) $\Gamma_2 \succ \langle 7= \rangle$ except when $\Gamma_2 \in \mathscr{K}_6^3$. But if this is so then $\Gamma \in \mathscr{K}_6^3$ contrary to hypothesis. Hence $\Gamma \succ \langle 7= \rangle$ in this case.

2) $p=2$.

Then $e_1 \geqq 4n_1 - \dfrac{19}{2}$. Assume w.l.g. $(s_1, s_2) \notin E(\Gamma)$. Consider $P_2\Gamma_1 = \Gamma_1 \cup (s_1, s_2)$. $n(P_2\Gamma_1)=n_1 \geqq 5$ by (3. 1), $e(P_2\Gamma_1) = e_1+1 > 4n_1-9$. By (2) and (A) $P_2\Gamma_1 \succ \langle 7= \rangle$.

3) $p \leqq 1$.

Then $e_1 \geqq 4n_1 - \frac{21}{2}$. Assume w.l.g. $(s_1, s_2), (s_1, s_3) \notin E(\Gamma)$. Consider $P_2\Gamma_1 = \Gamma_1 \cup (s_1, s_2) \cup (s_1, s_3)$. $n(P_2\Gamma_1) = n_1 \geqq 5$, $e(P_2\Gamma_1) = e_1 + 2 > 4n_1 - 9$. By (2) and (A) $P_2\Gamma_1 \succ \langle 7= \rangle$.

iii) $\sigma \leqq 2$.

Then $e_1 \geqq 4n_1 - 8$, hence by (3. 1), (2) and (A) $\Gamma_1 \succ \langle 7= \rangle$.

This completes the proof of (3).

(4) If Γ is 4-fold connected and Γ is not separated by a $\langle 4 \rangle$, $\langle 4- \rangle$, or $\langle 4= \rangle$, then $\Gamma \succ \langle 7= \rangle$.

Proof of (4): Assume Γ has the properties stated in (4). The proof will be by the steps (4. 1)—(4. 8).

(4. 1) If $\Gamma \supseteqq \langle 6 \rangle$, then $\Gamma \succ \langle 7= \rangle$.

Proof of (4. 1): Let Λ be a $\langle 6 \rangle \subseteqq \Gamma$. $\Gamma - \Lambda \neq \emptyset$, because $n > 7$. Let C be a connected component of $\Gamma - \Lambda$. Γ is 4-fold connected, hence C is joined to at least 4 of the vertices of Λ and by contracting C into one vertex Γ is contracted into a graph containing a $\langle 7= \rangle$ as a subgraph. This proves (4. 1).

Let x_0 be a vertex of minimal valency in Γ. $v(x_0, \Gamma) = j$, say. $j \geqq 4$ because Γ is 4-fold connected. $j \geqq 8$ implies $e \geqq 4n$ contrary to the fact that $e = 4n - 9$. Hence

(4. 2) $$4 \leqq j \leqq 7.$$

Let the vertices joined to x_0 be denoted by $x_1, x_2, \ldots, x_j$. $\Gamma(x_1, x_2, \ldots, x_j)$ is denoted by Γ_j.

(4. 3) If $\Gamma_j = \langle j \rangle$, then $\Gamma \succ \langle 7= \rangle$.

Proof of (4. 3): $j \geqq 5$ implies $\Gamma_j \cup x_0 = \langle 6 \rangle$, hence by (4. 1) $\Gamma \succ \langle 7= \rangle$. $j = 4$ implies $e(\Gamma - x_0) = 4(n-1) - 9$. By (2) with $n' = n - 1 \geqq 7$ $\Gamma - x_0 \succ \langle 7= \rangle$ except when $\Gamma - x_0 \in \mathscr{K}_6^3$. But in the latter case $\Gamma \supseteqq \langle 6 \rangle$ and by (4. 1) $\Gamma \succ \langle 7= \rangle$.

(4. 4) If $\exists x_i \in V(\Gamma_j)$: $v(x_i, \Gamma_j) \leqq 2$, then $\Gamma \succ \langle 7= \rangle$.

Proof of (4. 4): Assume w.l.g. $i = 1$. By contracting $\Gamma(x_0, x_1)$ into one vertex Γ is contracted into Γ' say. $n(\Gamma') = n - 1 \geqq 7$, $e(\Gamma') \geqq e - j + j - 3 \geqq 4(n-1) - 8$. By (2) and (A) $\Gamma' \succ \langle 7= \rangle$.

(4. 5) If $\forall x_k \in V(\Gamma_j)$: $v(x_k, \Gamma_j) \geqq 3$ and

$$\exists x_i \in V(\Gamma_j): v(x_i, \Gamma_j) = 3, \text{ then } \Gamma \succ \langle 7= \rangle.$$

Proof of (4. 5): $j = 4$ implies $\Gamma_j = \langle 4 \rangle$, hence because of (4. 3) $j \geqq 5$ may be assumed.

Assume w.l.g. $i = 1$ and x_1 joined to x_2, x_3, x_4. By contracting $\Gamma(x_0, x_1)$ into one vertex Γ is contracted into $(\Gamma - x_0) \cup \bigcup_{k=5}^{j} (x_1, x_k) = \Gamma'$. $n(\Gamma') = n - 1 \geqq 7$, $e(\Gamma') = e - j + j - 4 = 4(n-1) - 9$. By (2) $\Gamma' \succ \langle 7= \rangle$ except when $\Gamma' \in \mathscr{K}_6^3$. Assume then that $\Gamma' \in \mathscr{K}_6^3$. $\Gamma' \neq \langle 6 \rangle$ because $n - 1 \geqq 7$.

If $j \geqq 6$: By the contraction of $\Gamma(x_0, x_1)$ only the vertices x_2, x_3, x_4 have their valency decreased; the valency of x_1 is not decreased because $j > 4$ and $(x_1, x_j) \notin E(\Gamma)$. The valencies of x_2, x_3, x_4 each decrease by 1. Hence the minimal valency of Γ' is $\geqq j - 1 \geqq 5$ and at most three vertices of Γ' have valency 5, contrary to Lemma 1A.

The case $j=5$ remains to be dealt with. Then x_1 is joined to x_2, x_3, x_4 and not to x_5. Hence $\Gamma' = (\Gamma - x_0) \cup (x_1, x_5)$. Assume that Γ' contains a $\langle 6 \rangle$ to which $x_1, \dots, x_5$ all belong; then in Γ x_0 is joined to five vertices of a $\langle 6- \rangle \subseteqq \Gamma - x_0$, hence $\Gamma \supseteqq \langle 7= \rangle$. Assume next that this is not the case. Then one of the edges (x_2, x_3), (x_3, x_4), (x_4, x_5) is not in Γ_5, assume w.l.g. that $(x_2, x_3) \notin E(\Gamma)$. By contracting $\Gamma(x_0, x_2)$ into one vertex Γ is contracted into a graph containing $(\Gamma - x_0) \cup (x_2, x_3) = \Gamma' \cup (x_2, x_3) - (x_1, x_5) \succ \langle 7= \rangle$ by Lemma 1B, because $(x_2, x_3) \notin E(\Gamma')$.

This proves (4. 5).

As a consequence of (4. 4) and (4. 5) it may be assumed that

(4. 6) $$\forall x_k \in V(\Gamma_j)\colon\ v(x_k, \Gamma_j) \geqq 4.$$

Then necessarily $j \geqq 5$; $j=5$ implies $\Gamma_j = \langle j \rangle$, hence by (4. 3) $\Gamma \succ \langle 7= \rangle$ in this case. It may then be assumed that

(4. 7) $$j \geqq 6.$$

Assume $\Gamma - \Gamma_j - x_0 = \emptyset$. Then $\Gamma = \Gamma_j \cup x_0$, hence $n = j+1$. j is the minimal valency of Γ, hence $\Gamma = \langle j+1 \rangle \supseteqq \langle 7 \rangle$ by (4. 7). Consequently

(4. 8) $$\Gamma - \Gamma_j - x_0 \neq \emptyset$$

may be assumed from now on. Every connected component of $\Gamma - \Gamma_j - x_0$ is joined to at least four vertices of Γ_j because Γ is 4-fold connected.

By (4. 2) and (4. 7) $j=6$ or 7.

Let C be a connected component of $\Gamma - \Gamma_j - x_0$. C is joined to at least four vertices of Γ_j.

Assume there exists a simple projection P from C onto Γ_j (possibly P is the identical mapping on $V(\Gamma_j)$) such that $e(P\Gamma_j) \geqq e(\Gamma_j) + j - 5$. Then $e(P\Gamma_j) \geqq \frac{4}{2} j + j - 5 = 3j - 5$. $P\Gamma_j \notin \mathscr{K}_5^2$ because a member of $\mathscr{K}_5^2$ cannot have just six or seven vertices. Hence by (B) $P\Gamma_j \succ \langle 6= \rangle$ and consequently $\Gamma \succ \langle 7= \rangle$.

Assume then that $j-5$ new edges cannot be provided for Γ_j by any simple projection from C onto Γ_j. For $j=6$: If C is joined to two vertices of Γ_6 not joined by an edge, one new edge can be provided, contradiction. But a $\langle 4 \rangle$ does not separate Γ according to the assumptions of (4), hence C is joined to at least five vertices of Γ_6 spanning a $\langle 5 \rangle$, but then $\Gamma \succ \langle 7- \rangle$. For $j=7$: two new edges cannot be provided by any simple projection from C, but Γ is not separated by a $\langle 4 \rangle$, $\langle 4- \rangle$, or $\langle 4= \rangle$, hence C is joined to at least five vertices of Γ_7 and these vertices necessarily span a graph containing a $\langle 5= \rangle$ as a subgraph. It may w.l.g. be assumed that C is joined to x_1, x_2, x_3, x_4, x_5. If $\Gamma_7 \supseteqq \langle 5 \rangle$, then $\Gamma \supseteqq \langle 6 \rangle$ and by (4. 1) $\Gamma \succ \langle 7= \rangle$, hence assume w.l.g. $(x_1, x_2) \notin E(\Gamma)$. If $(x_6, x_7) \notin E(\Gamma)$ then by (4. 6) x_6 is joined to four of $x_1, x_2, \dots, x_5$; let P denote the projection from C onto Γ_7 obtained by contracting $C \cup x_1$ into one vertex; $P\Gamma_7 \supseteqq \langle 6= \rangle$, hence $\Gamma \succ \langle 7= \rangle$. If $(x_6, x_7) \in E(\Gamma)$, then x_1 and x_2 are by (4. 6) both joined to $\Gamma(x_6, x_7)$. By contracting each of C and $\Gamma(x_1, x_6, x_7)$ into one vertex $\Gamma - x_0$ is contracted into a graph containing a $\langle 6- \rangle$ as a subgraph, five vertices of which are joined to x_0, hence $\Gamma \succ \langle 7= \rangle$.

This completes the proof of (4).

(3) and (4) exhaust all possibilities, hence Theorem 1 is proved.

5. An Application to the Conjecture of Hadwiger

We now turn to the application of Theorem 1 to the Conjecture of Hadwiger. First we prove a lemma:

LEMMA 2. *Let Γ be a 5-fold connected graph and let a, b, c, d, e, f, g be seven different vertices of Γ. If $\Gamma(a, b, c)=\langle 3\rangle$ and $\Gamma(d, e, f, g)=\langle 4\rangle$, then Γ can be contracted into a $\langle 6=\rangle$ so that six of the vertices a, b, c, d, e, f, g are mapped by the contraction into different vertices of the $\langle 6=\rangle$.*

PROOF. By (C) there exist four $(\Gamma(a, b, c))(\Gamma(d, e, f, g))$-paths with nothing in common in $\Gamma-a-b-c-d-e-f-g$ and such that two of them have b as an end-vertex, one has a, one has c, and each of d, e, f, g is an end-vertex of exactly one of the paths. The notation may be chosen such that the four paths are an $(a)(g)$-path Π_1, a $(b)(f)$-path Π_2, a $(b)(e)$-path Π_3 and a $(c)(d)$-path Π_4.

$\Gamma-a-c-e-f$ is connected, hence contains a $((\Pi_1-a)\cup(\Pi_4-c))((\Pi_2-f)\cup \cup(\Pi_3-e))$-path Π. Assume w.l.g. that Π is a $(\Pi_1-a)(\Pi_2-f)$-path with end-vertices $z\in\Pi_1-a$, $y\in\Pi_2-f$.

By contracting each of $\Pi_1[z, g]$, $\Pi_1[a, z]-a$, $\Pi_2[b, y]$, $\Pi_2[y, f]-f$, $\Pi-y$, Π_3-b, and Π_4 into one vertex Γ is contracted into a graph containing a $\langle 6=\rangle$ as a subgraph such that six of the vertices a, b, c, d, e, f, g are mapped into different vertices of the $\langle 6=\rangle$ by this contraction. This proves Lemma 2.

THEOREM 2. *Let Γ be a 7-chromatic graph. Then $\Gamma\succ\langle 7=\rangle$.*

PROOF. Because every 7-chromatic graph is homomorphic to a contraction-critical 7-chromatic graph (see (6) p. 50 in [3]), it is sufficient to prove the theorem for contraction-critical 7-chromatic graphs. Let Γ be a contraction-critical 7-chromatic graph.

$n(\Gamma)\geqq 7$. For $\Gamma=\langle 7\rangle$ the theorem is trivial, assume therefore that $\Gamma\neq\langle 7\rangle$. By (D). 2. Γ is 7-fold connected. Then each vertex of Γ has valency $\geqq 7$.

Assume that each vertex of Γ has valency $\geqq 8$. Then $e(\Gamma)\geqq 4n(\Gamma)>4n(\Gamma)-9$, $n(\Gamma)>7$. By (A) and Theorem 1 $\Gamma\succ\langle 7=\rangle$ in this case.

Hence it may be assumed that there exists a vertex x_0 in Γ of valency 7. Let the vertices joined to x_0 be denoted by $x_1, x_2, \ldots, x_7$, and let $\Gamma(x_1, \ldots, x_7)$ be denoted by Γ_7.

Assume now

(1) $$\exists x_i\in V(\Gamma_7)\colon v(x_i, \Gamma_7)\leqq 2.$$

Assume w.l.g. that $i=1$ and x_1 is not joined to x_2, x_3, x_4, x_5. By (D). 1. $\Gamma(x_2, x_3, x_4, x_5)=\langle 4\rangle$. If x_1 is not joined by an edge to x_6 or to x_7 then $\Gamma(x_2, x_3, x_4, x_5, x_j)=\langle 5\rangle$, $j=6$ or 7 respectively, by the same theorem, so $\Gamma\supseteqq\langle 6\rangle$. By (C) x_1 is joined by paths having only x_1 in common to all six vertices of this $\langle 6\rangle$, hence $\Gamma\succ\langle 7\rangle$ in this case. Assume then that x_1 is joined by edges to both x_6 and x_7. If $(x_6, x_7)\in E(\Gamma)$ it follows from Lemma 2 that $\Gamma-x_0$, which is 6-fold connected, is homomorphic to a $\langle 6=\rangle$ all the vertices of which are joined to x_0, therefore $\Gamma\succ\langle 7=\rangle$ in this case. Hence it may be assumed that $(x_6, x_7)\notin E(\Gamma)$. One of x_6, x_7 is joined to at least 2 of the vertices x_2, x_3, x_4, x_5, otherwise there would be 3 independent vertices in Γ_7 contrary to (D). 1. Assume w.l.g. that x_6 is joined by edges to x_4 and x_5. $\Gamma-x_0-x_1-x_6$ is 4-fold connected, hence by (C) contains four paths $\Pi_1, \Pi_2, \Pi_3, \Pi_4$ with only x_7 in common from x_7 to each of the vertices x_2, x_3, x_4, x_5. By contracting each of $\Gamma(x_1, x_6), \Pi_j - x_7, j = 1, 2, 3, 4$, into one vertex $\Gamma - x_0$ is contracted into a graph containing a $\langle 6 = \rangle$ as a subgraph all the vertices of which are joined to x_0, hence $\Gamma \succ \langle 7 = \rangle$. So in case (1) $\Gamma \succ \langle 7 = \rangle$.

Assume next:

$$(2) \qquad \forall x_k \in V(\Gamma_7): v(x_k, \Gamma_7) \geqq 4.$$

If no vertex of Γ_7 has valency exactly 4 in Γ_7, then $e(\Gamma_7) \geqq \frac{1}{2} 5 \cdot 7 > 3 \cdot 7 - 5$; hence by (A) and (B) $\Gamma_7 \succ \langle 6 = \rangle$ and consequently $\Gamma \succ \langle 7 = \rangle$. Assume then that there exists a vertex $x_i \in V(\Gamma_7)$: $v(x_i, \Gamma_7) = 4$. Assume w.l.g. that $i = 1$ and x_1 is joined to x_2, x_3, x_4, x_5, but not to x_6, x_7. The minimal valency in Γ is 7, but $v(x_1, \Gamma_7 \cup x_0) = 5$, hence $\Gamma - \Gamma_7 - x_0 \neq \emptyset$. Let C be a connected component of $\Gamma - \Gamma_7 - x_0$. Γ is 7-fold connected hence C is joined to all vertices of Γ_7. Let P be the projection from C onto Γ_7 defined by contracting $C \cup x_1$ into one vertex. $P\Gamma_7 = \Gamma_7 \cup (x_1, x_6) \cup (x_1, x_7)$. $e(P\Gamma_7) \geqq \frac{1}{2} 4 \cdot 7 + 2 = 16 = 3 \cdot 7 - 5$; no member of $\mathscr{K}_5^2$ has seven vertices, hence by (B) $P\Gamma_7 \succ \langle 6 = \rangle$ and consequently $\Gamma \succ \langle 7 = \rangle$. So in case (2) $\Gamma \succ \langle 7 = \rangle$.

In view of (1) and (2) the only case left to consider is

$$(3) \qquad \forall x_k \in V(\Gamma_7): v(x_k, \Gamma_7) \geqq 3 \quad \text{and} \quad \exists x_i \in V(\Gamma_7): v(x_i, \Gamma_7) = 3.$$

First it will be proved that Γ_7 contains two disjoint triangles.

Assume w.l.g. that $i = 1$ and x_1 is joined to x_2, x_3, x_4, but not to x_5, x_6, x_7. By (D). 1. $\Gamma(x_5, x_6, x_7) = \langle 3 \rangle$ and x_2, x_3, x_4 cannot be independent, hence it may w.l.g. be assumed that $(x_2, x_3) \in E(\Gamma)$. Then Γ_7 contains the two disjoint triangles $\Gamma(x_1, x_2, x_3)$ and $\Gamma(x_5, x_6, x_7)$.

Assume now w.l.g. that the two disjoint triangles are $\Gamma(x_1, x_2, x_3)$ and $\Gamma(x_5, x_6, x_7)$ (and $i = 1$ need not be the case). Then x_4 is joined to two vertices in at least one of the two disjoint triangles, because $v(x_4, \Gamma_7) \geqq 3$. Assume w.l.g. that x_4 is joined to x_1 and x_2. If $(x_4, x_3) \in E(\Gamma)$ as well then it follows from Lemma 2 that $\Gamma - x_0$ is homomorphic to a $\langle 6 = \rangle$ all the vertices of which are joined to x_0, hence $\Gamma \succ \langle 7 = \rangle$ in this case. Therefore assume that $(x_4, x_3) \notin E(\Gamma)$. Then x_4 is joined to at least one of x_5, x_6, x_7, assume w.l.g. to x_7.

If x_3 is joined to neither x_5 nor x_6, then by (D). 1. x_4 is joined to both x_5 and x_6 and by Lemma 2 $\Gamma \succ \langle 7 = \rangle$ in this case. Assume therefore w.l.g. that $(x_3, x_5) \in E(\Gamma)$.

By (D). 1. either $(x_3, x_6) \in E(\Gamma)$ or $(x_4, x_6) \in E(\Gamma)$. By the symmetry between x_3 and x_4 it may be assumed w.l.g. that $(x_4, x_6) \in E(\Gamma)$.

$\Gamma - x_0 - x_3 - x_4 - x_5 - x_6$ is 2-fold connected, hence by (C) it contains an $(x_1)(x_7)$-path Π_1 and an $(x_2)(x_7)$-path Π_2 such that Π_1 and Π_2 have only x_7 in common.

$\Gamma - x_0 - x_3 - x_4 - x_5 - x_7$ is 2-fold connected, hence contains an $(x_6)(\Pi_1 \cup \Pi_2 - x_7)$-path Π. Assume w.l.g. Π has end-vertex z on $\Pi_2 - x_7$. By contracting each of $\Pi_1 - x_1$, $\Pi_2 - x_7$, $\Pi - z$, and $\Gamma(x_3, x_5)$ into one vertex $\Gamma - x_0$ is contracted into a graph containing a $\langle 6 = \rangle$ as a subgraph all the vertices of which are joined to x_0, hence $\Gamma \succ \langle 7 = \rangle$.

This completes the proof of Theorem 2.

REFERENCES

[1] DIRAC, G. A.: Extensions of Menger's Theorem, *Journal. Lond. Math. Soc.* **38** (1963), 148—161.
[2] DIRAC, G. A.: Homomorphism theorems for graphs, *Math. Ann.* **153** (1964), 69—80.
[3] DIRAC, G. A.: On the structure of 5- and 6-chromatic abstract graphs, *Journal f. d. reine u. angew. Math.* **214/215** (1964), 43—52.
[4] HADWIGER, H.: Über eine Klassifikation der Streckenkomplexe, *Vierteljahrschr. Naturforsch. Ges. Zürich* **88**, (1943), 133—142.
[5] MADER, W.: Über trennende Eckenmengen in homomorphiekritischen Graphen, *Math. Ann.* **175**, (1968), 243—252.
[6] MADER, W.: Homomorphiesätze für Graphen, *Math. Ann.* **178**, (1968), 154—168.
[7] WAGNER, K.: Bemerkungen zu Hadwiger's Vermutung, *Math. Ann.* **141** (1960), 433—451.

Matematisk Institut, Aarhus Universitet, Ny Munkegade, 8000 Aarhus C, Denmark

(Received December 1, 1969. Revised October 5, 1970.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 161—162.

EIGENVALUES OF POWERS OF FUNCTIONS

by

E. MAKAI

Let us consider the two eigenvalue problems

$$y''+\lambda r(x)y = 0, \quad y(0)=y(T)=0 \tag{1}$$

and

$$y''+\lambda[r(x)]^q y = 0, \quad y(0)=y(T)=0. \tag{2}$$

We denote by λ_1 and λ_q, respectively, the smallest eigenvalues of these two problems. A. M. FINK [2] has shown recently that if λ_1 is positive, then $\lambda_2 \leqq (T\lambda_1/\pi)^2$.

If we restrict ourselves to the class of those functions $r(x)$ which are non-negative in $(0, T)$ we can prove somewhat more. Indeed we have the following

THEOREM. *If* $q \geqq 1$, $r(x) \geqq 0$ *in* $(0, T)$ *then the smallest eigenvalue* λ_q *of* (2) *satisfies* $\lambda_q \leqq (T/\pi)^{2(q-1)}\lambda_1^q$ *with equality only if* $r(x)=\text{const}$.

Let namely y_q be the eigenfunction of (2) belonging to the smallest eigenvalue λ_q and A the class of differentiable functions vanishing in 0 and T. Then we have

$$\lambda_q = \min_{y\in A} \int_0^T y'^2\,dx \Big/ \int_0^T r^q y^2\,dx = \int_0^T y_q'^2\,dx \Big/ \int_0^T r^q y_q^2\,dx.$$

Using in turn HÖLDER's and WIRTINGER's [3, p. 184] inequalities we have if $p^{-1}+q^{-1} = 1$,

$$\lambda_1 = \frac{\int_0^T y_1'^2\,dx}{\int_0^T y_1^{\frac{2}{p}}\, r y_1^{\frac{2}{q}}\,dx} \geqq \frac{\left(\int_0^T y_1'^2\,dx\right)^{\frac{1}{p}}\left(\int_0^T y_1'^2\,dx\right)^{\frac{1}{q}}}{\left(\int_0^T y_1^2\,dx\right)^{\frac{1}{p}}\left(\int_0^T r^q y_1^2\,dx\right)^{\frac{1}{q}}} =$$

$$= \left(\frac{\int_0^T y_1'^2\,dx}{\int_0^T y_1^2\,dx}\right)^{\frac{1}{p}}\left(\frac{\int_0^T y_1'^2\,dx}{\int_0^T r^q y_1^2\,dx}\right)^{\frac{1}{q}} \geqq \left(\frac{\pi^2}{T^2}\right)^{\frac{1}{p}}\left(\frac{\int_0^T y_q'^2\,dx}{\int_0^T r^q y_q^2\,dx}\right)^{\frac{1}{q}} = \left(\frac{\pi}{T}\right)^{2\left(1-\frac{1}{q}\right)}\lambda_q^{\frac{1}{q}}.$$

Equality holds here everywhere if and only if $r=\text{const}$.

With the help of the variational principle combined with Hölder's inequality one is able to treat analogous multidimensional eigenvalue problems, too. Let us consider e.g. the equation

$$\Delta u+\lambda r^q u = 0$$

of an inhomogeneous membrane, where r is a non-negative function in a domain D, q a constant not less than 1 and u satisfies a condition $\partial u/\partial n+\sigma u = 0$ $(\sigma \geqq 0)$ on the boundary B of D. If λ_q is the first eigenvalue of this problem and λ_0 is the first eigenvalue of a homogeneous membrane $(r\equiv 1)$ with the same boundary conditions, then one has (cf. [1], p. 210)

$$\lambda_q = \min_u \frac{\iint\limits_D (\nabla u)^2\, dA+\int\limits_B \sigma u^2\, ds}{\iint\limits_D r^q u^2\, dA}$$

and in the same way as above

$$\lambda_q \leqq \lambda_0^{q-1}\lambda_1^q .$$

REFERENCES

[1] Courant, R. and Hilbert, D.: *Methods of Mathematical Physics*, vol. I, Interscience, New York—London 1953.
[2] Fink, A. M.: Eigenvalue of the square of a function, *Proc. Amer. Math. Soc.* **20** (1969), 73—74.
[3] Hardy, G. A., Littlewood, J. E. and Pólya, G.: *Inequalities*, 2nd ed., Cambridge Univ. Press, 1952.

Mathematical Institute of the Hungarian Academy of Sciences

(Received December 10, 1969.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 163—166.

ON BOUNDEDNESS PROPERTY OF THE SOLUTION OF CERTAIN NONLINEAR DIFFERENTIAL EQUATIONS

by

M. BEHZAD and B. MEHRI

I. Abstract. In this note we consider second order nonlinear differential equations of the forms:

(1) $$x''+a(t)x'+b(t)x+g(t)f(x^2) = ke(t),$$

and

(2) $$x''+a(t)x'+b(t)x+g(t)f(x) = ke(t).$$

Under certain conditions imposed on the functions involved we prove that every solution of (1) and every solution of (2) is bounded as $t\to\infty$.

II. Results. First we consider the differential equation (1) and prove the following theorem.

THEOREM 1. *Assume that a, b, g, and e are continuous functions of t and k is a constant. Furthermore, assume that:*

(i) *b is non-negative, nondecreasing, continuously differentiable, and* $b(0)\neq 0$;

(ii) $\int\limits_0^t \frac{e^2(t)}{\sqrt{b(t)}} \leqq M$ *for all* $t\in[0,\infty]$,

(iii) $\int\limits_0^\infty |a(t)|\,dt$, $\int\limits_0^\infty \frac{|g(t)|}{\sqrt{b(t)}}\,dt$, *and* $\int\limits_0^\infty \frac{1}{\sqrt{b(t)}}$ *are*

convergent;

(iv) *f is a continuous nondecreasing function of its argument, and* $\int\limits_0^\infty \frac{du}{u^3+f(u^2)}=\infty$.

Then every solution of (1) *is bounded as* $t\to\infty$.

PROOF. Let

(3) $$A(t) = \frac{x'^2}{b}+x^2, \qquad x\neq 0$$

we have:

$$A(t) = \frac{x'^2(0)}{b(0)}+x^2(0)+\int\limits_0^t \frac{d}{dt}\left(\frac{x'^2}{b}+x^2\right)dt.$$

Assuming $C = \dfrac{x'^2(0)}{b(0)} + x^2(0)$, we have:

$$A(t) = c + \int_0^t \left(2xx' + \frac{2x'x''}{b} - \frac{x'^2}{b^2}\frac{db}{dt}\right) dt. \tag{4}$$

Using (1) and (i) we obtain:

$$A(t) \leqq c - \int_0^t \left(\frac{2a}{b} x'^2 + \frac{2x'}{b} gf(x^2) - \frac{2kx'}{b} e\right) dt. \tag{5}$$

Taking the absolute value of both sides of (5) and using the triangular inequality we obtain:

$$A(t) \leqq c + \int_0^t \left(\frac{2|a|x'^2}{b} + \frac{|g|}{\sqrt{b}} \cdot \frac{|2x'f(x^2)|}{\sqrt{b}} + \frac{|k|}{\sqrt{b}} \cdot \frac{2|x'e|}{\sqrt{b}}\right) dt.$$

Employing the inequality $2\alpha\beta \leqq \alpha^2 + \beta^2$ and the assumption (iv) we obtain:

$$A(t) \leqq c + \int_0^t \left(2|a| A(t) + \frac{|g|}{\sqrt{b}} A(t)\right) dt +$$

$$+ \int_0^t \left(\frac{|g|}{\sqrt{b}} f^2(A(t)) + \frac{|k|}{\sqrt{b}} A(t) + \frac{|k|}{\sqrt{b}} e^2\right) dt.$$

Next, we use (ii) to obtain:

$$A(t) \leqq c + |k| M + \int_0^t \left(2|a| + \frac{|g|}{\sqrt{b}} + \frac{|k|}{\sqrt{b}}\right) \left(A(t) + f^2(A(t))\right) dt. \tag{6}$$

Now suppose $D = C + |k|M$, and

$$E(t) = 2|a| + \frac{|g|}{\sqrt{b}} + \frac{|k|}{\sqrt{b}}.$$

From (6) we obtain:

$$A(t) \leqq D + \int_0^t E(t)\left(A(t) + f^2(A(t))\right) dt.$$

Assume $\omega(s) = s + f(s^2)$, and $\Omega(u) = \displaystyle\int_{u_0}^{u} \frac{ds}{\omega(s)}$ approaches ∞ as $u \to \infty$. It follows from (iv) that ω is a nondecreasing function of s. Hence by the well-known BELLMAN—BIHARI Lemma [2, 3] we have:

$$A(t) \leqq \Omega^{-1}\left(\Omega(D) + \int_0^t E(t)\, dt\right).$$

Now the assumption (iii) implies that $A(t)$ is bounded as $t \to \infty$; This completes the proof of the theorem.

It should be mentioned that if in (1) we substitute $k=0$ and $b(t)=0$, we obtain a result of BIHARI [1].

Next we consider the differential equation (2) and prove the following theorem,

THEOREM 2. *Assume that:*

(i) *a is a positive function,*

(ii) *b is a positive continuous nondecreasing function $t \geqq t_0 > 0$,*

(iii) *$\frac{e}{b}$ and $\frac{g}{b}$ are continuously differentiable for large t satisfying the inequalities*

$$\left|\frac{g}{b}\right| < \frac{\alpha}{t}, \quad \left|\frac{e}{b}\right| < \frac{\alpha_1}{t}, \quad \left|\left(\frac{g}{b}\right)'\right| < \frac{\alpha}{t^2}, \quad \text{and} \quad \left|\left(\frac{e}{b}\right)'\right| < \frac{\alpha_1}{t^2}$$

for t sufficiently large, and constants α and α_1.

(iv) *$D[f]$ is the set of all real numbers and $f \in \mathrm{Lip}\,(1)$*

(v) *Furthermore, assume the boundedness of $\frac{|F(x)|}{x^2}$ for all x, where $F(x) = \int_0^x f(s)\,ds$.*

Then every solution of (2) is bounded as $t \to \infty$.

PROOF. Let $A(t) = x^2 + \frac{x'^2}{b}$. Then, as we have seen in the proof of Theorem 1, we have:

$$A(t) \leqq A(t_0) + \int_{t_0}^{t} \left\{-2\frac{ax'^2}{b} - \frac{2x'f(x)}{b}g + \frac{2x'ke}{b}\right\} d\tau.$$

Since a is positive, we have:

$$A(t) \leqq A(t_0) + \int_{t_0}^{t} -2\frac{x'f(x)}{b}g\,d\tau + \int_{t_0}^{t} \frac{2x'ke}{b}\,d\tau. \tag{7}$$

In (7) we use integration by part to obtain:

$$A(t) \leqq K - 2F(x)\frac{g}{b} + 2kx\frac{e}{b} + 2\int_{t_0}^{t} F(x)\left(\frac{g}{b}\right)' d\tau - 2K\int_{t_0}^{t} x\left(\frac{e}{b}\right)' dt,$$

where

$$K = A(t_0) + 2F(x(t_0))\frac{g(t_0)}{b(t_0)} - 2kx(t_0)\frac{e(t_0)}{b(t_0)}.$$

Take the absolute values of both sides to obtain:

$$A(t) \leqq |K| + 2|F(x)|\left|\frac{g}{b}\right| + 2|k|\cdot|x|\cdot\left|\frac{e}{b}\right| +$$

$$+ 2\int_{t_0}^{t} |F(x)|\left|\left(\frac{g}{b}\right)'\right| d\tau + 2|k|\int_{t_0}^{t} |x|\left|\left(\frac{e}{b}\right)'\right| d\tau.$$

Suppose M is the maximum value of x on $[t_0, t]$ attained at μ, where $t_0 \leqq \mu \leqq t$. Then

$$M^2 \leqq A(\mu) \leqq |K| + 2|F(M)| \frac{\alpha}{\mu} + 2|k| M \frac{\alpha_1}{\mu} + \tag{8}$$
$$+ 2|F(M)| \int_{t_0}^{\mu} \frac{\alpha}{t^2}\, dt + 2|k|\,|M| \int_{t_0}^{\mu} \frac{\alpha_1}{t^2}\, dt.$$

Next, we find the values of the integrals in (8) to obtain:

$$M^2 \leqq |K| + 2|F(M)| \frac{\alpha}{t_0} + 2|k| M \frac{\alpha_1}{t_0}.$$

This inequality can be written in the form:

$$M^2 \leqq |K| + 2|F(M)| \frac{\alpha}{t_0} + \frac{|k|}{t_0^{1/2}} \left\{ 2M \frac{\alpha_1}{t_0^{1/2}} \right\}.$$

Since $2ab \leqq a^2 + b^2$, we have:

$$M^2 \leqq |K| + 2|F(M)| \frac{\alpha}{t_0} + \frac{|k|}{t_0^{1/2}} \left\{ M^2 + \frac{\alpha_1^2}{t_0} \right\}. \tag{9}$$

The inequality (9) reduces to:

$$M^2 \left\{ 1 - \frac{|k|}{t_0^{1/2}} - 2 \frac{|F(M)|}{M^2} \frac{\alpha}{t_0} \right\} \leqq |K| + \frac{|k|}{t_0^{3/2}} \alpha_1^2.$$

We can select t_0 in such a way that

$$\left\{ 1 - \frac{|k|}{t_0^{1/2}} - 2 \frac{F(M)}{M^2} \frac{\alpha}{t_0} \right\} \geqq \frac{1}{2}.$$

Hence with this value of t_0 we have:

$$\frac{M^2}{2} \leqq M^2 \left\{ 1 - \frac{|k|}{t_0^{1/2}} - 2 \frac{F(M)}{M^2} \frac{\alpha}{t_0} \right\} \leqq |K| + \frac{|k|}{t_0^{3/2}} \alpha_1^2. \tag{10}$$

Now, since the right-hand side of (10) is bounded, M is bounded as well. This in turn implies that x is bounded as $t \to \infty$. This completes the proof of the theorem.

Again Theorem 2 is a generalization of another result of BIHARI [1], for $a(t) = 0$ and $k = 0$ yields that result.

REFERENCES

[1] BIHARI, I.: On the Nonlinear Differential Equation $u'' + a(t)u + q(t) f(u^2) = 0$, *Acta Math. Acad. Sci. Hung.*, **6** (1961), 287—290.

[2] BIHARI, I.: Researches of the Boundedness and Stability of the Solutions of Nonlinear Differential Equations, *Acta Math. Acad. Sci. Hung.*, **8**, (1957), 278—291.

[3] BELLMAN, R.: *Stability Theory of Differential Equations*, McGraw-Hill, New-York, 1953.

Pahlavi University, Arya-Mehr University

(Received January 6, 1970. Revised March 17, 1970.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 167—168.

ОБ ОДНОРОДНЫХ ГАУССОВСКИХ МАРКОВСКИХ ПРОЦЕССАХ

А. КРАМЛИ

В настоящей статье доказывается следующая теорема:

Теорема:

Единственными однородными вероятностными плотностями, принадлежащими непрерывному гауссовскому марковскому процессу являются те, которые описывают решения стохастического дифференциального уравнения $d\xi(t) = -A\xi\, dt + M\, dt + dw(t)$ $\big(w(t)$ — стандартный винеровский процесс.$\big)$

Несмотря на то, что доказательство опирается на простые расчеты, оно не встречается в известной литературе.

Проблема возникла по вопросу Ю. А. Розанова.

Доказательство: Известно, что переходная вероятностная плотность имеет вид

$$p(y, t|x, s) = \frac{1}{\sqrt{2\pi(1-\varrho^2)\sigma^2(t)}} \exp\left[-\frac{\left(y-m(t)-\varrho\frac{\sigma(t)}{\sigma(s)}(x-m(s))\right)^2}{2(1-\varrho^2)\sigma^2(t)}\right].$$

Здесь предполагается, что $t \geqq s$; $m(s)$ и $\sigma(s)$ математическое ожидание и дисперсия (произвольные непрерывные функции от времени) а $\varrho(s, t)$ коэффициент корреляции имеет вид: $\varrho(s, t) = e^{F(s)-F(t)}$ $\big(F(s)$ неубывающая функция$\big)$. Так как $p(y, t|x, s)$ может рассматриваться как вероятностная плотность гауссовской случайной величины имеющей математическое ожидание $M_1(t, s)x + {} + M_2(t, s) = \varrho\frac{\sigma(t)}{\sigma(s)}x + m(t) - \varrho\frac{\sigma(t)}{\sigma(s)}m(s)$ и дисперсию $\sqrt{\Sigma(t, s)} = \sqrt{(1-\varrho^2)\sigma^2(t)}$, необходимое и достаточное условие однородности — зависимость функций $M_1(t, s)$, $M_2(t, s)$ и $\Sigma(t, s)$ только от $(t-s)$.

Рассмотрим отдельно вышеуказанные функции:

$$M_1(t-s) = e^{F(s)-F(t)}\frac{\sigma(t)}{\sigma(s)} = \frac{e^{-F(t)}\sigma(t)}{e^{-F(s)}\sigma(s)}.$$

Этому функциональному уравнению — среди непрерывных функций — удовлетворяет только функция формы

$$e^{-F(t)}\sigma(t) = Ce^{-At} \quad \text{т. е.} \quad M_1(t-s) = e^{-A(t-s)}$$

(A некоторая постоянная). Имея это в виду, получаем

$$\text{(I)} \qquad \sum(t-s) = \sigma^2(t) - e^{-2A(t-s)}\sigma^2(s).$$

Переписывая правую часть оно переходит в

$$\sum(t-s) = e^{-2At}\big(\sigma^2(t)e^{2At} - \sigma^2(s)e^{2As}\big).$$

Обозначим $\sigma^2(t)e^{2At}$ через $G(t)$.

Докажем, что предел $\lim\limits_{\tau\searrow 0}\dfrac{\Sigma(\tau)}{\tau}$ существует и так $G(t)$ дифференцируема и $G'(t)=e^{2At}\left(\lim\limits_{\tau\searrow 0}\dfrac{\Sigma(\tau)}{\tau}\right)$.

Существует такая последовательность $\{\tau^n\}$, что $\dfrac{\Sigma(\tau^n)}{\tau^n}\to B$ и предел конечен. (Противный случай вел бы к нелепости.) Если τ^n достаточно мало, из непрерывности функции e^{-2At} вытекает

$$\sum(\tau^n)=e^{-2A(s+\tau^n)}[G(s+\tau^n)-G(s)]\leqq\left(1+\frac{\varepsilon}{4}\right)\sum_{k=0}^{N-1}e^{-2A\left(s+\frac{k+1}{N}\tau^n\right)}\cdot$$

$$\cdot\left(G\left(s+\frac{k+1}{N}\tau^n\right)-G\left(s+\frac{k}{N}\tau^n\right)\right)\leqq\left(1+\frac{\varepsilon}{2}\right)N\sum\left(\frac{\tau^n}{N}\right)$$

Таким же образом получается неравенство

$$\frac{\sum(\tau^n)}{\tau^n}\geqq\left(1-\frac{\varepsilon}{2}\right)\frac{N\left(\sum\left(\frac{\tau^n}{N}\right)\right)}{\tau^n}.$$

По непрерывности функции $\Sigma(\tau)$ при достаточно малом δ мы можем установить

$$(1-\varepsilon)\frac{\sum(\tau)}{\tau}\leqq\frac{\sum(\tau^n)}{\tau^n}\leqq(1+\varepsilon)\frac{\sum(\tau)}{\tau}$$

при $\tau<\delta$.

Так как ε произвольное и $\dfrac{\Sigma(\tau^n)}{\tau^n}\to B,\ \lim\limits_{\tau\searrow 0}\dfrac{\Sigma(\tau)}{\tau}=B.$

Простой расчет показывает, что $G(t)=be^{2At}+C$ если $A\neq 0$ и $Bt+C$ если $A=0$

$\left(\text{где } b=-\dfrac{B}{2A}\right)$ и

$$\sum(t-s)=\begin{cases}b-be^{-2A(t-s)} & \text{при}\quad A\neq 0\\ B(t-s) & \text{при}\quad A=0\end{cases}$$

Аналогичные рассуждения относятся к функции $M_2(t,s)$:

$$M_2(t-s)=m(t)-e^{-A(t-s)}m(s)$$

$$M_2(t-s)=\begin{cases}m-me^{-A(t-s)} & \text{при}\quad A\neq 0\\ M(t-s) & \text{при}\quad A=0\end{cases}$$

$\left(\text{где } m=\dfrac{-M}{A}\right)$.

Подведя итоги, случаи $A>0$, $A=0$ и $A<0$ принадлежат соответственно к стационарному, винеровскому и эксплозионному процессу; даже параметры A, B и M совпадают с параметрами стохастического дифференциального уравнения.

ЛИТЕРАТУРА

[1] Дуб, Дж. Л.: *Вероятностные процессы*, ИЛ, 1956. *Москва*

(Поступила 1-ого января 1970 г.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 169—179.

КВАЗИМАРКОВСКИЕ СЛУЧАЙНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ

Т. ГЕРГЕЙ, И. И. ЕЖОВ

1. Основные определения

Пусть $\{\xi_n\}_0^\infty$ — произвольная последовательность случайных величин со значениями из $I = \{E_0, E_1, \ldots, E_n, \ldots\}: \mathsf{P}\{\xi_n \in I; n=\overline{0, \infty}\} = 1$. Состояние E_i называется *марковским* (по отношению к $\{\xi_n\}_0^\infty$), если из того, что $\xi_n = E_i$ следует независимость $\{\xi_k\}_{n+1}^\infty$ от $\{\xi_k\}_0^{n-1}$ для любого натурального n. Более точно это означает следующее: для любых натуральных n, m $(n < m)$ и $E_{i_0}, E_{i_1}, \ldots, E_{i_m}$ $(E_{i_n} = E_i)$

$$\mathsf{P}\{\xi_k = E_{i_k}; k = \overline{0, m}\} = \mathsf{P}\{\xi_k = E_{i_k}; k = \overline{0, n}\} \mathsf{P}\{\xi_k = E_{i_k}; k = \overline{n+1, m}/\xi_n = E_i\} \quad (1)$$

Момент попадания $\{\xi_n\}_0^\infty$ в марковские состояния будем называть *марковскими моментами.* Множество всех марковских состояний обозначим через D и положим $I \setminus D = D'$. Состояния из D' будем иногда называть *немарковскими.* Случайную последовательность $\{\xi_n\}_0^\infty$ назовем к в а з и м а р к о в с к о й, если соответствующее ей множество марковских состояний D не пусто Если $D = \emptyset$, то последовательность $\{\xi_n\}_0^\infty$ назовем *антимарковской.* В том частном случае, когда $D = I$, $\{\xi_n\}_0^\infty$ — цепь Маркова (далее ЦМ). Верно, конечно, и обратное: если $\{\xi_n\}_0^\infty$ — ЦМ с фазовым пространством I, то $\{\xi_n\}_0^\infty$ — квазимарковская последовательность с $D = I$. Пусть $\{\xi_n\}_0^\infty$ — некоторая квазимарковская последовательность (сокращено, КМП) и $D(D')$ — множество ее марковских (немарковских) состояний. Введем обозначения:

$$\mathsf{P}\{\xi_k = E_{j_k}; k = \overline{0, \infty}\} = Q_{j_0, j_1, \ldots}(E_{jk} \in D', k = \overline{0, \infty}),$$

$$\mathsf{P}\{\xi_k = E_{j_k}, k = \overline{0, m}; \xi_{m+1} = E_i\} = Q_{j_0, j_1, \ldots, j_m}(i) \quad (E_{j_k} \in D', k = \overline{0, m}; E_i \in D),$$

$$\mathsf{P}\{\xi_{n+1} = E_j | \xi_n = E_i\} = R_{ij}^{[n]} \quad (E_i, E_j \in D),$$

$$\mathsf{P}\{\xi_{n+k} = E_{j_k}, k = \overline{1, m}; \xi_{n+m+1} = E_r | \xi_n = E_i\} = R_{ir}^{[n]}(j_1, \ldots, j_m),$$

$$(E_i, E_r \in D; E_{j_1}, \ldots, E_{j_m} \in D'). \quad (2)$$

$$\mathsf{P}\{\xi_{n+k} = E_{j_k}; k = \overline{1, \infty} | \xi_n = E_i\} = R_i^{[n]}(j_1, j_2, \ldots) \quad (E_i \in D, E_{j_k} \in D', k = \overline{1, \infty}),$$

$$\mathsf{P}\{\xi_k = E_{j_k}; k = \overline{0, m}\} = Q_{j_0, j_1, \ldots, j_m} \quad (E_{j_k} \in D', k = \overline{0, m});$$

$$\mathsf{P}\{\xi_{n+k} = E_{j_k}, k = \overline{1, m} | \xi_n = E_i\} = R_i^{[n]}(j_1, \ldots, j_m) \quad (E_i \in D, E_{j_1}, \ldots, E_{j_m} \in D').$$

В дальнейшем мы ограничимся рассмотрением только таких КМП, у которых с вероятностью 1 происходит бесконечное число попаданий в D. Для таких КМП всегда

$$Q_{j_0, j_1 \dots} = \lim_{m \to \infty} Q_{j_0, \dots, j_m} = 0$$

$$R_i^{[n]}(j_1, j_2, \dots) = \lim_{n \to \infty} R_i^{[n]}(j_1, j_2, \dots, j_m) = 0 \tag{3}$$

Свяжем с КМП $\{\xi_n\}_0^\infty$ случайную последовательность $\{\xi_n^*\}_0^\infty$, множеством состояний которой являются символы

$$(E_{j_1}), \quad (E_{j_1, j_2}), \dots, (E_{j_1}, \dots, E_{j_k}), \dots \quad (E_{j_k} \in D'; k = \overline{1, \infty})$$

$$[E_i], \quad [E_i; E_{j_1}], \dots, [E_i; E_{j_1}, \dots, E_{j_k}], \dots \quad (E_i \in D; E_{j_k} \in D', k = \overline{1, \infty}),$$

причем

$$\xi_n^* = (E_{j_0}, \dots, E_{j_n})$$

тогда и только тогда, когда

$$\xi_0 = E_{j_0}, \quad \xi_1 = E_{j_1}, \dots, \xi_n = E_{j_n},$$

а

$$\xi_n^* = [E_i; E_{j_1}, \dots, E_{j_k}] \qquad (0 \leqq k \leqq n),$$

тогда и только тогда, когда

$$\xi_{n-k} = E_i, \qquad \xi_{n-k+1} = E_{j_1}, \dots, \xi_n = E_{j_k}.$$

Согласно (1)—(3) $\{\xi_n^*\}_0^\infty$ — ЦМ (назовем ее цепью *Маркова, натянутой на* КМП $\{\xi_n\}_0^\infty$), для которой

$$\mathsf{P}\{\xi_0^* = (E_j)\} = \mathsf{P}\{\xi_0 = E_j\}, \qquad E_j \in D';$$

$$\mathsf{P}\{\xi_0^* = [E_i]\} = \mathsf{P}\{\xi_0 = E_i\}, \qquad E_i \in D;$$

$$\mathsf{P}\{\xi_{n+1}^* = (E_{j_0}, \dots, E_{j_{n+1}}) | \xi_n^* = (E_{j_0}, \dots, E_{j_n})\} = \frac{Q_{j_0, \dots, j_{n+1}}}{Q_{j_0, \dots, j_n}},$$

$$\mathsf{P}\{\xi_{n+1}^* = [E_i] | \xi_n^* = (E_{j_0, \dots,} E_{j_n})\} = \frac{Q_{j_0, \dots, j_n}(i)}{Q_{j_0, \dots, j_n}},$$

$$\mathsf{P}\{\xi_{n+1}^* = [E_j] | \xi_n^* = [E_i]\} = R_{ij}^{[n]}, \tag{4}$$

$$\mathsf{P}\{\xi_{n+1}^* = [E_i; E_{j_1}] | \xi_n^* = [E_i]\} = R_i^{[n]}(j_1) \qquad (E_{j_1} \in D'),$$

$$\mathsf{P}\{\xi_{n+1}^* = [E_i; E_{j_1}, \dots, E_{j_{k+1}}] | \xi_n^* = [E_i, E_{j_1}, \dots, E_{j_k}]\} = \frac{R_i^{[n-k]}(j_1, \dots, j_{k+1})}{R_i^{[n-k]}(j_1, \dots, j_k)},$$

$$\mathsf{P}\{\xi_{n+1}^* = [E_r] | \xi_n^* = E_i; E_{j_1, \dots,} E_{j_k}]\} = \frac{R_{ir}^{[n-k]}(j_1, \dots, j_k)}{R_i^{[n-k]}(j_1, \dots, j_k)}$$

Пусть

$$0 \leqq \tau_0 < \tau_1 < \dots < \tau_n < \dots$$

последовательность всех марковских моментов КМП $\{\xi_n\}_0^\infty$. Если $\xi_{\tau_n}=\hat{\xi}_n$, то двумерная последоательность $\{\hat{\xi}_n, \tau_n\}_0^\infty$ образует однородную цепь Маркова (ОЦМ), для которой

$$\mathsf{P}\{\hat{\xi}_0=E_i, \tau_0=0\}=P\{\xi_0=E_i\},$$

$$\mathsf{P}\{\hat{\xi}_0=E_i, \tau_0=k\}=\sum_{E_{j_0}, \ldots, E_{j_{k-1}} \in D'} Q_{j_0, \ldots, j_{k-1}}(i) \qquad (k \geqq 1), \tag{5}$$

$$\mathsf{P}\{\hat{\xi}_{n+1}=E_r, \tau_{n+1}=l/\hat{\xi}_n=E_i, \tau_n=m\}=\sum_{E_{j_{m+1}}, \ldots, E_{j_{l-1}} \in D'} R_{ir}^{[m]}(j_{m+1}, \ldots, j_{l-1})$$

ОЦМ $\{\hat{\xi}_n, \tau_n\}_0^\infty$ назовем *вложенной цепью Маркова* соответствующей КМП $\{\xi_n\}_0^\infty$.

2. Классификация состояния

$\{\xi_n\}_0^\infty$ будем называть *однородной квазимарковской последовательностью* (сокращенно, ОКМП), если $R_i^{[n]}(j_1, \ldots, j_k)$ и $R_{ir}^{[n]}(j_1, \ldots, j_k)$, фигурирующие в (2), не зависят от n. Из (4) следует, что если $\{\xi_n\}_0^\infty$ — ОКМП, то $\{\xi_n^*\}_0^\infty$ — ОЦМ. Более того, в рассматриваемом случае $\{\hat{\xi}_n\}_0^\infty$ также образует ОЦМ (будем называть ее в дальнейшем *урезанной вложенной цепью Маркова*, соответствующей ОКМП $\{\xi_n\}_0^\infty$). В самом деле, используя (5), имеем:

$$\mathsf{P}\{\hat{\xi}_0 = E_i\} = \mathsf{P}\{\xi_0 = E_i\}+\sum_{k=1}^{\infty} \sum_{E_{j_0}, \ldots, E_{j_{k-1}} \in D'} Q_{j_0, \ldots, j_{k-1}}(i);$$

$$\mathsf{P}\{\hat{\xi}_k=E_{i_k};\, k=\overline{0, n}\}=\sum_{0 \leqq l_0<\cdots<l_n<\infty} \mathsf{P}\{\hat{\xi}_k=E_{i_k}, \tau_k=l_k;\, k=\overline{0, n}\}=$$

$$=\sum_{0 \leqq l_0<l_1<\cdots<l_n<\infty} \mathsf{P}\{\hat{\xi}_0=E_{i_0}, \tau_0=l_0\} \prod_{k=1}^{n} \mathsf{P}\{\hat{\xi}_k=E_{i_k}, \tau_k=l_k/\hat{\xi}_{k-1}=E_{i_{k-1}}, \tau_{k-1}=l_{k-1}\}=$$

$$=\sum_{0 \leqq l_0<l_1<\cdots<l_n} \mathsf{P}\{\hat{\xi}_0=E_{i_0}, \tau_0=l_0\} \prod_{k=1}^{n} \sum_{E_{j_1}, \ldots, E_{j_{l_k-l_{k-1}-1}} \in D'} \cdot$$

$$\cdot R_{i_{k-1} i_k}(j_1, \ldots, j_{l_k-l_{k-1}-1}) = \mathsf{P}\{\hat{\xi}_0 = E_{i_0}\} \cdot$$

$$\cdot \sum_{m_1, \ldots, m_n} \prod_{k=1}^{n} \sum_{E_{j_1}, \ldots, E_{j_{m_k}} \in D'} R_{i_{k-1} i_k}(j_1, \ldots, j_m) =$$

$$=\mathsf{P}\{\hat{\xi}_0=E_{i_0}\} \prod_{k=1}^{n} \sum_{m_k=0}^{\infty} \sum_{E_{j_1}, \ldots, E_{j_{m_k}} \in D'} R_{i_{k-1} i_k}(j_1, \ldots, j_{m_k})$$

Этим доказано, что $\{\hat{\xi}_n\}$ — ОЦМ, для которой

$$\mathsf{P}\{\hat{\xi}_0=E_i\}=T_i,$$

$$\mathsf{P}\{\hat{\xi}_{n+1}=E_j/\hat{\xi}_n=E_i\}=T_{ij}, \tag{6}$$

где

$$T_i = \mathsf{P}\{\xi_0 = E_i\} + \sum_{k=1}^{\infty} \sum_{E_{j_0}, \dots, E_{j_{k-1}} \in D'} Q_{j_0, \dots, j_{k-1}}(i) \tag{7}$$

$$T_{ij} = \sum_{m=0}^{\infty} R_{ij}(m), \quad R_{ij}(m) = \sum_{E_{j_1}, \dots, E_{j_m} \in D'} R_{ij}(j_1, \dots, j_m)$$

В силу (3) $\sum_{E_i \in D} T_i = \sum_{E_j \in D} T_{ij} = 1$ (т. е. матрица $\|T_{ij}\|$ стохастическая).

Займемся теперь классификацией состояния ОЦМ $\{\xi_n^*\}_0^\infty$, натянутой на ОКМП $\{\xi_n\}_0^\infty$. При этом мы будем существенно использовать соответствующую классификацию состояний урезанной вложенной цепи Маркова $\{\hat{\xi}_n\}_0^\infty$ с матрицей переходных вероятностей $\|T_{ij}\|$. Пусть

$$D = \bigcup_{i=0}^{\infty} F_i,$$

где F_0 — подмножество всех невозвратных состояний, а $F_i (i \geqq 1)$ — замкнутый класс сообщающихся возвратных состояний (ясно, что $F_i \cap F_j = \emptyset$ если только $i \neq j$). Из определения ЦМ $\{\xi_n^*\}_0^\infty$ следует, что множество всех ее возможных состояний сосредоточено на тех и только тех символах

$$(E_{j_0}), \dots, (E_{j_0}, \dots, E_{j_k}), \dots, (E_{j_k} \in D', k \geqq 0),$$

$$[E_i], [E_i; E_{j_1}], \dots, [E_i; E_{j_1}, \dots, E_{j_k}], \dots, (E_i \in D; E_{j_k} \in D', k \geqq 1),$$

на которых

$$\mathsf{P}\{\xi_0 = E_{j_0}\} > 0, \quad Q_{j_0, \dots, j_k} > 0 \quad (E_{j_k} \in D', k \geqq 0),$$

$$R_i(j_1, \dots, j_k) > 0 \quad (E_{j_1}, \dots, E_{j_k} \in D'; k \geqq 1)$$

Так как ОКМП $\{\xi_n\}_0^\infty$ с вероятностью 1 принимает значения из D, то $\{(E_{j_0}, \dots, E_{j_k});\ k \geqq 0\} \subset F_0^*$, где F_0^* — подмножество всех невозвратных состояний ОЦМ $\{\xi_n^*\}_0^\infty$. Впрочем, то что каждое состояние вида $(E_{j_0}, \dots, E_{j_k})$ невозвратно, следует из того, что

$$\mathsf{P}\{\xi_m^* \neq (E_{j_0}, \dots, E_{j_k}), \quad m > n | \xi_n^* = (E_{j_0}, \dots, E_{j_k})\} = 1$$

Если $F_0 = D$ (т. е. все состояния ЦМ $\{\hat{\xi}_n\}_0^\infty$ невозвратные), то и F_0^* совпадает со множеством всех состояний ЦМ $\{\xi_n^*\}_0^\infty$. Действительно, если бы сообщались $[E_i; E_{j_1}, \dots, E_{j_k}]$ и $[E_r; E_{l_1}, \dots, E_{l_m}]$, то сообщались бы и $[E_i]$ с $[E_r]$, что противоречит тому, что $E_i, E_r \in F_0$. Значит либо $[E_i, E_{j_1}, \dots, E_{j_k}]$, $[E_r, E_{l_1}, \dots, E_{l_m}] \in F_0^*$, либо $[E_i; E_{j_1}, \dots, E_{j_r}] \in F_m^*$, $[E_r; E_{l_1}, \dots, E_{l_m}] \in F_n^*$, где F_m^* и F_n^* — разные замкнутые классы возвратных состояний ОЦМ $\{\xi_n^*\}_0^\infty$. Покажем, что последнее невозможно. В самом деле, из того что

$$[E_i; E_{j_1}, \dots, E_{j_r}] \in F_m^* \qquad ([E_r; E_{l_1}, \dots, E_{l_m}] \in F_n^*)$$

следует существование хотя бы одного состояния $[E_{i'}]$ $([E_{r'}])$ такого, что $[E_{i'}] \in F_m^* ([E_{r'}] \in F_n^*)$. Но тогда $[E_{i'}]$ и $[E_i]$ ($[E_{r'}]$ и $[E_r]$) сообщается, что противоречит тому, что E_i и E_r — невозвратные состояния ОЦМ $\{\hat{\xi}_n\}_0^\infty$.

Пусть теперь $F_0 \subset D$. Из рассуждений предыдущего обзаца следует, что если $E_i \in F_0$, то $[E_i; E_{j_1}, \ldots, E_{j_k}] \in F_0^*$ $(k \geqq 1)$.

Рассмотрим теперь тот случай, когда $E_i \in F_m$ $(m > 0)$. Если $F_m = \{E_i\}$ т. е. E_i — поглощающее состояние, то состояния

$$[E_i], \quad [E_i; E_{j_1}], \ldots, [E_i; E_{j_1}, \ldots, E_{j_k}], \ldots, [E_{j_r} \in D', r \geqq 1)$$

образуют замкнутый класс сообщающихся состояний ОЦМ. Так как

$$\mathsf{P}\{\xi_{n+k}^* = [E_i]/\xi_n^* = [E_i]\} = R_{ii}^{\langle k \rangle} = R_{ii}(k-1) + \sum_{m=1}^{k-1} R_{ii}(m-1) R_{ii}(k-m-1) + \\ + \cdots + \underbrace{R_{ii}(0) R_{ii}(0) \ldots R_{ii}(0)}_{k}$$

то состояние $[E_i]$ (а значит и все состояния $[E_i; E_{j_1}, \ldots, E_{j_k}]$ имеет период равный наибольшему общему делителю (н. о. д.) тех k, для которых $R_{ii}^{\langle k \rangle} > 0$ $(k \geqq 1)$. Нетрудно видеть, что

$$\sum_{k=1}^{\infty} R_{ii}^{\langle k \rangle} z^k = A_{ii}(z) = \frac{B_{ii}(z)}{1 - B_{ii}(z)}, \quad \text{где} \quad B_{ii}(z) = \sum_{k=1}^{\infty} R_{ii}(k-1) z^k$$

Если $E_i \in F_m$ $(m > 0)$ и F_m содержит более одного состояния, то все состояния

$$[E_r], \quad [E_r; E_{j_1}], \ldots, [E_r; E_{j_1}, \ldots, E_{j_k}], \ldots, (E_r \in F_m; E_{j_k} \in D', k \geqq 1)$$

образуют замкнутый класс F_m^* сообщающихся состояний ОЦМ $\{\xi_n^*\}_0^{\infty}$. Это очевидно. Исследуем теперь период всех состояний из класса F_m^*. Введем обозначения:

$$\mathsf{P}\{\xi_{m+n}^* = [E_j]/\xi_m^* = [E_i]\} = R_{ij}^{\langle n \rangle} \quad (E_i, E_j \in F_m)$$

Используя определение $R_{ij}(m)$ и формулу полной вероятности имеем:

$$R_{ij}^{\langle n \rangle} = R_{ij}(n-1) + \sum_{\substack{k_1+k_2=n \\ E_r \in F_m}} R_{ir}(k_1-1) R_{rj}(k_2-1) + \\ + \cdots + \sum_{E_{r_1}, \ldots, E_{r_{n-1}} \in F_m} R_{ir_1}(0) R_{r_1 r_2}(0) \ldots R_{r_{n-1}, j}(0) \tag{8}$$

Если положить $\sum_{n=1}^{\infty} R_{ij}^{\langle n \rangle} z^n = a_{ij}(z)$, $\sum_{n=1}^{\infty} R_{ij}(n-1) z^n = b_{ij}(z)$,

$$\|a_{ij}(z)\| = A(z), \quad \|b_{ij}(z)\| = B(z) \quad (E_i, E_j \in F_m),$$

то из (8) следует, что

$$A(z) = B(z)[I - B(z)]^{-1} \tag{9}$$

Пусть E_j — произвольный элемент из F_m. Так как все состояние из F_m^* сообщаются, то все они имеют один и тот же период. Этот период равен н. о. д. тех n, для которых $R_{jj}^{\langle n \rangle} > 0$. Вероятность $R_{jj}^{\langle n \rangle}$ могут быть определены из матричного соотношения (9) либо непосредственно (через $R_{ij}(m)$) из (8)

Из вышесказанного следует

Теорема 1. *Пусть* $\{\xi_n\}_0^\infty$ *— ОКМП, со значениями из* $I = \{E_0, E_1, \ldots, E_m, \}$ *и множеством марковских состояний* D;

$\{\xi_n^*\}_0^\infty$ *— ОЦМ, натянутая на* $\{\xi_n\}_0^\infty$, а $\{\hat{\xi}_n\}_0^\infty$ *— урезанная вложенная ОЦМ, соответствующая* $\{\xi_n\}_0^\infty$.

1. *Если* $D' = I \setminus D$ *и*

$$\mathsf{P}\{\xi_m = E_{j_m}, m = \overline{0, k}\} = Q_{j_0, \ldots, j_k} \quad (E_{j_m} \in D' \; m = \overline{0, k}),$$

$$\mathsf{P}\{\xi_{n+m} = E_{j_m}; m = \overline{1, k} | \xi_n = E_i\} = R_i(j_1, \ldots, j_k)(E_i \in D, E_{j_m} \in D'),$$

то фазовое пространство ОЦМ $\{\xi_n^*\}_0^\infty$ *состоит из тех и только из тех символов*

$$(E_{j_0}, E_{j_1}, \ldots, E_{j_k}) \quad (E_{j_r} \in D', r = \overline{0, k}, k = \overline{0, \infty}),$$

$$[E_i], \quad [E_i; E_{j_0}, \ldots, E_{j_k}] \quad (E_i \in D, E_{j_r} \in D', r = \overline{1, k}; k = \overline{1, \infty}),$$

из которых

$$Q_{j_0, \ldots, j_k} > 0, \; R_i(j_1, \ldots, j_k) > 0.$$

Переходные вероятности ОЦМ $\{\xi_n^*\}_0^\infty$ *и ее начальное распределение определяются соотношениями* (4);

2. *Фазовое пространство ОЦМ* $\{\hat{\xi}_n\}_0^\infty$ *совпадает с* D, *а ее начальное распределение и матрица переходных вероятностей имеют вид* $\{T_i\}$, $\|T_{ij}\|$, *где*

$$T_i = \mathsf{P}\{\xi_0 = E_i\} + \sum_{k=1}^{\infty} \sum_{E_{j_0}, \ldots, E_{j_{k-1}} \in D'} \mathsf{P}\{\xi_k = E_i, \xi_r = E_{jr}, r = \overline{0, k-1}\},$$

$$T_{ij} = \sum_{m=0}^{\infty} R_{ij}(m), \quad R_{ij}(m) = \sum_{E_{j_1}, \ldots, E_{j_m} \in D'} \mathsf{P}\{\xi_{n+m+1} = E_j; \xi_{n+r} = E_{j_r}, r = \overline{1, m} | \xi_n = E_i\};$$

3. *Если* F_0 *— множество всех невозвратимых состояний* $\{\hat{\xi}_n\}_0^\infty$, *а* F_m $(m > 0)$ *— некоторый замкнутый класс всех сообщающихся между собой состояний, то*

$$F_0^* = \{(E_{j_0}, \ldots, E_{j_k}), \quad [E_i], \quad [E_i, E_{j_1}, \ldots, E_{j_r}] : E_i \in F_0; E_{j_n} \in D', n \geqq 0\} -$$

множество всех невозвратимых состояний $\{\xi_n^*\}_0^\infty$, *а*

$$F_m^* = \{[E_i], [E_i, E_{j_1}, \ldots, E_{j_k}] : E_i \in F_m; E_{j_n} \in D', n \geqq 1\} -$$

замкнутый класс сообщающихся состояний этой же цепи. При этом если

$$D = \bigcup_{m=0}^{\infty} F_m \text{ и } I^* \text{ — фазовое пространство } \{\xi_n^*\}_0^\infty,$$

то

$$I^* = \bigcup_{m=0}^{\infty} F_m^*$$

4. *Период всех состояний из* $F_m^*(m>0)$ *равен н. о. д. тех* n, *для которых* $R_{jj}^{\langle n\rangle}>0$ (j *фиксировано так, что* $E_j \in F_m$), *где*

$$\sum_{n=1}^{\infty} R_{ij}^{\langle n\rangle} z^n = a_{ij}(z), \quad \|a_{ij}(z)\| = \|b_{ij}(z)\|\big(I-\|b_{ij}(z)\|\big)^{-1},$$

$$(i, j\colon E_i,\, E_j \in F_m)$$

а

$$b_{ij}(z) = \sum_{m=1}^{\infty} R_{ij}(m-1)\, z^m.$$

3. Эргодическая теорема

Займемся теперь выяснением тех условий, при выполнении которых все состояния из $F_m^*(m>0)$ имеют конечное среднее время возвращения. Если π_{ik} — математическое ожидание того промежутка времени, за которое ОЦМ $\{\xi_n^*\}_0^\infty$, выходя из $[E_i]$ впервые попадает в $[E_k]$ ($[E_i]$, $[E_k] \in F_m^*$), то

$$\pi_{ik} = \sum_{m=1}^{\infty} mR_i(m-1) + \sum_{r\neq k} T_{ir}\pi_{rk}, \tag{10}$$

где

$$R_i(m-1) = \sum_k R_{ik}(m-1).$$

Отметим, что $\sum\limits_{m=1}^{\infty} mR_i(m-1)$ совпадает со средним временем того промежутка времени, за который КМП $\{\xi_n\}_0^\infty$ выходя из E_i ($E_i \in D$), впервые попадает в D. Если это среднее обозначить через m_i, то согласно (10)

$$\pi_{ik} = m_i + \sum_{r\neq k} T_{ir}\tau_{rk} \qquad (E_i,\, E_k,\, E_r \in F_m) \tag{11}$$

Предположим, что система уравнений

$$\varrho_i = \sum_{E_k \in F_m} \varrho_k T_{ki} \qquad (E_i \in F_m) \tag{12}$$

обладает единственным (с точностью до нормирующего множителя) неотрицательным решением. Решение, удовлетворяющее условию $\sum\limits_{E_i \in F_m} \varrho_i = 1$ обозначим через $\{\varrho_i^*\}$. *

* Отметим, что если F_m имеет конечное число состояний, то существование и единственность $\{\varrho_i^*\}$ следует из теоремы Фробениуса о собственных числах и векторах неразложимой матрицы с неотрицательными элементами ([1], стр. 355). В общем случае теорема Фробениуса, к сожалению, не имеет места. Однако, если предположить, что все состояния из F_m ОЦМ $\{\hat{\xi}_n\}_0^\infty$ апериодичны и положительны (т. е. имеют конечное среднее время возвращения), то решение $\{\varrho_i^*\}$ существует, единственно, положительно и совпадает со стационарным распределением $\{\hat{\xi}_n\}_0^\infty$ при условии, что $\hat{\xi}_0 \in E_m$.

Умножая обе части (11) на ϱ_i^* и суммируя их по всем i $(E_i \in F_m)$ имеем:

$$\sum_{E_i \in F_m} \varrho_i^* \pi_{ik} = \sum_{E_i \in F_m} \varrho_i^* m_i + \sum_{r \neq k} \pi_{rk} \sum_{E_i \in F_m} \varrho_i^* T_{ik},$$

или

$$\sum_{E_i \in F_m} \varrho_i^* \pi_{ik} = \sum_{E_i \in F_m} \varrho_i^* m_i + \sum_{E_r \in F_m} \varrho_r^* \pi_{rk} - \varrho_k^* \pi_{kk},$$

откуда

$$\pi_{kk} = \frac{1}{\varrho_k^*} \sum_{E_i \in F_m} \varrho_i^* m_i \tag{13}$$

Из (13) следует, что для того, чтобы все состояния из F_m^* были положителными необходима и достаточна сходимость ряда $\sum\limits_{E_i \in F_m} \varrho_i^* m_i$;

Предполагая теперь, что все состояния из F_m^* апериодичны, заново выведем (13) и установим явный вид стационарного распределения ОЦМ $\{\xi_n^*\}_0^\infty$ при условии, что

$$\xi_0^* \in F_m^* \quad \text{и} \quad \sum_{E_i \in F_m} \varrho_i^* m_i < \infty$$

Введем обозначения:

$$\lim_{n \to \infty} \mathsf{P}\{\xi_n^* = [E_k]\} = p_k, \tag{14}$$

$$\lim_{n \to \infty} \mathsf{P}\{\xi_n^* = [E_k; E_{j_1}, \ldots, E_{j_r}]\} = p_k(j_1, \ldots, j_r)$$

$$(E_k \in F_m; E_{j_1}, \ldots, E_{j_r} \in D'; r \geqq 1)$$

Пределы (14) всегда существуют и либо все равны нулю, либо все строго положительны. Это зависит от того, как мы увидим, будет ряд $\sum\limits_{E_i \in F_m} \varrho_i^* m_i$ расходиться или сходиться. Используя переходные вероятности $\{\xi_n^*\}_0^\infty$ (4) имеем:

$$p_k(j_1, \ldots, j_r, j_{r+1}) = p_k(j_1, \ldots, j_r) \frac{R_k(j_1, \ldots, j_r, j_{r+1})}{R_k(j_1, \ldots, j_r)} \tag{15}$$

$$p_k = \sum_{E_i \in F_m} \sum_{r=0}^{\infty} \sum_{E_{j_1}, \ldots, E_{j_r} \in D'} p_i(j_1, \ldots, j_r) \frac{R_{ik}(j_1, \ldots, j_r)}{R_i(j_1, \ldots, j_r)}$$

Из первого равенства в (15) следует, что

$$p_k(j_1, \ldots, j_r) = p_k R_k(j_1, \ldots, j_r) \qquad (r \geqq 0) \tag{16}$$

Используя (16) и второе равенство в (15), имеем:

$$p_k = \sum_{E_i \in F_m} p_i \sum_{r=0}^{\infty} \sum_{E_{j_1}, \ldots, E_{j_r} \in D'} R_{ik}(j_1, \ldots, j_r),$$

или

$$p_k = \sum_{E_i \in F_m} p_i T_{ik} \tag{17}$$

Согласно предположениям относительно решения системы уравнений (12) имеем:

$$p_k = C\varrho_k^* \quad (E_k \in F_m) \tag{18}$$

C находится из условия нормировки:

$$1=\sum_{E_k\in F_m}\sum_{r=0}^{\infty}\sum_{E_{j_1},\dots,E_{j_r}\in D'}p_k(j_1,\dots,j_r)=C\sum_{E_k\in F_m}\varrho_k^*\sum_{r=0}^{\infty}\sum_{E_{j_1},\dots,E_{j_r}\in D'}R_k(j_1,\dots,j_r)$$

Так как

$$\sum_r\sum_{E_{j_1},\dots,E_{j_r}}R_k(j_1,\dots,j_r)=m_k$$

то

$$C^{-1}=\sum_{E_k\in F_m}\varrho_k^* m_k$$

откуда

$$p_k(j_1,\dots,j_r)=\frac{\varrho_k^* R_k(j_1,\dots,j_r)}{\sum\limits_{E_i\in\varrho_m}\varrho_i^* m_i} \tag{19}$$

Из (19) видно, что $p_k(j_1,\dots,j_r)$ положительно тогда и только тогда, когда ряд стоящий в знаменателе сходится. Так как $\pi_{kk}=p_{k]}^{-1}$([2] стр. 381) то (13) равносильно (19) при $r=0$.

Итак, нами доказана

Теорема 2. *Если* 1. F_m^* — *замкнутый класс сообщающихся апериодических состояний ОЦМ* $\{\xi_n^*\}_0^\infty$; 2. $\xi_0^*\in F_m^*$; 3. *система уравнений*:

$$\varrho_i=\sum_{E_k\in F_m}\varrho_k T_{ki}\quad(E_i\in F_m)$$

обладает единственным (с точностью до нормировки) неотрицательным решением, то

$$\lim_{n\to\infty}\mathsf{P}\{\zeta_n^*=[E_k;E_{j_1},\dots,E_{j_r}]\}=\frac{\varrho_k R_k(j_1,\dots,j_r)}{\sum\limits_{E_i\in F_m}\varrho_i m_i}\ *$$

где m_i — *среднее число шагов, за которое КМП* $\{\xi_n\}_0^\infty$, *выходя из* E_i *впервые попадает в* D.

4. Распределение моментов времени достижения заданной области фазового пространства

Пусть $\Gamma\subset I$ — произвольное подмножество состояний ОКМП $\{\xi_n\}_0^\infty$, а $\xi_0^*=[E_i;E_{j_1},\dots,E_{j_r}]$, где $\{\xi_n^*\}_0^\infty$ — соответствующая ОЦМ, натянутая на $\{\xi_n\}_0^\infty$. Нас интересует распределение того первого момента времени $t=\tau\{[E_i;E_{j_1},\dots,E_{j_r}],\Gamma\}$, для которого впервые $\xi_t\in\Gamma$. Очевидно, что $t=0$ тогда и только тогда, когда $E_{j_r}\in\Gamma$, ибо из условия $\xi_0^*=[E_i,E_{j_1},\dots,E_{j_r}]$ следует, что $\xi_0=E_{j_r}$.

* Если ряд $\sum\limits_{E_i\in F_m}\varrho_i m_i$ расходится, то предел следует считать равным нулю; в этом случае все состояние из F_m^* являются возвратно-нулевыми.

Введем обозначения:

$$
\begin{gathered}
Mz^{\tau\{[E_i;\, E_{j_1},\, \ldots,\, E_{j_r}],\, \Gamma\}} = S_\Gamma^*([E_i;\, E_{j_1},\, \ldots,\, E_{j_r}],\, z),\\
R_i(j_1,\, \ldots,\, j_r)\, S_\Gamma^*([E_i;\, E_{j_1},\, \ldots,\, E_{j_r}],\, z) = S_\Gamma([E_i;\, E_{j_1},\, \ldots,\, E_{j_r}],\, z),\\
S_\Gamma([E_i],\, z) = S_i(\Gamma,\, z), \quad \Gamma' = I \setminus \Gamma,\\
b_k([E_i;\, E_{j_1},\, \ldots\, E_{j_r}],\, \Gamma) = \sum_{\substack{E_{j_{r+1}},\, \ldots,\, E_{j_{r+k-1}} \in D' \cap \Gamma' \\ E_{j_{r+k}} \in \Gamma}} R_i(j_1,\, j_2,\, \ldots,\, j_{r+k})
\end{gathered} \tag{20}
$$

Если $E_{j_r} \in \Gamma'$, то согласно (4)

$$
\tau\{[E_i;\, E_{j_1},\, \ldots,\, E_{jr}];\, \Gamma\} =
$$

$$
= 1 + \begin{cases}
0 & : \dfrac{b_1([E_1;\, E_{j_1},\, \ldots,\, E_{j_r}],\, \Gamma)}{R_i(j_1,\, \ldots,\, j_r)} \\[2ex]
\tau\{[E_i;\, E_{j_1},\, \ldots,\, E_{j_{r+1}}],\, \Gamma\} & : \dfrac{R_i(j_1,\, \ldots,\, j_{r+1})}{R_i(j_1,\, \ldots,\, j_r)}, \quad E_{j_{r+1}} \in \Gamma' \cap D' \\[2ex]
\tau\{[E_k],\, \Gamma\} & : \dfrac{R_{ik}(j_1,\, \ldots,\, j_r)}{R_i(j_1,\, \ldots,\, j_r)}, \quad E_k \in \Gamma' \cap D
\end{cases} \tag{21}
$$

Из стохастических соотношений (20) и (21) следует, что

$$
\begin{gathered}
S_\Gamma([E_i;\, E_{j_1},\, \ldots,\, E_{j_r}],\, z) = z b_1([E_i;\, E_{j_1},\, \ldots,\, E_{j_r}],\, \Gamma) + \\
+ z \sum_{E_k \in \Gamma' \cap D} R_i(j_1,\, \ldots,\, j_r,\, k)\, S_k(\Gamma,\, z) + z \sum_{E_{j_{r+1}} \in \Gamma' \cap D'} S_\Gamma([E_i;\, E_{j_1},\, \ldots,\, E_{j_{r+1}}],\, z)
\end{gathered} \tag{22}
$$

Система линейных уравнений (22) может быть преобразована так, чтобы правая часть в (22) содержала только неизвестные вида $S_k(\Gamma,\, z)$. Действительно, нетрудно видеть, что

$$
\tau\{[E_i;\, E_{j_1},\, \ldots,\, E_{j_r}],\, \Gamma\} =
$$

$$
= \begin{cases}
k & : \dfrac{b_k([E_i;\, E_{j_1},\, \ldots,\, E_{j_r}],\, \Gamma)}{R_i(j_1,\, \ldots,\, j_r)}; \quad k \geqq 1 \\[2ex]
k + \tau\{[E_l],\, \Gamma\} & : \dfrac{b_{kl}([E_i;\, E_{j_1},\, \ldots,\, E_{j_r}],\, \Gamma)}{R_i(j_1,\, \ldots,\, j_r)}; \quad k \geqq 1, \quad E_l \in \Gamma' \cap D,
\end{cases}
$$

где

$$
b_{kl}([E_i;\, E_{j_1},\, \ldots,\, E_{j_r}],\, \Gamma) = \sum_{E_{j_{r+1}},\, \ldots,\, E_{j_{r+k-1}} \in D' \cap \Gamma'} R_{il}(j_1,\, \ldots,\, j_{r+k-1})
$$

Поэтому

$$
\begin{gathered}
S_\Gamma([E_i;\, E_{j_1},\, \ldots,\, E_{j_r}],\, z) = \sum_{k=1}^{\infty} b_k([E_i;\, E_{j_1},\, \ldots,\, E_{j_r}],\, \Gamma) z^k + \\
+ \sum_{k=1}^{\infty} z^k \sum_{E_l \in \Gamma' \cap D} b_{kl}([E_i;\, E_{j_1},\, \ldots,\, E_{j_r}],\, \Gamma)\, S_l(\Gamma,\, z),
\end{gathered}
$$

или

$$S_\Gamma([E_i; E_{j_1}, \dots, E_{j_r}], z) = B_\Gamma([E_i; E_{j_1}, \dots, E_{j_r}], z) + \\ + \sum_{E_l \in \Gamma' \cap D} B_\Gamma^{(l)}([E_i; E_{j_1}, \dots, E_{j_r}], z) S_l(\Gamma, z), \tag{23}$$

где

$$B_\Gamma([E_i; E_{j_1}, \dots, E_{j_r}], z) = \sum_{k=1}^{\infty} b_k([E_i; E_{j_1}, \dots, E_{j_r}], \Gamma) z^k,$$

$$B_\Gamma^{(l)}([E_i; E_{j_1}, \dots, E_{j_r}], z) = \sum_{k=1}^{\infty} b_{kl}([E_i; E_{j_1}, \dots E_{j_r}], \Gamma) z^k.$$

Из (23) видно, что для определения $S_\Gamma([E_i; E_{j_1}, \dots, E_{j_r}], z)$ достаточно знать $S_l(\Gamma, z)$ $(E_l \in \Gamma' \cap D)$. Полагая в (23) $r=0$ получим следующую систему линейных алгебраических уравнений для определения $S_l(\Gamma, z)$:

$$S_i(\Gamma, z) = B_\Gamma([E_i], z) + \sum_{E_l \in \Gamma' \cap D} B_\Gamma^{(l)}([E_i], z) S_l(\Gamma, z) \tag{24}$$

$$(E_i \in \Gamma' \cap D).$$

(24) определяет $S_i(\Gamma, z)$ однозначно для всех z с $|z| \leqq 1$.

Система уравнений (24) является регулярной [3]. Поэтому ее решение существует, единственно и может быть найдено методом редукции [3].

ЛИТЕРАТУРА

[1] Гантмахер, Ф. Р.: *Теория матриц*, «Наука», Москва, 1967.
[2] Феллер, В: *Введение в теорию вероятностей и ее приложения*, „МИР", Москва, 1964.
[3] Канторович, Л. В., Крылов, В. И.: *Приближенные методы высшего анализа*, ГИТТП, Москва—Ленинград, 1950.

(Поступила 16-ого января 1970 г.)

ON THE ERROR EXPONENT FOR SOURCE CODING AND FOR TESTING SIMPLE STATISTICAL HYPOTHESES

by

I. CSISZÁR and G. LONGO[1]

Summary

The problems of estimating the optimum probability (i) of incorrect decoding when messages of length n from a discrete memoryless source are being encoded by means of 2^{nR} codewords and (ii) of the error of the second kind in testing simple alternative hypotheses when the probability of the error of the first kind decreases exponentially with a prescribed exponent, are very tightly connected. In this paper these problems are dealt with by the aid of a properly chosen auxiliary probability distribution[2]. This approach provides a simpler and more motivated derivation of the known results; moreover a sharp theorem of STRASSEN is used to obtain more accurate estimates.

Acknowledgement. G. Longo thanks for the support of Consiglio Nazionale delle Ricerche, Italy.

§ 1. Introduction

Let $X=\{x_1, \ldots, x_m\}$ be a finite set, $\mathscr{P}=\{p_1, \ldots, p_m\}$ a probability distribution on X and $a_1, \ldots, a_m$ positive numbers whatsoever. Let X^n denote the set of the n-length sequences, $u=x_{i_1}\ldots x_{i_n}$, of elments of X (for the sake of simplicity, commas will be omitted). For any $u=x_{i_1}\ldots x_{i_n}\in X^n$ we put:

$$p(u)=p_{i_1}\cdots p_{i_n}, \quad a(u)=a_{i_1}\cdots a_{i_n}, \tag{1}$$

and for any subset E of X^n:

$$p(E)=\sum_{u\in E} p(u), \quad a(E)=\sum_{u\in E} a(u). \tag{2}$$

We shall be interested in the limiting behaviour for $n\to\infty$ of the minimum of[3] $a(E)$ (or of $p(E^c)$) under the condition that the value of $p(E^c)$ (or of $a(E)$, respectively) is given (possibly depending on n). This interest is motivated by the following two problems.

(i) Let us be given a discrete memoryless source having alphabet X and probability distribution $\mathscr{P}$. Its messages of length n are required to be encoded into 2^{nR} codewords in such a way that the probability of incorrect decoding be as small

[1] This work was done while G. LONGO was a guest of the Mathematical Institute of the Hungarian Academy of Sciences. He is indebted to Prof. A. RÉNYI for helpful suggestions.

[2] A similar idea was used by RÉNYI [5] p. 473 in connection with the statistical interpretation of the entropy of order α and by ARUTJUNJAN [1] in connection with channel coding.

[3] Here E^c denotes the complement of the set E with respect to X^n.

as possible. The minimal probability of error $P_e = P_e(n, R)$ is obviously achieved if the 2^{nR} most probable sequences of length n are encoded into different codewords (and the others quite arbitrarily); then decoding consists in stating that each codeword comes from the most probable of the sequences of the form $u = x_{i_1} \dots x_{i_n}$ from which it could arise. In this case, taking $a_i = 1$ for $i = 1, \dots, m$, the probability of error equals the minimum of $p(E^c)$ under the condition $a(E) = 2^{nR}$. Conversely, we may also fix the error probability γ we are prepared to tolerate in decoding, and look for the minimum number of codewords, i.e. the minimum of $a(E)$ under the condition $p(E^c) \leqq \gamma$.

(ii) Suppose that $\mathcal{P} = \{p_1, \dots, p_m\}$ and $\mathcal{Q} = \{q_1, \dots, q_m\}$ are two a priori possible probability distributions on the set X; on the basis of an independent sample of size n, say $u = x_{i_1} \dots x_{i_n}$, a decision should be taken as to which distribution is the true one. If the sample belongs to a suitably chosen set $E \subset X^n$, then hypothesis $\mathcal{P}$ will be accepted, otherwise hypothesis $\mathcal{Q}$ will. Setting $a_i = q_i$ $(i = 1, \dots, m)$ the minimum of $a(E) = q(E)$ under the condition $p(E^c) \leqq \gamma$ is just the minimum probability of the error of the second kind under the condition that the probability of the error of the first kind is not greater than γ. It should be emphasized that γ need not be constant, as it is usually assumed; actually, we shall obtain an accurate estimate for the probability of the second kind error also in case the probability of the first kind error is required to vanish exponentially as $n \to \infty$, with a given positive exponent.

§ 2. Preliminary lemmas

Our starting point will be the following well-known elementary lemma (the notations of § 1 are used):

LEMMA 1. *Define:*

$$b(n, \gamma) = \min_{E \subset X^n,\ p(E^c) \leqq \gamma} a(E), \tag{3}$$

then

$$\lim_{n \to \infty} \frac{1}{n} \log_2 b(n, \gamma) = -\sum_{i=1}^{m} p_i \log_2 \frac{p_i}{a_i} \overset{\text{def}}{=} M. \tag{4}$$

PROOF. Let $F_n(\delta)$ denote the set of sequences $u = x_{i_1} \dots x_{i_n}$ satisfying the condition

$$\left| \log_2 \frac{p(u)}{a(u)} + nM \right| < n\delta,$$

where δ is any positive constant. Since

$$\sum_{u \in X^n} p(u) \log_2 \frac{p(u)}{a(u)} = -nM,$$

Chebyshev's inequality implies

$$\sum_{u \in F_n^c(d)} p(u) \leqq \frac{\sum_{u \in X^n} p(u) \left(\log_2 \frac{p(u)}{a(u)} + nM \right)^2}{(n\delta)^2} = \frac{n \sum_{i=1}^{m} p_i \left(\log_2 \frac{p_i}{a_i} + M \right)^2}{n^2 \delta^2} \tag{5}$$

so that

(6) $$p(F_n^c(\delta)) = \sum_{u\in F_n^c(\delta)} p(u) \to 0 \quad \text{as} \quad n\to\infty.$$

The definition of $F_n(\delta)$ entails:

(7) $$2^{n(-M-\delta)}a(u) \leqq p(u) \leqq 2^{n(-M+\delta)}a(u) \quad \text{if} \quad u\in F_n(\delta),$$

whence, on account of (6), for sufficiently large n

(8) $$b(n,\gamma) \leqq a(F_n(\delta)) = \sum_{u\in F_n(\delta)} a(u) \leqq 2^{n(M+\delta)} \sum_{u\in F_n(\delta)} p(u) \leqq 2^{n(M+\delta)}$$

On the other hand, for any $E\subset X^n$, (7) yields also

(9) $$a(E) \geqq a(E\cap F_n(\delta)) \geqq \sum_{u\in E\cap F_n(\delta)} 2^{n(M-\delta)}p(u) = 2^{n(M-\delta)}p(E\cap F_n(\delta)).$$

Now in view of (6), if $P(E^c)\leqq\gamma$ and n is large enough, we have $p(E\cap F_n(\delta)) \geqq \geqq \frac{1-\gamma}{2}$ and consequently (9) gives

(10) $$b(n,\gamma) = \inf_{E\subset X^n,\ p(E^c)\leqq\gamma} a(E) \geqq 2^{n(M-\delta)}\frac{1-\gamma}{2}.$$

Since $\delta>0$ is arbitrary, (8) and (10) yields (4).

We shall use also the following considerable sharpening of lemma 1, due to STRASSEN [6] (logarithms are taken to the base e):

LEMMA 2. *The exact asymptotic expression for $b(n,\gamma)$ is given by*

(11) $$\log b(n,\gamma) = nM + \sqrt{n}\,\lambda S - \frac{1}{2}\log n + \frac{T^3}{6S^2}(\lambda^2-1) - $$
$$-\frac{1}{2}\lambda^2 - \log(\sqrt{2\pi}\,S) + o(1),$$

where M, S^2 and T^3 are the expectation, variance and third central moment, respectively, of the random variable $h(x_i) = -\log\frac{p_i}{a_i}$, with respect to the probability distribution $\mathscr{P}$, and λ is defined by the equation

(12) $$\int_{-\infty}^{\lambda} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,dt = 1-\gamma.$$

Expression (11) is valid if the distribution of the random variable $h(\cdot)$ is non-lattice; in case $h(\cdot)$ has a lattice distribution, a slightly different expression, also given by STRASSEN [6], holds.

The only further tool we need is the following simple particular case of the Neymann—Pearson lemma.

LEMMA 3. *If $E^*\subset X^n$ is a set such that $p(E^*) = 1-\gamma$ and that $v\notin E^*$ implies $\frac{a(v)}{p(v)} \geqq \sup_{u\in E^*}\frac{a(u)}{p(u)}$, then*

(13) $$b(n,\gamma) = a(E^*).$$

Similarly, if $E^{**}\subset X^n$ *is a set such that* $a(E^{**})=a_0$ *and that* $v\notin E^{**}$ *implies*

$$\frac{p(v)}{a(v)}\geqq\sup_{u\in E^{**}}\frac{p(u)}{a(u)}$$

then

$$\inf_{E\subset X^n,\,a(E)\geqq a_0} p(E)=p(E^{**}). \tag{14}$$

PROOF. If $E^*\subset X^n$ is an arbitrary set satisfying $p(E^c)\leqq\gamma$, i.e. $p(E)\geqq 1-\gamma$, then $\sum_{v\in E\setminus E^*} p(v) = p(E\setminus E^*)\geqq p(E^*\setminus E) = \sum_{v\in E^*\setminus E} p(v)$. Thus

$$a(E\setminus E^*)\geqq\sup_{u\in E^*}\frac{a(u)}{p(u)}\sum_{v\in E\setminus E^*}p(v)\geqq\sup\frac{a(u)}{p(u)}\sum_{u\in E^*\setminus E}p(u)\geqq a(E^*\setminus E)$$

implying

$$a(E) = a(E\cap E^*)+a(E\setminus E^*)\geqq a(E\cap E^*)+a(E^*\setminus E) = a(E^*).$$

The assertions concerning (14) can be proved similarly.

§ 3. The results

Our aim is to estimate the minimum of $p(E^c)$, $E\subset X^n$, under the condition $a(E)=2^{nA}$, being A a constant whose value ranges within an interval to be specified later. Set

$$P(n,A)=\min_{E\subset X^n,\,a(E)\leqq 2^{nA}} p(E^c). \tag{15}$$

Let us introduce an auxiliary distribution $\mathcal{Q}_\alpha=(q_{\alpha 1},\dots,q_{\alpha m})$ defined as follows:

$$q_{\alpha i}=\frac{p_i^\alpha a_i^{1-\alpha}}{\sum_{j=1}^m p_j^\alpha a_j^{1-\alpha}}\qquad i=1,\dots,m \tag{16}$$

(this form for the auxiliary distribution is suggested by the requirement that we want $q_{\alpha i}$ and $q_\alpha(u)=q_{\alpha i_1}\dots q_{\alpha i_n}$ $(u=x_{i_1}\dots x_{i_n})$ to be similar functions of p_i and a_i and of $p(u)$ and $a(u)$, respectively). In (16) α is a non-negative parameter whose value will be fixed later.

Consider the function

$$h(\alpha)\stackrel{\text{def}}{=}-\sum_{i=1}^m q_{\alpha i}\log_2\frac{q_{\alpha i}}{a_i}\qquad(\alpha\geqq 0). \tag{17}$$

An elementary calculation shows that

$$h'(\alpha)=-\alpha\left\{\sum_{i=1}^m q_{\alpha i}\left(\log_2\frac{p_i}{a_i}\right)^2-\left(\sum_{i=1}^m q_{\alpha i}\log_2\frac{p_i}{a_i}\right)^2\right\}\log_2 e, \tag{18}$$

so that $h(\alpha)$ is a strictly decreasing function of α, unless $p_i=ca_i$ for every i $(1\leqq i\leqq m)$ for which $p_i>0$ (in the latter case $h(\alpha)$ is obviously constant).

Since

$$h(0) = \log_2 \sum_{i=1}^{m} a_i \quad \text{and} \quad h(1) = -\sum_{i=1}^{m} p_i \log_2 \frac{p_i}{a_i} = M$$

it follows that the equation

$$h(\alpha) = A \tag{19}$$

always has a unique solution α^*, with $0 < \alpha^* < 1$, whenever $M < A < \log_2 \sum_{i=1}^{m} a_i$.

THEOREM 1. *In case* $M < A < \log_2 \sum_{i=1}^{m} a_i$ *we have*

$$\lim_{n\to\infty} \frac{1}{n} \log_2 P(n, A) = -I(\mathcal{Q}_{\alpha^*} \| \mathcal{P}) \tag{20}$$

where α^* *is the (unique) solution of equation* (19) *and* $I(\mathcal{Q}_\alpha \| \mathcal{P})$ *denotes the I-divergence (or* KULLBACK—LEIBLER *information number) of the probability distributions* $\mathcal{P}$ *and* $\mathcal{Q}_\alpha$:

$$I(\mathcal{Q}_\alpha \| \mathcal{P}) = \sum_{i=1}^{m} q_{\alpha i} \log_2 \frac{q_{\alpha i}}{p_i} = \sum_{i=1}^{m} \frac{p_i^\alpha a_i^{1-\alpha}}{\sum_{i=1}^{m} p_j^\alpha a_j^{1-\alpha}} \log_2 \frac{p_i^{\alpha-1} a_i^{1-\alpha}}{\sum_{j=1}^{m} p_j^\alpha a_j^{1-\alpha}}. \tag{21}$$

PROOF. Let us order the sequences of the form $u = x_{i_1} \dots x_{i_n}$ according to decreasing ratios $\frac{p(u)}{a(u)}$; then, see (16), the sequences $u \in X^n$ are ordered also according to decreasing ratios $\frac{q_\alpha(u)}{a(u)}$ for every $\alpha > 0$ and according to decreasing ratios $\frac{p(u)}{q_\alpha(u)}$ for $0 < \alpha < 1$.

Let $E_n(A)$ denote the set made up by as many sequences as possible chosen in this order and satisfying the condition

$$a(E_n(A)) = \sum_{u \in E_n(A)} a(u) \leqq 2^{nA} \tag{22}$$

From (15), (22) and from lemma 3 (with $E_n^c(A+\varepsilon')$ playing the role of E^{**}) we obtain for any $\varepsilon' > 0$ and n sufficiently large

$$p(E_n^c(A+\varepsilon')) = \min_{E \subset X_n^c,\, a(E) \geqq a(E_n(A+\varepsilon'))} p(E) \leqq P(n, A) \leqq p(E_n^c(A)). \tag{23}$$

Let us fix $\varepsilon > 0$ such that $0 < \alpha^* - \varepsilon < \alpha^* + \varepsilon < 1$, and let $\varepsilon' > 0$ be such that $h(\alpha^* + \varepsilon) < A < A + \varepsilon' < h(\alpha^* - \varepsilon)$; applying lemma 1 first with $\mathcal{Q}_{\alpha^*+\varepsilon}$ and second with $\mathcal{Q}_{\alpha^*-\varepsilon}$ instead of $\mathcal{P}$ and taking into account lemma 3 (with $E_n^c(A)$ resp. $E_n^c(A+\varepsilon')$ playing the role of E^{**}) we can conclude that

$$\begin{aligned} q_{\alpha^*+\varepsilon}(E_n^c(A)) &\to 0 \\ q_{\alpha^*-\varepsilon}(E_n^c(A+\varepsilon')) &\to 1 \end{aligned} \qquad (n \to \infty) \tag{24}$$

We are going to apply lemma 1 with $a_i = p_i$ $(i = 1, \ldots, m)$ and with $\mathcal{Q}_\alpha$ in the role of $\mathcal{P}$ (first for $\alpha = \alpha^* + \varepsilon$ and second for $\alpha = \alpha^* - \varepsilon$). Observe preliminarily that in this case an application of lemma 3 with $E_n^c(A)$ resp. $E_n^c(A+\varepsilon')$ instead of E^* gives for n sufficiently large

$$(25) \qquad \inf_{E \subset X^n,\, q_{\alpha^*-\varepsilon}(E) \geq \frac{1}{2}} p(E) \leq p\big(E_n^c(A+\varepsilon')\big) \leq p\big(E_n^c(A)\big) \leq \inf_{E \subset X^n,\, q_{\alpha^*+\varepsilon}(E) \geq \frac{1}{2}} p(E),$$

taking into account that (24) implies $q_{\alpha^*+\varepsilon}\big(E_n^c(A)\big) < \frac{1}{2}$ and $q_{\alpha^*-\varepsilon}\big(E_n^c(A+\varepsilon')\big) > \frac{1}{2}$ if n is large enough.

Now applying lemma 1 yields, in force of (23) and (25)

$$(26) \qquad -I(\mathcal{Q}_{\alpha^*-\varepsilon}\|\mathcal{P}) \leq \varliminf_{n\to\infty} \frac{1}{n} \log_2 P(n, A) \leq \varlimsup_{n\to\infty} \frac{1}{n} \log_2 P(n, A) \leq -I(\mathcal{Q}_{\alpha^*+\varepsilon}\|\mathcal{P}).$$

Since $\varepsilon > 0$ may be arbitrarily small, and $I(\mathcal{Q}_\alpha\|\mathcal{P})$ is a continuous function of α, this completes the proof of theorem 1.

Remark 1. If $A < M$, equation (19) still has a soultion $\alpha = \alpha^*$, provided $A > \lim_{\alpha\to\infty} h(\alpha)$, but now $\alpha^* > 1$. Observe that whenever $\alpha > 1$, ordering the sequences $u \in X^n$ according to decreasing ratios $\dfrac{p(u)}{a(u)}$ corresponds to ordering them according to increasing ratios $\dfrac{p(u)}{q_\alpha(u)}$. This means that in the last application of lemma 3 in proving theorem 1, $E_n^c(A)$ cannot play the role of E^*; it is rather $E_n(A)$ that plays this role. This leads to the conclusion that $1 - P(n, A)$ and not $P(n, A)$ decreases exponentially as $n \to \infty$:

$$\lim_{n\to\infty} \frac{1}{n} \log_2 \big(1 - P(n, A)\big) = -I(\mathcal{Q}_{\alpha^*}\|\mathcal{P})$$

where $\alpha^* > 1$ is the (unique) solution of the equation $h(\alpha) = A$.

Remark 2. It is easy to check that the "error exponent"

$$e(A) = \lim_{n\to\infty} \left(-\frac{1}{n} \log_2 P(n, A) \right)$$

is a strictly convex function of A $(M < A < \log_2 \sum_{i=1}^m a_i)$. In fact, an elementary computation shows that

$$(28) \qquad \frac{d}{d\alpha} I(\mathcal{Q}_\alpha\|\mathcal{P}) = (\alpha - 1) \left\{ \sum_{i=1}^m q_{\alpha i} \left(\log_2 \frac{p_i}{a_i} \right)^2 - \left(\sum_{i=1}^m q_{\alpha i} \log_2 \frac{p_i}{a_i} \right)^2 \right\} \log_2 e,$$

whence, using (18) and (19)

$$(29) \qquad \frac{de(A)}{dA} = -\left[\frac{d}{d\alpha} I(\mathcal{Q}_\alpha\|\mathcal{P}) \Big/ \frac{d}{d\alpha} h(\alpha) \right]_{\alpha=\alpha^*} = \frac{1}{\alpha^*} - 1.$$

Since the function $h(\alpha)$ is decreasing, the solution α^* of equation (19) is a decreasing function of A, thus from (29) we see that $\dfrac{de(A)}{dA}$ is an increasing function of A.

Specializing theorem 1 to the case $a_i=1$, $i=1, \dots, m$, we obtain a solution for problem (i), § 1. In fact, in this case, setting $A=R$, $P(n, A)$ becomes $P_e=P_e(n, R)$ i.e. the minimum probability of erroneous decoding, when the outputs of length n of the source are encoded by using 2^{nR} different codewords. In this case the constant M of theorem 1 is the entropy of the source, i.e. $H=-\sum_{i=1}^{m} p_i \log_2 p_i$ and $\sum_{i=1}^{m} a_i$ becomes simply m; thus from theorem 1 and remark 1 we obtain the following

Corollary 1. Given the discrete memoryless source of § 1, (i), if R satisfies $H<R<\log_2 m$ then

$$P_e = 2^{-n(I(\mathcal{Q}_{\alpha^*}\|\mathcal{P})+o(1))} \tag{30}$$

being α^* $(0<\alpha^*<1)$ the (unique) solution of the equation

$$H(\mathcal{Q}_\alpha) \stackrel{\text{def}}{=} -\sum_{i=1}^{m} \frac{p_i^\alpha}{\sum_{j=1}^{m} p_j^\alpha} \log_2 \frac{p_i^\alpha}{\sum_{j=1}^{m} p_j^\alpha} = R. \tag{31}$$

If, on the contrary $H>R>\log_2 r$[4], where r is the number of the indices i for which $p_i = \max_{1\le j\le m} p_j$, we have

$$P_e = 1-2^{-n(I(\mathcal{Q}_{\alpha^*}\|\mathcal{P})+o(1))} \tag{32}$$

where α^* $(\alpha^*>1)$ is again the unique solution of equation (31).

Specializing theorem 1 to the case $a_i=q_i$, $i=1, \dots, m$ where $\mathcal{Q}=\{q_1, \dots, q_m\}$ is another probability distribution on X, we arrive at a solution of problem (ii) of § 1. In this case the constant M appearing in theorem equals $-\sum_{i=1}^{m} p_i \log_2 \frac{p_i}{q_i} = -I(\mathcal{P}\|\mathcal{Q})$ and $\sum_{i=1}^{m} a_i = \sum_{i=1}^{m} q_i = 1$. Thus we have (writing $-A$ instead of A):

Corollary 2. If in problem (ii), § 1, of testing hypothesis $\mathcal{P}$ against hypothesis $\mathcal{Q}$, the probability of the first kind error is required to decrease as 2^{-nA}, where $0<A<I(\mathcal{P}\|\mathcal{Q})$, then the minimum probability of the second kind error is $2^{-n(I(\mathcal{Q}_{\alpha^*}\|\mathcal{P})+o(1))}$ where α^* $(0<\alpha^*<1)$ is the unique solution of the equation

$$I(\mathcal{Q}_\alpha\|\mathcal{Q}) \stackrel{\text{def}}{=} \sum_{i=1}^{m} \frac{p_i^\alpha q_i^{1-\alpha}}{\sum_{j=1}^{m} p_j^\alpha q_j^{1-\alpha}} \log_2 \frac{p_i^\alpha q_i^{-\alpha}}{\sum_{j=1}^{m} p_j^\alpha q_j^{1-\alpha}} = A.$$

The results of corollaries 1 and 2 are known in the literature of information theory and mathematical statistics, respectively.

Formula (30) is derived in [4] and the result of corollary 2 can be deduced from the results of [3]. Both references rely on the method of large deviations (or Chernoff bounding technique) and obtain the estimates in a somewhat different algebraic form. Our method seems us more direct and more motivated. It should

[4] Observe that $\log_2 r = \lim_{\alpha\to\infty} H(\mathcal{Q}_\alpha)$.

be noted, however, that the proof of the "large deviations theorem" is also based on an auxiliary distribution, cf. e.g. [2], Appendix 5A, thus no substantial difference exists between the two approaches. Formula (32) is closely related to Theorem 1, p. 473 in [5].

If instead of lemma 1 we use the sharper estimate given in lemma 2, the following improvement of theorem 1 can be obtained:

THEOREM 2. *Under the conditions of theorem 1,* $\log_2 P(n, A)$ *has the following asymptotic expression:*

$$\log_2 P(n,A) = -nI(\mathcal{Q}_{\alpha^*}\|\mathcal{P}) - \frac{1}{2\alpha^*}\log_2 n + \tag{33}$$
$$+\frac{(1-\alpha^*)T_1^3}{6\alpha^* S_1^2} - \frac{T_2^3}{6S_2^2} - \frac{1-\alpha^*}{2\alpha^*}\log_2 S_1^2 - \frac{1}{2}\log_2 S_2^2 - \frac{1}{\alpha^*}\log_2\frac{2\pi}{\log_2 e} + o(1)$$

where $\mathcal{Q}_\alpha$ is defined by (16), α^* is the solution of equation (19), S_1^2 and S_2^2 are the variances and T_1^3 and T_2^3 are the third central moments of the random variables $h_1(x_i) = -\log_2 \frac{q_{\alpha^* i}}{a_i}$ and $h_2(x_i) = -\log_2 \frac{q_{\alpha^* i}}{p_i}$, respectively, in terms of the auxiliary distribution $\mathcal{Q}_{\alpha^*}$.

To be more precise, (33) holds in case the distributions of the mentioned random variables are non-lattice; otherwise a slightly different formula holds. Let us remark that the distributions of h_1 and h_2 are non-lattice if and only if the distribution of the random variable $h_0(x_i) = -\log_2 \frac{p_i}{a_i}$ is non-lattice.

PROOF. Choose α_n such that $q_{\alpha_n}(E_n(A)) = \frac{1}{2}$, where $E_n(A)$ is the same as in (22); for n large enough this is always possible (cf. (24)). Setting $\gamma = \frac{1}{2}$, i.e. $\lambda = 0$, in lemma 2 and applying it with $\mathcal{Q}_{\alpha_n}$ instead of $\mathcal{P}$ we obtain keeping in mind lemma 3 (attention should be paid to the bases of the logarithms):

$$nA = nh(\alpha_n) - \frac{1}{2}\log_2 n - \frac{T_1^3}{6S_1^2} - \frac{1}{2}\log_2(2\pi S_1^2) + \log_2\log_2 e + o(1). \tag{34}$$

Actually, we should have used in (34) the moments of the random variable $-\log_2 \frac{q_{\alpha_n i}}{a_i}$ with respect to the distribution $\mathcal{Q}_{\alpha_n}$; however, since from (34) one immediately sees that $\alpha_n \to \alpha^*$ as $n \to \infty$, the error introduced in this way vanishes in the limit.

Now we apply lemma 2 once more, still with $\gamma = \frac{1}{2}$ and $\mathcal{Q}_{\alpha_n}$ in the role of $\mathcal{P}$, but with p_i instead of a_i. In force of lemma 3 we get

$$\log_2 P(n,A) = -nI(\mathcal{Q}_{\alpha_n}\|\mathcal{P}) - \frac{1}{2}\log_2 n - \tag{35}$$
$$-\frac{T_2^3}{6S_2^2} - \frac{1}{2}\log_2(2\pi S_2^2) + \log_2\log_2 e + o(1)$$

Expanding both $h(\alpha)$ and $g(\alpha) \stackrel{\text{def}}{=} I(\mathcal{Q}_\alpha \| \mathcal{P})$ around $\alpha = \alpha^*$ and taking into account that $h(\alpha^*) = A$ we get

$$h(\alpha_n) = A + (\alpha_n - \alpha^*) h'(\alpha^*) + o(\alpha_n - \alpha^*) \tag{36}$$

$$I(\mathcal{Q}_{\alpha_n} \| \mathcal{P}) = I(\mathcal{Q}_{\alpha^*} \| \mathcal{P}) + (\alpha_n - \alpha^*) g'(\alpha^*) + o(\alpha_n - \alpha^*); \tag{37}$$

(34), (36) and (37) give rise to

$$\begin{aligned} I(\mathcal{Q}_{\alpha_n} \| \mathcal{P}) &= I(\mathcal{Q}_{\alpha^*} \| \mathcal{P}) + \\ &+ \frac{g'(\alpha^*)}{h'(\alpha^*)} \cdot \frac{1}{n} \left(\frac{1}{n} \log_2 n + \frac{T_1^3}{6S_1^2} + \frac{1}{2} \log_2 (2\pi S_1^2) - \log_2 \log_2 e + o(1) \right), \end{aligned} \tag{38}$$

Since $\dfrac{g'(\alpha^*)}{h'(\alpha^*)} = 1 - \dfrac{1}{\alpha^*}$ (cf. (29)), substituting (38) into (35) completes the proof of theorem 2.

Of course, theorem 2 can now be used to solve problems (i) and (ii) of § 1 with greater accuracy. Namely, under the same conditions and with the same notations as in Corollary 1 we obtain the following

Corollary 3. If the distribution of the random variable $h_0(x_i) = -\log_2 p_i$ is non-lattice, the exact limiting behaviour of the minimum probability of erroneous decoding in case $H < R < \log_2 m$ is given by

$$\begin{aligned} P_e = \exp_2 \Big\{ &-nI(\mathcal{Q}_{\alpha^*} \| \mathcal{P}) - \frac{1}{2\alpha^*} \log_2 n + \frac{(1-\alpha^*) T_1^3}{6\alpha^* S_1^2} - \\ &- \frac{T_2^3}{6S_2^2} - \frac{1-\alpha^*}{2\alpha^*} \log_2 S_1^2 - \frac{1}{2} \log_2 S_2^2 - \frac{1}{\alpha^*} \log_2 \frac{2\pi}{\log_2 e} + o(1) \Big\} \end{aligned} \tag{39}$$

being α^* the (unique) solution of equ. (31).

Conversely, if $H > R > \log_2 r$, we have

$$\begin{aligned} P_e = 1 - \exp_2 \Big\{ &-nI(\mathcal{Q}_{\alpha^*} \| \mathcal{P}) - \frac{1}{2\alpha^*} \log_2 n + \frac{(1-\alpha^*) T_1^3}{6\alpha^* S_1^2} - \\ &- \frac{T_2^3}{6S_2^2} - \frac{1-\alpha^*}{2\alpha^*} \log_2 S_1^2 - \frac{1}{2} \log_2 S_2^2 - \frac{1}{\alpha^*} \log_2 \frac{2\pi}{\log_2 e} + o(1) \Big\}. \end{aligned} \tag{40}$$

Under the same conditions and with the same notations as in Corollary 2, from theorem 2 we obtain the following:

Corollary 4. If the distribution of the random variable $h_0(x_i) = -\log_2 \dfrac{p_i}{q_i}$ is non-lattice, the minimum probability of the second kind error in testing hypothesis $\mathcal{P}$

against hypothesis $\mathcal{Q}$ is given by

$$\exp_2\left\{-nI(\mathcal{Q}_{\alpha^*}\|\mathcal{P})-\frac{1}{2\alpha^*}\log_2 n+\frac{(1-\alpha^*)T_1^3}{6\alpha^* S_1^2}-\frac{T_2^3}{6S_2^2}-\right.$$

$$\left.-\frac{1-\alpha^*}{2\alpha^*}\log_2 S_1^2-\frac{1}{2}\log_2 S_2^3-\frac{1}{\alpha^*}\log_2\frac{2\pi}{\log_2 e}+o(1)\right\}$$

whenever the probability of the first kind error is bound to decrease as 2^{-nA}, $0<A<$ $<I(\mathcal{P}\|\mathcal{Q})$. Here α^* $(0<\alpha^*<1)$ is the unique solution of the equation at the end of Corollary 2.

For the case of lattice distributions slightly different asymptotic formulas hold, which can be derived in a completely similar way.

§ 4. Concluding remarks

For the sake of simplicity, we have restricted our attention to the case when X is a finite set. With slight modifications and under some regularity conditions, theorems 1 and 2 remain valid for the general case as well. If X is an arbitrary set and $\mathcal{X}$ is a σ-algebra of subsets of X, one may consider a probability measure μ and a σ-finite measure λ on $\mathcal{X}$, instead of the p_i's and a_i's, respectively. Assuming $\mu\ll\lambda$ and letting $p(x)=\frac{\mu(dx)}{\lambda(dx)}$ be its Radon—Nikodym derivative, the set $E_n(A)$ may now be defined as a subset of X^n with the property that

$$\prod_{i=1}^{n} p(y_i)\leqq \inf_{(x_1,\dots,x_n)\in E_n(A)}\prod_{i=1}^{n} p(x_i)$$

whenever $(y_1,\dots,y_n)\notin E_n(A)$ and such that $\lambda^n(E_n(A))=2^{nA}$ (or, if equality cannot be reached, $\lambda^n(E_n(A))$ should be as large as possible below 2^{nA}). The auxiliary distribution playing the role of $\mathcal{Q}_\alpha$ can be defined as the probability measure ν_α having density (i.e. Radon—Nikodym derivative with respect to λ) $q_\alpha(x)=\frac{p^\alpha(x)}{\int p^\alpha(x)\lambda(dx)}$ provided the integral in the denominator is finite. Unlike the case of a finite set X, the auxiliary distribution (with density $q_\alpha(x)$) need not be defined for every $\alpha\geqq 0$; it can be shown, however, that the function $h(\alpha)=-\int q_\alpha(x)\log_2 q_\alpha(x)\lambda(dx)$ possesses all the properties of the function (17) in the interval where $q_\alpha(x)$ is defined, apart, conceivably, from the endpoints of this interval. Thus, if the constant A is given, $(A>M=-\int p(x)\log_2 p(x)\lambda(dx))$ and is such that the equation $h(\alpha)=A$ has a root $\alpha=\alpha^*$ belonging to the interior of the domain of $h(\alpha)$, theorems 1 and 2 remain valid, and their proofs run along the same lines.

As regards the problem of source coding, only a finite or countably infinite set X makes sense if we stick to the adopted formulation. In the case of a countably infinite X corollaries 1 and 2 remain valid with the only change that the inequality $H<R<\log_2 m$ should be replaced by $H<R<R_0$ where $R_0=\lim_{\alpha\to\alpha_0+0} h(\alpha)$, being α_0 the infimum of those α's for which $\sum_{i=1}^{\infty} p_i^\alpha<\infty$. Let us point out, however, that even the case $\alpha_0=1$ can occur $\left(\text{e.g. if } p_i = \frac{c}{i\log^3(i+1)},\ i=1,2,\ldots\right)$ for which the assertion becomes vacuous. If $R_0<\infty$ and $R\geqq R_0$, the asymptotic behaviour of P_e remains unknown.

REFERENCES

[1] Арутюнян, Е. А.: Оценки экспоненты вероятности ошибки для полунепрерывного канала без памяти. *Проблемы передачи информации*, **4** (1968), 37—48.
[2] Gallager, R.: *Information Theory and Reliable Communication,* J. Wiley and S., New York, 1968.
[3] Hoeffding, W.: Asymptotically Optimal Tests for Multinomial Distributions, *Annals of Math. Stat.* **36** (1965), 369—400.
[4] Jelinek, F.: *Probabilistic Information Theory* ,McGraw-Hill, New York 1968.
[5] Rényi, A.: *Wahrscheinlichkeitsrechnung,* VEB Deutscher Verlag der Wissenschaften, Berlin, 1962.
[6] Strassen, V.: Asymptotische Abschätzungen in Shannon's Informationstheorie, *Trans. of the Third Prague Conf. on Information Theory* etc., Prague 1964, pp. 689—723.

Mathematical Institute of the Hungarian Academy of Sciences

(Received January 29, 1970)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 193—196.

ON THE MINIMUM VALUE OF A QUADRATIC FUNCTION UNDER LINEAR CONSTRAINTS

by

G. KÉRI

Introduction. The theorem being proved in this paper states a proposition, the validity of which is presupposed in several quadratic programming contributions. Its proof is outlined, as far as I know, only in [1] by FRANK and WOLFE. It is mentioned there that in case of unbounded polyhedron K the statement can be proved by means of the Motzkin theorem; this proof is referred to, for instance in [3] and [4]. In case of bounded polyhedron the proposition is self-evident. However, I could not find out how to complete the proof by means of the instructions given in [1], so here I am going to show a straightforward inductive proof.

THEOREM. *If a quadratic function*

$$q(x) = x'Qx+2c'x$$

is bounded from below in an (unbounded) polyhedron

$$K=\{x: Ax \leqq b\}$$

then there exists a vector $x^0 \in K$, such that

$$q(x^0)=\min \{q(x):\ x \in K\}.$$

PROOF. The theorem will be proved by induction on the number of the variables of $q(x)$, that is for the number of the components of x. In the 1-dimensional case the statement is trivial. Denoting

$$x'=(x_1, x_2, \ldots, x_n),$$

let us suppose that the theorem has been proved for any quadratic objective function of $n-1$ variables and for any polyhedron in the $n-1$ dimensional Euclidean space.

Let us denote

$$\hat{x}'=(x_2, x_3, \ldots, x_n),$$

$$c'=(c_1, \hat{c}')=(c_1, c_2, \ldots, c_n).$$

Then we can write

$$q(x) = p(x_1; \hat{x})+q^*(\hat{x}),$$

where $p(x_1; \hat{x}) = ax_1^2+2(d'\hat{x})x_1+2c_1x_1$ (the vector $\begin{pmatrix} a \\ d \end{pmatrix} = \begin{pmatrix} a \\ d_2 \\ d_3 \\ \vdots \\ d_n \end{pmatrix}$ being the first column of the matrix Q), while $q^*(\hat{x})$ is a quadratic function of the $n-1$ variables $x_2, x_3, \ldots, x_n$.

Further notations:

R^n for the n-dimensional Euclidean space;

$d' = (d_2, d_3, \ldots, d_n)$;

$$A = \begin{pmatrix} a_{11} & a_{12} \ldots a_{1n} \\ a_{21} & a_{22} \ldots a_{2n} \\ \vdots & \\ a_{m1} & a_{m2} \ldots a_{mn} \end{pmatrix}$$

$$a_{01} = a;$$

$$a_{0j} = d_j \qquad (j = 2, 3, \ldots, n);$$

$$b_0 = -c_1;$$

let $j \in J$ if and only if $0 \leqq j \leqq m$ (j integer) and $a_{j1} \neq 0$;

$$l_j(\hat{x}) = \frac{b_j - \sum_{k=2}^{n} a_{jk} x_k}{a_{j1}} \qquad (j \in J);$$

$$l_{m+1}(\hat{x}) = p(x_2; \hat{x}) = ax_2^2 + 2\left(\sum_{k=2}^{n} d_k x_k\right) x_2 + 2c_1 x_2;$$

$$J' = J \cup \{m+1\};$$

$$I(\hat{x}) = \{x_1 : (x_1, \hat{x}) \in K\}.$$

Now we can write that

$$\inf\{q(x) : x \in K\} = \inf_{\hat{x} \in R^{n-1}} \{q^*(\hat{x}) + \inf\{p(x_1; \hat{x}) : x_1 \in I(\hat{x})\}\}. \tag{1}$$

Let $\hat{x}^0 \in R^{n-1}$ be such a vector that $I(\hat{x}^0)$ is non-empty. Then

$$\inf\{p(x_1; \hat{x}^0) : x_1 \in I(\hat{x}^0)\} = \min\{p(l_j(\hat{x}^0); \hat{x}^0) : j \in J', (l_j(\hat{x}^0), \hat{x}^0) \in K\} \tag{2}$$

To prove the validity of (2), let us denote

$$S_{\text{left}} = \{p(x_1; \hat{x}^0) : x_1 \in I(\hat{x}^0)\}$$

and

$$S_{\text{right}} = \{p(l_j(\hat{x}^0); \hat{x}^0) : j \in J', (l_j(\hat{x}^0), \hat{x}^0) \in K\}.$$

First we shall demonstrate that S_{right} is non-empty. On account of the convexity of K, $I(\hat{x}^0)$ is a finite or infinite interval. Let us first consider the case where $I(\hat{x}^0) \neq (-\infty, +\infty)$, and let x_1^0 be an end point of the interval $I(\hat{x}^0)$. Then $(x_1^0, \hat{x}^0)$ must lie on the boundary of the polyhedron K, consequently

$$a_{j1} x_1^0 + \sum_{k=2}^{n} a_{jk} x_k^0 = b_j \tag{3}$$

for at least one j ($1 \leqq j \leqq m$). Let J^0 be the set of those indices j, for which x_1^0 satisfies the equality (3). If $a_{j1} = 0$ for all $j \in J^0$ then, for sufficiently small $\varepsilon > 0$, $(x_1^0 + \varepsilon, \hat{x}^0) \in K$

and $(x_1^0-\varepsilon, \hat{x}^0)\in K$, which contradicts to the fact that x_1^0 has been chosen as an end point of $I(\hat{x}^0)$. For this reason, there must be a j, for which x_1^0 satisfies (3), and $a_{j1}\neq 0$. For this j, $j\in J'$ and

$$x_1^0 = \frac{b_j - \sum_{k=2}^{n} a_{jk} x_k^0}{a_{j1}} = l_j(\hat{x}^0),$$

that is $(l_j(\hat{x}^0), \hat{x}^0)\in K$. If, on the other hand, $I(\hat{x}^0)=(-\infty, +\infty)$ then $m+1\in J'$ and $(l_{m+1}(\hat{x}^0), \hat{x}^0)\in K$. Now it can be seen that S_{right} is a finite, non-empty set. Moreover, since obviously $S_{\text{left}}\supset S_{\text{right}}$, we can state that

$$\inf\{p(x_1; \hat{x}^0): x_1\in I(\hat{x}^0)\} \leqq \min\{p(l_j(\hat{x}^0); \hat{x}^0): j\in J', (l_j(\hat{x}^0), \hat{x}^0)\in K\}.$$

In order to prove the inverse inequality first we can observe that $\inf\{p(x_1; \hat{x}^0): x_1\in I(\hat{x}^0)\}<+\infty$ because of $I(\hat{x}^0)\neq\emptyset$, then

$$\inf\{p(x_1; \hat{x}^0): x_1\in I(\hat{x}^0)\} > -\infty,$$

which follows from $\inf\{q(x): x\in K\} > -\infty$. Since $p(x_1, \hat{x}^0)$ is now being regarded as a quadratic function of only x_1, there exists an $x_1^0\in I(\hat{x}^0)$ satisfying

$$p(x_1^0, \hat{x}^0) = \inf\{p(x_1; \hat{x}^0): x_1\in I(\hat{x}^0)\}. \tag{4}$$

If this x_1^0 is unique and lies in the interior of $I(\hat{x}^0)$, then $a=a_{01}>0$ and

$$x_1^0 = -\frac{d'\hat{x}^0+c_1}{a} = l_0(\hat{x}^0),$$

consequently the minimal element of S_{left} is among the elements of S_{right}. If $I(\hat{x}^0)= =(-\infty, +\infty)$ then, because of the finiteness of $\inf S_{\text{left}}$, $p(x_1; \hat{x}^0)$ is independent of x_1 in the whole $(-\infty, +\infty)$ interval, so $x_1^0=l_{m+1}(\hat{x}^0)$ may be chosen. Finally, in any other case an $x_1^0\in I(\hat{x}^0)$ satisfying (4) may be chosen as an end point of $I(\hat{x}^0)$. Then according to the argumentation of the previous paragraph, there exists a $j\in\{1, 2, \ldots, m\}$ such that $a_{j1}\neq 0$ and $x_1^0=l_j(\hat{x}^0)$. Now the validity of (2) has been justified for each possible case.

From this point the proof can be easily completed. From (1) and (2) we gain that

$$\inf\{q(x): x\in K\} = \min_{j\in J'}\{\inf\{q_j^*(\hat{x}): \hat{x}\in K_j\}\}, \tag{5}$$

where

$$q_j^*(\hat{x}) = q^*(\hat{x})+p(l_j(\hat{x}); \hat{x})$$

and

$$K_j = \{\hat{x}: (l_j(\hat{x}), \hat{x})\in K\}.$$

Since $\inf\{q(x): x\in K\}$ is finite, $\inf\{q_j^*(\hat{x}): \hat{x}\in K_j\} > -\infty$ for arbitrary $j\in J'$. $q_j^*(\hat{x})$ is a quadratic function of $\hat{x}$ and K_j is a convex polyhedron in the $n-1$ dimensional space, thus by the inductive assumption

$$\inf\{q_j^*(\hat{x}): \hat{x}\in K_j\} = q_j^*(\hat{x}^{0,j}) \tag{6}$$

provided that $j \in J'$ and K_j is non-empty. If K_j is empty for all $j \in J'$ then $I(\hat{x})$ is empty for any $\hat{x} \in R^{n-1}$, consequently K itself is empty. If this is not the case then by (5) and (6) we can arrive at

$$\inf\{q(x) : x \in K\} = \min_{\substack{j \in J' \\ K_j \neq \emptyset}} \{q_j^*(\hat{x}^{0,j})\} = q_r^*(\hat{x}^{0,r}) =$$

$$= p\big(l_r(\hat{x}^{0,r}); \hat{x}^{0,r}\big) + q^*(\hat{x}^{0,r}) = q\big(l_r(\hat{x}^{0,r}); \hat{x}^{0,r}\big) = q(x^0),$$

where

$$\hat{x}^{0,r} \in K_r,$$

that is

$$x^0 = \big(l_r(\hat{x}^{0,r}), \hat{x}^{0,r}\big) \in K.$$

The theorem is thus completely proved.

Acknowledgement. I am indebted to A. MAJTHAY for letting me acquainted with the problem and his proof in the case of convex quadratic objective function.

REFERENCES

[1] FRANK, M. and WOLFE, P.: An algorithm for Quadratic Programming, *Nav. Res. Log. Qu.*, **3** (1956), 95—110.
[2] DORN, W. S.: Duality in Quadratic Programming, *Qu. Appl. Math.*, **18** (1960), 155—162.
[3] DORN, W. S.: Self Dual Quadratic Programs, *J. Soc. Ind. Appl. Math.*, **9** (1961), 51—55.
[4] COTTLE, R. W.: Note on a Fundamental Theorem in Quadratic Programming, *J. Soc. Ind. Appl. Math.*, **12** (1964), 663—665.

Computing Centre of the Hungarian Academy of Sciences

(Received February 3, 1970.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 197—203.

EXISTENCE THEOREM FOR WEAK SOLUTIONS OF ORDINARY DIFFERENTIAL EQUATIONS IN REFLEXIVE BANACH SPACES

by

A. SZÉP

I. The general problem

Let E be a topological vector space, denote R the real line, and let $x(t)$ be a function with values in E, defined in a neighborhood of the point $t_0 \in R$. Denote τ the topology of the space E. The function $x(t)$ is said to be differentiable according to the topology τ, or briefly τ-differentiable if the quotient $\frac{x(t)-x(t_0)}{t-t_0}$ converges to a certain $x'(t_0) \in E$ for $t \to t_0$, in the topology τ. Let Ω be a (temporarily arbitrary) but non-void subset of the product space $R \times E$, and $f(t, x)$ a function, defined on $\Omega, f(t, x) \in E$ for $(t, x) \in \Omega$. Let (t_0, x_0) be a fastened point of Ω. Consider the problem

$$\left.\begin{aligned} x' &= f(t, x) \\ x(t_0) &= x_0 \end{aligned}\right\} \tag{1}$$

Definition: An E-valued function $x(t)$, defined on some non-degenerated real interval I, containing t_0, is a solution of the problem (1), if:

a) $x(t)$ is τ-differentiable for $t \in I$,

b) $(t, x(t)) \in \Omega$ for $t \in I$

c) $x(t_0) = x_0$

d) $x'(t) = f(t, x(t))$ for $t \in I$

The problem cannot be expected to be solved in such generality. Various conditions on E, τ, Ω and f enable to be proved existence theorems.

Such theorems are the Cauchy—Peano and the Caratheodory existence theorem (both are valid only in finite dimensional spaces), the Picard—Lindelöf existence theorem, which can be extended to arbitrary Banach spaces (see [1]), the Krasnoselskii—Krein-theorem [2] and so on.

II. The special problem

In this paper we shall investigate the following special case. The space E will be a reflexive Banach space, the topology τ will be the weak-topology, which by the assumption coincides with the weak* topology of the adjoint space of E. The set Ω will be closed in the norm-topology of E and the function $f(t, x)$ will be continuous on Ω, in the weak-weak sense, that is for every $(t', x') \in \Omega$ and arbitrary weak neighbourhood U of the point $f(t', x') \in E$ there exist an $\varepsilon > 0$ and a weak neighbourhood V of x' so that for every $x \in V$, $|t-t'| < \varepsilon$ $(t, x) \in \Omega$, $f(t, x) \in U$ is valid.

III. The tools

The first property, which will be used repeatedly is the weak completeness of reflexive Banach spaces. Strictly speaking:

1. Let $\{x_n\}_{n=1}^{\infty}$ be a (denumerable) sequence in the reflexive Banach space E. Suppose, that for all $x^* \in E^*$ — that is for all linear continuous functionals on E, the sequence $x^*(x_n)$ converges. Then there exists an element $x_0 \in E$, for which $\lim_{n\to\infty} x^*(x_n) = x^*(x_0)$ for $x^* \in E^*$ arbitrary. For proof see [3], p. 69, 29 Corollary).

2. A Banach space is reflexive if and only if its closed unit sphere (hence all closed spheres) is weakly compact. ([3] p. 425 Theorem).

3. A subset C of a reflexive Banach space is weakly compact if and only if it is closed in the weak topology and bounded in the norm-topology (see [3], p. 424 3. Cor.).

4. Our main tool is the EBERLEIN—ŠMULIAN theorem:

Let A be a subset of a Banach space E. Then the following statements are equivalent:

(i) A is weakly sequentially compact, i.e. any sequence in A has a subsequence which converges to an element of E.

(ii) every infinit subset of A has a weak limit point in E.

(iii) the closure of A in the weak topology is weakly compact ([3] p. 430 1. Theorem).

COROLLARY 1. *Let E be a reflexive Banach space $\{x_n\}_{n=1}^{\infty}$ a sequence of elements of E, bounded in norm, that is there exists a number $K \geqq 0$ for which $\|x_n\| \leqq K$ for all n. Then there exists a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ of the original sequence which converges weakly to an element of E.*

PROOF. Let S be the closed sphere in E with center in the origin and with radius K. S is compact by 2, it follows by 4, that our sequence contains a weakly convergent subsequence. Q.e.d.

COROLLARY 2. *Let E be a Banach space, R the the real line, $a, b > 0$ given real constants. Denote P the "cilyndre"*

$$\{(t, x): t_0 \leqq t \leqq t_0 + a, \ \|x - x_0\| \leqq b\}; \quad P \subset R \times E$$

where t_0 is a fixed real number and $x_0 \in E$ a fixed elements. Let $f(t, x)$ be a function, defined on P, with values in E. Suppose that $f(t, x)$ weak-weak continuous on P (cf. Ch. II). *Then there exists an upper bound M, so that*

$$\|f(t, x)\| \leqq M \quad \textit{for} \quad (t, x) \in P.$$

PROOF. According to Tychonoff's theorem on topological products, P is compact in the "real × weak" topology. The continuity of f implies the weak compactness of its range $f(P)$, hence by 3., the range is bounded in norm. Qu.e.d.

We shall use the following simple consequence of the Hahn—Banach extension theorem:

5. Let x_0 be an element of the Banach space E with the property $\|x^*(x_0)\| \leqq K$ for all $x^* \in E^*$, $\|x^*\| = 1$, then $\|x_0\| \leqq K$.

IV. Some remarks on Banach space analysis

Let E be a Banach space — reflexive or not — denote E^* its adjoint space. In this chapter we shall deal with functions, the values of which lie in E. These functions will be called briefly vector-valued-, or vector-functions.

A more detailed theory of such functions is developed in [1] and [3]. Here we shall give only the definitions and theorems essential to our purposes.

Let $x(t)$ be a function of the real variable t with values in E.

1. Definition. The function $x(t)$ is said to be weakly continuous at $t=t_0$, if $t\to t_0$ implies $x(t)\to x(t_0)$ in the weak topology. Our condition is equivalent to the following one:

$$t\to t_0 \quad \text{implies} \quad x^*(x(t))\to x^*(x(t_0)) \quad \text{for all} \quad x^*\in E^*.$$

2. Definition. The function $x(t)$ is said to be weakly (Riemann) integrable in some interval $[a, b]$, if in any choice of the points τ_i, $t_{i-1}\leqq\tau_i\leqq t_i$ $i=1, \dots, n$ the sums $\sum_{i=1}^{n} x(\tau_i)(t_i-t_{i-1})$ converge weakly to the same element $x_0\in E$, provided $\max_{1\leqq i\leqq n} |t_i-t_{i-1}| \to 0$. Let the space E be weakly complete, then our condition is equivalent to the Riemann integrability of all complex-valued function of the form $x^*(x(t))$. In this case there exists an element x_0 of E, for which

$$x^*(x_0)=\int_a^b x^*(x(t))\,dt \quad \text{for all} \quad x^*\in E^*.$$

3. Definition. The function $x(t)$ is said to be weakly differentiable at $t=t_0$, if the difference-quotient $x(t)-x(t_0)/(t-t_0)$ converges weakly for $t\to t_0$ to an element $x'(t_0)\in E$. In the weakly complete case this condition is equivalent to the differentiability of all functions $x^*(x(t))$, $x^*\in E^*$. It can be easy to be proved that weak differentiability implies weak continuity, the latter implies weak integrability, the integral of a weakly continuous function is weakly differentiable with respect to the right endpoint of the integration interval and its derivative equals to the integrand (at the same point).

4. Definition. Let I be an interval on the real line, and $x_n(t)\in E$ for $t\in I$, $n=1, 2, \dots$. The sequence $\{x_n(t)\}_{n=1}^{\infty}$ converges weakly uniformly to the vector function $x(t)$, if for all $x^*\in E^*$, $\varepsilon>0$ there exists a $n_0=n_0(x^*, \varepsilon)$, so that $n\geqq n_0$ implies $|x^*(x_n(t))-x^*(x(t))| < \varepsilon$ for all $t\in I$

5. Theorem. *If all terms of the sequence $\{x_n(t)\}_{n=1}^{\infty}$ are weakly-continuous in I (that is in all points of I) and the sequence converges weakly uniformly to the function $x(t)$, then $x(t)$ is weakly continuous.*

Proof. Let $x^*\in E^*$, $\varepsilon>0$ arbitrary. Choose $n_0=n_0\left(x^*, \frac{\varepsilon}{3}\right)$ by means of the weak uniform convergence. Let $n\geqq n_0$ be a fixed integer, $t_0\in I$ arbitrary, but fixed too. The function $x^*(x_n(t))$ is continuous at t_0, therefore we can choose a $\delta > 0$ to $\frac{\varepsilon}{3}$. Let $|t-t_0| < \delta$, then

$$x^*(x(t))-x^*(x(t_0)) = x^*(x(t))-x^*(x_n(t))+x^*(x_n(t))-$$
$$-x^*(x_n(t_0))+x^*(x_n(t_0))-x^*(x(t_0)),$$

hence

$$|x^*(x(t))-x^*(x(t_0))| < \frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3} = \varepsilon \quad \text{Qu.e.d.}$$

6. DEFINITION. The sequence $\{x_n(t)\}_{n=1}^{\infty}$, defined on the real interval I is weakly uniformly fundamental, if for all $x^* \in E^*$ and $\varepsilon > 0$ there exists an $n_0 = n_0(x^*, \varepsilon)$ so that $m, n \geqq n_0$ implies $|x^*(x_n(t))-x^*(x_m(t))| < \varepsilon$ for all $t \in I$.

7. THEOREM. *Let the Banach space E be weakly complete, the sequence* $\{x_n(t)\}_{n=1}^{\infty}$ *weakly fundamental in I. Then the sequence converges weakly uniformly to a certain vector function* $x(t)$.

PROOF. For all fixed $t \in I$ the sequence $\{x_n(t)\}_{n=1}^{\infty}$ is weakly fundamental, hence it has a weak limit. Denote this limit element by $x(t)$. Let $x^* \in E^*$, $\varepsilon > 0$ be arbitrary $n_0 = n_0\left(x^*, \frac{\varepsilon}{2}\right)$ by means of the weak fundamentality, n, $m \geqq n_0$. Then $|x^*(x_n(t))-x^*(x_m(t))| < \frac{\varepsilon}{2}$ for all $t \in I$. Let t and m be fixed, while $n \to \infty$, then $x^*(x_n(t)) \to x^*(x(t))$, that is

$$x^*(x_n(t))-x^*(x_m(t)) \to x^*(x(t))-x^*(x_m(t)),$$

hence

$$|x^*(x(t))-x^*(x_m(t))| \leqq \frac{\varepsilon}{2} < \varepsilon.$$

8. THEOREM. *Let E be a weakly complete Banach space,* $\{x_n(t)\}_{n=1}^{\infty}$ *a weakly uniformly fundamental sequence of weakly continuous functions defined on the real interval I. Then the sequence converges weakly uniformly to a well-determined weakly continuous function* $x(t)$, *defined on I.*

PROOF. The assertion is an easy consequence of Theorems 8 and 7.

9. DEFINITION. Let E be a Banach space, $\mathcal{F} = \{x_\alpha(t): \alpha \in A\}$ a family of vector valued functions defined on the real interval I, A some index set. The family $\mathcal{F}$ is said to be equicontinuous, if for $\varepsilon > 0$ there exists a $\delta > 0$, so that $|t'-t''| < \delta$ implies $\|x_\alpha(t')-x_\alpha(t'')\| < \varepsilon$ for all $\alpha \in A$. This implies, among others, the weak continuity of all $x_\alpha(t)$ in I.

10. THEOREM. *Let E be a reflexive (hence weakly complete) Banach space. Let* $\mathcal{F}$ *be, as in* 9. Definition, *a family of equicontinuous function. Moreover let the family be bounded that is let there exists a* $K \geqq 0$, *so that* $\|x_\alpha(t)\| \leqq K$ *for all* $\alpha \in A$, $t \in I$. *Then there exists a sequence* $\alpha(n)$ *so that the sequence* $\{x_{\alpha(n)}\}_{n=1}^{\infty}$ *is weakly uniformly fundamental.*

PROOF. Let $\{t_n\}_{n=1}^{\infty}$ be a denumerable dense set in I. The set $x_\alpha(t_1)$ is bounded by the assumption, hence by Corollary 1 of Chapter III there exists a sequence $\alpha_{11}, \alpha_{12}, \dots, \alpha_{1n}, \dots \in A$ so that $\{x_{\alpha_{1n}}(t_1)\}_{n=1}^{\infty}$ is weakly fundamental. Consider the set $\{x_{\alpha_{1n}}(t_2)\}_{n=1}^{\infty}$. It is bounded, hence contains a weakly fundamental subsequence, denoted with $x_{\alpha_{21}}(t_2), x_{\alpha_{22}}(t_2), \dots, x_{\alpha_{2n}}(t_2), \dots$. We construct in the same way the sequences $x_{\alpha_{ij}}(t)$ for all i, and put $\alpha(n)=\alpha_{nn}$. Then the subsequence $\{x_{\alpha(n)}(t)\}_{n=1}^{\infty}$ is weakly fundamental for all $t=t_1, t_2, \dots$. We shall prove that $\{x_{\alpha(n)}(t)\}_{n=1}^{\infty}$ is weakly uniformly fundamental. To this end let $x^* \in E^*$, $\varepsilon>0$ be arbitrary. Without loss of generality we can suppose, that $\|x^*\|=1$. Choose $\delta>0$ to $\varepsilon/3$ by means of the equicontinuity. The intervals with radius δ and centre in t_i $(i=1, 2, \dots)$ cover I, hence there exists a finite covering system.

Suppose that the covering system consists of the first N interval. Chose n_0 by means of the weak fundamentality in $t_1, \dots, t_N$, so that for $k_1, k_2 \geqq n_0$

$$\left|x^*\left(x_{\alpha(k_1)}(t_i)\right)-x^*\left(x_{\alpha(k_2)}(t_i)\right)\right| < \frac{\varepsilon}{3} \quad \text{for} \quad i=1, \dots, N,$$

Let $K_1, K_2 \geqq n_0$, $t \in I$ be arbitrary. There exists a term of the finite covering system, containing. We can suppose without loss of generality that it is the first one.

$$x^*\left(x_{\alpha(k_1)}(t)\right)-x^*\left(x_{\alpha(k_2)}(t)\right) = x^*\left(x_{\alpha(k_1)}(t)\right)-x^*\left(x_{\alpha(k_1)}(t_1)\right)+$$
$$+x^*\left(x_{\alpha(k_1)}(t_1)\right)-x^*\left(x_{\alpha(k_2)}(t_1)\right)+x^*\left(x_{\alpha(k_2)}(t_1)\right)-x^*\left(x_{\alpha(k_2)}(t)\right).$$

For $l=1, 2$ by means of the equicontinuity

$$\left|x^*\left(x_{\alpha(k_l)}(t)\right)-x^*\left(x_{\alpha(k_l)}(t_1)\right)\right| \leqq \|x^*\| \cdot \|x_{\alpha(k_l)}(t)-x_{\alpha(k_l)}(t_1)\| \leqq 1 \cdot \frac{\varepsilon}{3} = \frac{\varepsilon}{3}.$$

The weak fundamentality in t implies

$$\left|x^*\left(x_{\alpha(k_1)}(t_1)\right)-x^*\left(x_{\alpha(k_2)}(t_1)\right)\right| < \frac{\varepsilon}{3}$$

Hence

$$\left|x^*\left(x_{\alpha(k_1)}(t)\right)-x^*\left(x_{\alpha(k_2)}(t)\right)\right| < \frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3} = \varepsilon$$

Since n_0 is independent from t the proof is complete.

V. The existence theorem

Through this chapter let E be a reflexive Banach space, denote R the real line, $f(t, x)$ a weak-weak continuous function on the cylinder

$$P = \{t_0 \leqq t \leqq t_0+a, \ \|x_0-x\| \leqq b\}$$

in accordance with Chapter II. By Corollary 2 at Chapter III there exists an upper bound $M>0$ of f in norm. Let $\alpha=\min(a, b/M)$

THEOREM. *The problem* (1) *in Chapter* I *has at least one weak solution defined on* $[t_0, t_0+\alpha]$. (Strictly speaking we assert the existence of a solution in the sense of Chapter I, while τ is the weak topology of the space E.)

PROOF. The mean idea used in the sequel constructing approximate solutions due to TONELLI [4]. Problem (1) is equivalent to the following (set of) integral equation(s):

$$x^*(x(t_0)) = x^*(x_0) + \int_{t_0}^{t} x^*(f(s, x(s)))\, ds \quad \text{for all} \quad x^* \in E^*.$$

Let $\delta > 0$ be a suitable small real constant, $x_0(t)$ a weakly continuously differentiable function on $[t_0 - \delta, t_0]$, for which $x(t_0) = x_0$; $x_0'(t_0) = f(t_0, x_0)$; $\|x_0(t) - x_0\| \leqq b$ and $\|x_0'(t)\| \leqq M$. (The property "weakly continuously differentiable" means that $x_0(t)$ is weakly differentiable and its derivative, as a function of t is weakly continuous.)

Let us remark that for $\delta(>0)$ small enough the function

$$x_0(t) = x_0 + (t - t_0) \cdot f(t_0, x_0)$$

has the required properties, but we do not restrict the choice of $x_0(t)$ beyond the required properties. For $0 < \varepsilon \leqq \delta$ define $x_\varepsilon(t)$ on $[t_0 - \delta, t_0 + \alpha)$ by

$$x_\varepsilon(t) = \begin{cases} x_0(t) & \text{for} \quad t_0 - \delta \leqq t \leqq t_0 \\ x_0 + \int_{t_0}^{t} f(s, x_\varepsilon(s - \varepsilon))\, ds & \text{for} \quad t_0 \leqq t \leqq t_0 + \alpha \end{cases}$$

This formula defines $x_\varepsilon(t)$ first in the interval $[t_0, t_0 + \alpha_1]$, where $\alpha_1 = \min(\alpha, \varepsilon)$ The substitution is allowed, $f(s, x_\varepsilon(s - \varepsilon))$ is a weakly continuous function of s, hence it is weakly integrable. Let $x^* \in E$, $\|x^*\| = 1$ be arbitrary. For $t_0 \leqq t \leqq t_0 + \alpha_1$

$$x^*(x_\varepsilon(t) - x_0) = \int_{t_0}^{t} x^*(f(s, x_\varepsilon(s - \varepsilon)))\, ds,$$

whence

$$|x^*(x_\varepsilon(t) - x_0)| \leqq M \cdot |t - t_0| \leqq M \cdot \alpha \leqq M \cdot \frac{b}{M} = b$$

Hence by Chapter III, 5, $\|x_\varepsilon(t) - x_0\| \leqq b$ for $t_0 \leqq t \leqq t_0 + \alpha_1$. It follows that $x_\varepsilon(t)$ may be extended to the interval $[t_0 - \delta, t_0 + \alpha_2]$ where $\alpha_2 = \min(\alpha, 2\varepsilon)$. We extend the definition of $x_\varepsilon(t)$ in the same way to $[t_0 - \delta, t_0 + \alpha]$, conforming the relation $\|x_\varepsilon(t) - x_0\| \leqq b$. It follows, that the function family $\{x_\varepsilon(t)\}$; $0 < \varepsilon \leqq \delta$ is bounded as it was required in Theorem 10 in this Chapter. In addition, we shall point out its equicontinuity in the closed interval $[t_0 - \delta, t_0 + \alpha]$. To this end let $x^* \in E^*$, $\|x^*\| = 1$ arbitrary and consider the formula of $x_\varepsilon(t)$.

$$x_\varepsilon(t_1) - x_\varepsilon(t_2) = \int_{t_1}^{t_2} f(s, x_\varepsilon(s - \varepsilon))\, ds$$

Hence

$$|x^*(x_\varepsilon(t_1) - x_\varepsilon(t_2))| \leqq M|t_1 - t_2|$$

if

$$t_1, t_2 \in [t_0, t_0 + \alpha].$$

In the interval $[t_0-\delta, t_0]$ all $x_\varepsilon(t)$ equals to $x_0(t)$. The weak derivative of the latter is bounded by the assumption, hence $|x^*(x_\varepsilon(t_1)-x_\varepsilon(t_2))| \leqq M\cdot|t_1-t_2|$. In both cases from III. 5 it follows

$$\|x_\varepsilon(t_1)-x_\varepsilon(t_2)\| \leqq M|t_1-t_2|$$

It is an easy consequence of this inequality that for $\eta>0$ arbitrary, $\eta/2M$ possesses the property required from δ in 9. Definition of this Chapter, hence the assertion is proved. Applying theorem 10, we get a sequence $\varepsilon(1), \varepsilon(2), \ldots$ for which $n\to\infty$ implies $\varepsilon(n)\to 0$, and $x_{\varepsilon(n)}(t)$ is weakly uniformly fundamental on $[t_0-\delta, t_0+\alpha]$. Hence by the 8. Theorem there exists a weakly continuous function $x(t)$, so that $x_{\varepsilon(n)}(t)$ converges weakly uniformly to $x(t)$, as $n\to\infty$. All closed spheres in E are weakly compact, much rather weakly closed, hence for all t, considered above $\|x(t)-x_0\|\leqq b$. Hence $f(s, x(s))$ exists and it is weakly continuous.

Regarding the formula for $x_\varepsilon(t)$, it is to be pointed out, that for $x^*\in E^*$, $t_0\leqq t\leqq$ $\leqq t_0+\alpha$ arbitrary, $n\to\infty$ implies

$$\int_{t_0}^{t} x^*(f(s, x_{\varepsilon(n)}(s-\varepsilon(n))))\,ds \to \int_{t_0}^{t} x^*(f(s, x(s)))\,ds$$

Let x^* and t be fixed. The integrand converges pointwise, and can be majorized with a universal constant, hence our assertion follows from Lebesgue's classical theorem.

Corollary: If we suppose the "strong-weak" continuity of $f(t, x)$, then it follows that the solution constructed above is differentiable in the norm-topology (strongly). Hence we can get a "strong" existence theorem too.

Remark: It is easy to prove that the weak solutions are always strongly continuous. This remains true in non-reflexiv Banach spaces too if we suppose the boundedness of f.

REFERENCES

[1] Hille, E.—Philips, R.: *Functional analysis and semi-groups,* Amer. Math. Soc. Coll. Publ. XXXI (1957).

[2] Красноселский—Крейн: Нелокальные теоремы существования и теоремы единственности, ДАН СССР, 1955 101 № 1.

[3] Dunford, L. Schwartz: *Linear operators* I. Interscience Publ. (1964)

[4] Tonelli, L.: Sulle equazioni funzionali del tipo di Volterra, *Bull. Calcutta Math. Soc.* **20** (1928).

Mathematical Institute of the Hungarian Academy of Sciences

(Received February 11, 1970.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 205—209.

ON STRUCTURE SPACES

by

R. WIEGANDT

To the memory of Professor Alfréd Rényi

1. In [3] SULIŃSKI has given a category theoretical generalization of (strong) structure spaces. Structure spaces are well-known in the ring theory as well as in the theory of Banach algebras, further also the 0-dimensional compact spaces can be considered as structure spaces of Boolean algebras. SULIŃSKI has defined a closure operation on structure spaces on the usual way, but it remained the question whether the structure space has become a topological T_1-space. The purpose of this paper is to give an affirmative answer of SULIŃSKI's question.

In section 2 we shall define structure spaces on lattices, further we shall establish a condition which will involve that a structure space shall be a T_1-space.

Applying this condition, in section 3 we obtain that SULIŃSKI's structure spaces are always topological T_1-spaces.

I am grateful to Dr. E. T. SCHMIDT for his valuable remarks and for simplifying the proof of the Theorem.

2. Let L be a complete lattice. Denote M a set of dual atoms of L. An element $c \in L$ will be called *M-representable* if c can be represented as a complete intersection of elements belonging to M. Let L_M denote the sublattice generated by all M-representable elements.

A closure operation can be introduced into M as follows. For any subset N of M, let the *M-closure* $\mathrm{Cl}(N)$ consist of all elements $m \in M$ having property $m \geqq \bigcap_{n \in N} n$. In what follows, the *intersection* $\bigcap_{n \in N} n$ *will be denoted by* $\hat{N}$.

It is obvious that $N \to \mathrm{Cl}(N)$ is a closure operation, moreover, since the elements of M are dual atoms, therefore the M-closure of a single point is itself. The set M equipped with the M-closure operation will be called the *structure M-space of the lattice L*. It remains the question when the M-closure operation is topological, i.e. for any two subsets $N_1, N_2 \subseteqq M$ the relation $\mathrm{Cl}(N_1) \cup \mathrm{Cl}(N_2) = \mathrm{Cl}(N_1 \cup N_2)$ holds. Clearly a structure M-space is a T_1-space if and only if the M-closure operation is topological.

THEOREM. *If the sublattice L_M is modular and every element $m \in M$ has at most one relative complement, then the structure M-space is a topological T_1-space.*

PROOF. We have to show that $\mathrm{Cl}(N_1) \cup \mathrm{Cl}(N_2) = \mathrm{Cl}(N_1 \cup N_2)$ holds for any two subsets $N_1, N_2 \subseteqq M$. Clearly $\mathrm{Cl}(N_1) \cup \mathrm{Cl}(N_2) \subseteqq \mathrm{Cl}(N_1 \cup N_2)$ is always valid. Let us assume $\mathrm{Cl}(N_1) \cup \mathrm{Cl}(N_2) \subset \mathrm{Cl}(N_1 \cup N_2)$. Now there exists an element $m \in \mathrm{Cl}(N_1 \cup N_2)$ such that $m \notin \mathrm{Cl}(N_1) \cup \mathrm{Cl}(N_2)$. From $\mathrm{Cl}(N_1) \cup \mathrm{Cl}(N_2) \subset$ $\subset \mathrm{Cl}(N_1 \cup N_2)$ it follows that neither $\hat{N}_1 \leqq \hat{N}_2$ nor $\hat{N}_2 \leqq \hat{N}_1$ hold.

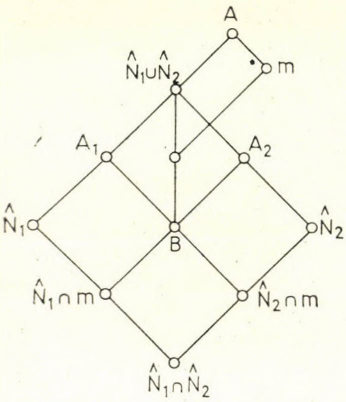


Fig. 1

In the most general case m, $\hat{N}_1$ and $\hat{N}_2$ generate a sublattice given by Fig. 1 such that $A_i=(\hat{N}_j\cap m)\cup\hat{N}_i$, $i=1,2$; $i\neq j$, and by the modularity

$$B=A_1\cap A_2=((\hat{N}_2\cap m)\cup\hat{N}_1)\cap((\hat{N}_1\cap m)\cup\hat{N}_1)=$$
$$=(\hat{N}_1\cap m)\cup(\hat{N}_2\cap m).$$

Let us observe that some elements of Fig. 1 may coincide. Since $\hat{N}_i\leqq A_i$ and $\hat{N}_i\not\leqq m$, so $A_i\not\leqq m$ follows for $i=1,2$. Hence $m\in M$ implies $A_i\cup m=1$. On the other hand by the modularity we have

$$m\cap A_i=m\cap(\hat{N}_i\cup(\hat{N}_j\cap m))=$$
$$=(m\cap\hat{N}_1)\cup(m\cap\hat{N}_2)=B.$$

Thus A_1 as well as A_2 are relative complements of m. To get a contradiction we have to show that A_1 and A_2 are different elements.

By $\hat{N}_i\not\leqq m$ it follows $m\cap\hat{N}_i<\hat{N}_i$, further $\hat{N}_i\cap(m\cap\hat{N}_j)=\hat{N}_1\cap\hat{N}_2$ holds for $i=1,2$; $i\neq j$. Hence the modularity implies that $(\hat{N}_i\cap m)\cup\hat{N}_j$ and $\hat{N}_1\cup\hat{N}_2$ are different elements i.e. $A_i=(\hat{N}_j\cap m)\cup\hat{N}_i<\hat{N}_1\cup\hat{N}_2$. Taking into account $\hat{N}_i\leqq A_i<$ $<\hat{N}_1\cup\hat{N}_2$, it follows $A_1\neq A_2$ and the Theorem is proved.

3. Now we turn to answer SULIŃSKI's question. We shall adopt the category theoretical notions and notations of [3] (see also [2]), moreover, let us assume that the considered category $\mathscr{C}$ satisfies all the conditions (C_1)—(C_{10}) of [3]. (We omit to recall them.) In such a category both of the Noetherian Isomorphism Theorems hold in the following sense.

1. *Consider the commutative diagram*

$$\begin{array}{ccccc} k & \xrightarrow{\varkappa_1} & d & \xrightarrow{\mu} & m \\ \| & & \downarrow\delta & & \downarrow\chi \\ k & \xrightarrow{\varkappa} & a & \xrightarrow{\alpha} & b \\ & & & & \downarrow\beta \\ & & & & c \end{array}$$

where $\varkappa_1$, $\varkappa$, δ, χ *are normal monomorphisms and* α, β, μ *normal epimorphisms such that* $\operatorname{Ker}\alpha=(k,\varkappa)$, $\operatorname{Ker}\beta=(m,\chi)$. *Then* $\operatorname{Ker}\mu=(k,\varkappa_1)$, *moreover* $\operatorname{Ker}\alpha\beta=(d,\delta)$ *and* $\alpha\beta$ *is a normal epimorphism.* (Cf. [5] Theorem 2, 1; the last statement is an easy consequence of [2] 9. 8).

2. *Consider the commutative diagram*

$$\begin{array}{ccccc} k & \xrightarrow{\varkappa_1} & d_1 & \xrightarrow{\beta} & m_1 \\ \varkappa_2\downarrow & & \downarrow\delta_1 & & \\ d_2 & \xrightarrow{\delta_2} & a & \xrightarrow{\alpha} & m \end{array}$$

where $\varkappa_1$, $\varkappa_2$, δ_1, δ_2 *are normal monomorphisms and* α, β *normal epimorphisms such that* $(k,\varkappa_1\delta_1)=(d_1,\delta_1)\cap(d_2,\delta_2)$ *and* $(d_1,\delta_1)\cup(d_2,\delta_2)=(a_1\varepsilon_a)$. *Then there exists* *an equivalence* $\mu: m_1 \to m$ *with commutativity preserved in the diagram* (Cf. [3] Proposition 2, 4).

Let $\mathcal{M}$ be an abstract class of simple objects such that $a \in \mathcal{M}$ and $a \approx b$ imply $b \in \mathcal{M}$. If $\alpha: a \to b$ is a normal epimorphism such that $b \in \mathcal{M}$, then the ideal $(m, \mu) = \operatorname{Ker} \alpha$ is called an *$\mathcal{M}$-maximal ideal* of a. The set of all $\mathcal{M}$-maximal ideals forms the *structure $\mathcal{M}$-space* of the object a and it is denoted by M_a. In [3] it is supposed:

($\ast$) *If* (p, π) *is an ideal of* a *and* $p \in \mathcal{M}$, *then there exists a unique $\mathcal{M}$-maximal ideal* (m, μ) *such that* $(p, \pi) \cap (m, \mu) = (0, \omega)$ *and* $(p, \pi) \cup (m, \mu) = (a, \varepsilon_n)$.

Let us mention that in [3] it is shown: by a suitable choice of $\mathcal{M}$ the category of rings, alternative rings and Lie rings satisfies condition ($\ast$).

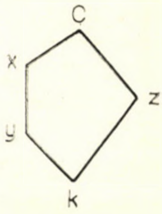
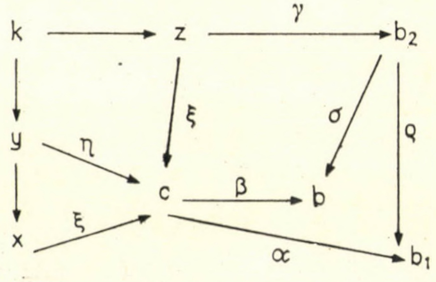


Fig. 2

PROPOSITION 1. *The lattice of all ideals of an object* $a \in \mathcal{C}$ *is modular.*

This statement follows from the Second Isomorphism Theorem. In contrary, suppose that there exists a nonmodular sublattice given by Fig. 2. Here (c, γ) is an ideal of a, moreover, there exist such normal monomorphisms ξ, η, ζ that (x, ξ), (y, η) and (z, ζ) are ideals of c. By the Second Isomorphism Theorem we obtain the commutative diagram



where α, β, γ are normal epimorphisms and ϱ, σ equivalences. Since β is a normal epimorphism, therefore by [2] 9, 8 there exists a map $\pi: b_2 \to b_1$ such that $\beta\pi = \alpha$. This implies that π is a normal epimorphism. Since $\sigma^{-1} = \pi\varrho^{-1}$, so π is a monomorphism too and hence an equivalence. From this $\xi\beta = \xi(\alpha\pi^{-1}) = \omega$ follows, and so there exists a map $\xi_1: x \to y$ such that $\xi_1\eta = \xi$. Hence $(x, \xi) \leqq (y, \eta)$ holds, contradicting our assumption.

Let us remark that Proposition 1 is essentially the converse statement of DEDEKIND's transponation principle (cf. [1] Theorem V. 6 or [4] Theorem 3 6).

The ideals being complete intersections of $\mathcal{M}$-maximal ideals, are called *M-representable ideals.* Again let L_M denote the lattice generated by all M-representable ideals of an object a.

PROPOSITION 2. *An element* $(m, \mu) \in M_a$ *has at most one relative complement in* L_M.

PROOF. Suppose that there exists an element $(m, \mu) \in M_a$ having two relative complements (x, ξ) and (y, η). Since by Proposition 1 L_M is modular, so it follows easily

$$(x, \xi) \cap (m, \mu) = (y, \eta) \cap (m, \mu) = (x, \xi) \cap (y, \eta) = (k, \varkappa).$$

Further by the Second Isomorphism Theorem the diagram

$$\begin{array}{ccccc} k & \rightarrow & x & \rightarrow & b_x \\ \downarrow & & \downarrow & & \downarrow\uparrow \\ m & \rightarrow & a & \rightarrow & b \end{array}$$

is commutative, and so $b_x \in \mathcal{M}$ holds.

(a) Assume $(x, \xi) \cup (y, \eta) = (a, \varepsilon_a)$. By the First Isomorphism Theorem the diagram

$$\begin{array}{ccccc} k & \rightarrow & x & \rightarrow & b_x \\ \| & & \downarrow & & \downarrow \chi \\ k & \rightarrow & a & \xrightarrow{\gamma} & c \\ & & & \searrow & \downarrow \\ & & & & d \end{array}$$

is commutative, further (b_x, χ) is an ideal of c. Consider the images (m_c, μ_c) and (y_c, η_c) of the ideals (m, μ) and (y, η) by the epimorphism γ. Now we have

$$(b_x, \chi) \cup (m_c, \mu_c) = (b_x, \chi) \cup (y_c, \eta_c) = (c, \varepsilon_c).$$

Hence by condition $(*)$ it follows $(m_c, \mu_c) = (y_c, \eta_c)$ which implies the contradiction $(m, \mu) = (y, \eta)$.

(b) If $(x, \xi) \cup (y, \eta) = (z, \zeta) < (a, \varepsilon_a)$, then there are normal monomorphisms ξ_z, η_z such that (x, ξ_z) and (y, η_z) are ideals of the object z. Since by Proposition 1 L_M is modular, so we have

$$(m_z, \nu) = (z, \zeta) \cap (m, \mu) \neq (x, \xi) \cap (y, \eta).$$

Moreover, there exists a normal monomorphism ν_z such that (m_z, ν_z) is an ideal of z. For $(k, \varkappa_z) = (x_z, \xi_z) \cap (y, \eta_z)$ the Second Isomorphism Theorem implies that the diagram

$$\begin{array}{ccccc} k & \rightarrow & x & \rightarrow & b_x \\ \downarrow & & \downarrow & & \downarrow\uparrow \\ m_z & \rightarrow & z & \rightarrow & b_z \end{array}$$

is commutative. Since $b_x \in \mathcal{M}$, therefore (m_z, ν_z) is an $\mathcal{M}$-maximal ideal of z. Hereby the $\mathcal{M}$-maximal ideal (m_z, ν_z) has two relative complements (x, ξ_z) and (y, η_z) such that $(x, \xi_z) \cup (y, \eta_z) = (z, \varepsilon_z)$, and we have traced case (b) back to case (a).

The Theorem and Proposition 1, 2 yield immediately

COROLLARY. *The structure $\mathcal{M}$-space of any object of $\mathcal{C}$ is a T_1-space.*

Remark: By the Corollary the assumption of [3] Theorem 5, 12 that the structure $\mathcal{M}$-space should be a T_1-space, has become superfluous.

REFERENCES

[1] BIRKHOFF, G.: *Lattice theory*, Providence 1948.
[2] KUROSCH, A. G.—LIWSCHITZ A. CH.—SCHULGEIFER E. G.—ZALENKO M. S.: *Zur Theorie der Kategorien* Berlin, 1963.
[3] SULIŃSKI, A.: The Brown-McCoy radical in categories, *Fund. Math.*, **59** (1966), 23—41.
[4] SZÁSZ, G.: *Introduction to lattice theory*, Budapest, 1963.
[5] WIEGANDT, R.: Radical and semi-simplicity in categories, *Acta Math. Acad. Sci. Hung.*, **19** (1968), 345—364.

Mathematical Institute of the Hungarian Academy of Sciences

(Received February 24, 1970.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 211—233.

ON THE DISTRIBUTION OF ADDITIVE AND THE MEAN VALUES OF MULTIPLICATIVE ARITHMETIC FUNCTIONS

by

G. HALÁSZ

1. Let the number theoretic function $g(n)$ be additive $\big(g(mn) = g(m)+g(n)$ for $(m, n)=1\big)$ and integral valued. We are going to study its "local" distribution, i.e. the behavior of

$$N(m, x) = \sum_{\substack{n \leq x \\ g(n)=m}} 1$$

(m is an integer). "Global" distribution in the sense of probability theory of additive (not necessarily integral valued) functions has been extensively investigated; a general theory is laid down in KUBILIUS's book [1].

Our problem is not new either. LANDAU [2] derived from the prime number theorem

$$N(m, x) \sim x \frac{(\log \log x)^{m-1}}{(m-1)! \log x}$$

for fixed m if $g(n)$ is $\omega(n)$, or $\Omega(n)$, the number of distinct and the number of all prime divisors, respectively. Much later SATHE [3] obtained similar formuli valid up to $m \leq c \log\log x$ $\big(c<2$ for $\Omega(n)$, $c<e$ for $\omega(n)$, the latter extended afterwards by A. SELBERG [4] for any value of $c\big)$ and independently Erdős [5] for a smaller interval. Sharp upper bounds had been known before for all m, HARDY and RAMANUJAN [6], and also lower bounds by ERDŐS and PILLAI in certain neighbourhoods of the average loglog x. (See [7].)

RÉNYI [8] considered $g(n) = \Omega(n)-\omega(n)$ proving the existence of the asymptotic densities

$$d(m) = \lim_{x\to\infty} \frac{N(m, x)}{x}, \quad \sum_{m=0}^{\infty} d(m) = 1$$

This function is characterized essentially by $g(p)=0$ for all primes p, and it was this case that first lent itself for generalization: KUBILIUS [1] extended RÉNYI's theorem to any function satisfying

$$\text{(1)} \qquad \sum_{g(p)\neq 0} \frac{1}{p} < +\infty$$

For larger exceptional sets the picture is entirely different as we shall presently see.

Recently KUBILIUS was able to generalize the case of $\omega(n)$ and $\Omega(n)$ $\big($i.e. $g(p)=1\big)$ as well, again by admitting an exceptional set of primes, this time an even sparser one:

$$\sum_{g(p)\neq 1} \frac{\log p}{p} < +\infty$$

Under certain additional conditions he obtains formuli for $m-\log\log x = O\big(\sqrt{\log\log x \ \log\log\log x}\big)$ or even $m-\log\log x = o(\log\log x)$ (See [9]).

In this paper we combine the above two cases by considering functions assuming 1 on an arbitrary set $\mathfrak{P}$ of primes and 0 at all the other primes. For the sake of simplicity only, we suppose $g(n)$ completely additive.

THEOREM 1. *Let $g(n)$ denote the number of prime divisors $p \in \mathfrak{P}$ of n, counted with proper multiplicity. With the definition*

$$E(x) = \sum_{\substack{p \leqq x \\ p \in \mathfrak{P}}} \frac{1}{p}$$

we have

$$N(m,x) = \sum_{\substack{n \leqq x \\ g(n)=m}} 1 = x \frac{E^m(x)}{m!} \exp\{-E(x)\} \left(1 + O\left(\frac{|m-E(x)|}{E(x)}\right) + O\left(\frac{1}{\sqrt{E(x)}}\right)\right)$$

uniformly in m and x for $\delta \leqq \dfrac{m}{E(x)} \leqq 2-\delta$, $E(x) \geqq 2$, with any fixed $\delta > 0$.

It should be noted that, contrary to the earlier results, neither the "sparseness" of $\mathfrak{P}$ nor that of its complement is assumed here. Should $E(x)$ tend to ∞, and in view of KUBILIUS's generalization (1) of RÉNYI's theorem this is the case left open, our theorem gives asymptotic relation whenever $m-E(x) = o(E(x))$. As the constants involved in the symbols $O(\ )$ are absolute constants depending only on δ, it also yields sharp lower bounds for the larger range $|m-E(x)| \leqq cE(x)$; the exact value of the universal constant c remains unknown. On the other hand, our formula supplies sharp upper bounds for the whole $\delta E(x) \leqq m \leqq (2-\delta)E(x)$.* This $2-\delta$ is caused by complete additivity and the prime 2, so to speak and is of less importance here. It will be clear from the proof that when prime divisors are counted without multiplicity it can be replaced by any number and also that one can get uniform bounds for all m contending with less sharp ones outside the above interval.

The author owes the problem to Professor ERDŐS. Actually he asked if

$$N(m+1,x) \sim N(m,x) \qquad \text{for } m-E(x) = o(E(x))$$

which is obvious from Theorem 1. His old conjecture, now WIRSING's theorem (see below) concerning multiplicative functions (i.e. $f(mn)=f(m)f(n)$ for $(m,n)=1$) with $f(n)=\pm 1$ would follow easily and he told the author he had tried to prove the conjecture this way.

Another connection with multiplicative functions is due to A. SELBERG [4]. In his analytic proof of the SATHE—ERDŐS formula he considers the multiplicative $f(n)=z^{g(n)}$ and represents $N(m,x)$ as

$$\sum_{\substack{n \leqq x \\ g(n)=m}} 1 = \frac{1}{2\pi i} \int\limits_{|z|=r} \frac{M(x,z)}{z^{m+1}}\, dz,$$

where

$$M(x,z) = \sum_{n \leqq x} f(n) = \sum_{n \leqq x} z^{g(n)} = \sum_{m=0}^{\infty} N(m;x)\, z^m,$$

* For some complementary results see the author's forthcoming paper in Acta Sci. Math. Hung.

thus reducing the problem to the summation of multiplicative functions. While, however, the special $\omega(n)$ and $\Omega(n)$ led to special function $f(n)$, we are led to general $f(n)$.

2. So we assume $f(n)$ to be completely multiplicative, $f(mn)=f(m)f(n)$ for all m, n and seek estimates or asymptotics for

$$\sum_{n\leqq x} f(n)$$

in terms of a majorant of $f(n)$. This problem was already considered by WIRSING in [10]. Here, for simplicity, we take only $|f(n)|$ as a majorant. In this case essentially under the condition

$$|e^{i\vartheta_p}-e^{i\vartheta_0}| \geqq \delta, \quad e^{i\vartheta_p} \stackrel{\text{def}}{=} \frac{f(p)}{|f(p)|}$$

for fixed ϑ_0, δ and all primes p, WIRSING proved very deep and general results of this type, but his formuli were just of the form

$$\sum_{n\leqq x} f(n) = (C+o(1))A(x),$$

clearly insufficient for our application. (Let us mention in passing that the case $C\neq 0$, at least when the majorant is 1, $|f(n)|\leqq 1$, had been completely solved before by DELANGE [11] and that afterwards in [12] we, again for the majorant 1 settled the case of any C, removing WIRSING's condition.) Therefore, we set out to estimate the errors. It would not be difficult to do it in complete generality as in [12], however, having in mind our application, we restrict ourselves to WIRSING's case. Our results are still more general than we actually need.

THEOREM 2. *Let $f(n)$ be completely multiplicative, $0<\delta\leqq|f(p)|\leqq 2-\delta$ and with the definition $\vartheta_p=\arg f(p)$*

$$|e^{i\vartheta_p}-e^{i\vartheta_0}| \geqq \delta$$

for fixed δ, ϑ_0 and all p. Then

$$\sum_{n\leqq x} f(n) = O\left(x \exp\left\{\sum_{p\leqq x} \frac{|f(p)|-1}{p} - c_1 \sum_{p\leqq x} \frac{|f(p)|-\operatorname{Re} f(p)}{p}\right\}\right).$$

Here and in that follows $c_1, c_2, \ldots$ are positive constants depending only on δ or are universal if δ is irrelevant. (The same is true for constants involved in $O(\;)$). This latter is the case e.g. when $f(n)=\pm 1$ and we get the conjecture of ERDŐS, now WIRSING's theorem of which we spoke earlier in an explicit form. We remark that we could then take $c_1=0{,}07$ but not $0{,}37$.

Corresponding to DELANGE's theorem with non-vanishing main term we state

THEOREM 3. *If $f(n)$ is completely multiplicative, $|f(p)-1|\leqq\eta<1$ for all primes then*

$$\sum_{n\leqq x} f(n) = x\exp\left\{\sum_{p\leqq x}\frac{f(p)-1}{p}\right\}+$$

$$+O(x)\left[\eta\exp\left\{\sum_{p\leqq x}\frac{\operatorname{Re} f(p)-1}{p}\right\}+\left(e^{-c_2/\eta}+\frac{1}{\log^{c_3} x}\right)\exp\left\{\sum_{p\leqq x}\frac{|f(p)|-1}{p}\right\}\right].$$

These are by no means the first error estimations for general multiplicative functions, (see e.g. [13], [14]). While, however, in the earlier results sharper asymptotic relations have been derived from rather regular behaviors of $f(p)$, here especially in Theorem 2, we are given an irregular distribution of $f(p)$ and are to find estimates in terms of some simple quantities.

As to the method of proof, we shall modify our analytic method of [12]. The original form would only work in the case $|f(p)|>\frac{1}{2}$ and therefore we now combine it with some elementary argument taken partly from WIRSING [10], although the proof remains basicly analytic and still much simpler than WIRSING's. Since we, too, assume his condition on the ϑ_p's we do not claim that our results cannot be proved by his elementary method. (According to a remark in [10], they can in a special case.) It will be clear from the proof, however, how to remove his condition, giving the behavior of the sum function in general with precise explicit extimations in terms of some more complicated expressions than in Theorems 2 and 3.

Our analytic approach depends of course on

$$F(s)=\sum_{n=1}^{\infty}\frac{f(n)}{n^s}\qquad (s=\sigma+it)$$

for which by multiplicativity

(2)

$$F(s)=\prod_p\left(1+\frac{f(p)}{p^s}+\frac{f^2(p)}{p^{2s}}+\cdots\right)=\prod_p\frac{1}{1-\frac{f(p)}{p^s}}=e^{\sum\limits_p\log\frac{1}{1-\frac{f(p)}{p^s}}}=e^{\sum\limits_{n=1}^{\infty}\frac{\lambda(n)f(n)}{n^s}}$$

where by $|f(p)|<2$ the product, hence the series converge absolutely for $\sigma=\operatorname{Re}s>1$ representing an analytic function there;

$$\lambda(n)=\begin{cases}\frac{1}{k} & \text{if}\quad n=p^k\\ 0 & \text{otherwise.}\end{cases}$$

We shall also use the symbol

$$\Lambda(n)=\lambda(n)\log n=\begin{cases}\log p & \text{if}\quad n=p^k\\ 0 & \text{otherwise.}\end{cases}$$

In order not to interrupt the proof we gather here some easy estimations.

(3)

$$\sum_{\substack{n=p^k\\ k\geq 2}}\frac{\Lambda(n)|f(n)|}{n^\sigma}\leq\sum_p\operatorname{lop}p\sum_{k=2}^{\infty}\frac{(2-\delta)^k}{p^k}\leq 4\sum_p\frac{\log p}{p(p-2+\delta)}=c_3<+\infty\qquad(\sigma\geq 1)$$

and also in the case $|f(p)-1|\leq\eta$ (Theorem 3), using $|z^k-1|\leq k|z-1|(2-\delta)^k$

for $|z| \leqq 2-\delta$

(4) $$\sum_{\substack{n=p^k\\k\geqq 2}} \frac{\Lambda(n)|f(n)-1|}{n^\sigma} \leqq \eta \sum_p \log p \sum_{k=2}^{\infty} \frac{(2-\delta)^k k}{p^k} = O(\eta) \qquad (\sigma \geqq 1).$$

(5) $$\sum_{\substack{n=p^k\leqq x\\k\geqq 2}} \Lambda(n)|f(n)| \leqq \sum_{k=2}^{\left[\frac{\log x}{\log 2}\right]} (2-\delta)^k \sum_{p\leqq x^{1/k}} \log p \leqq c_4 \sum_{k=2}^{\left[\frac{\log x}{\log 2}\right]} (2-\delta)^k x^{1/k} \leqq$$
$$\leqq c_4 \frac{\log x}{\log 2}\left((2-\delta)^2 x^{1/2} + 2x^{\frac{\log(2-\delta)}{\log 2}}\right) \leqq c_5 x^{1-c_6},$$

the general term being a convex function of k, thus assuming its maximum for the extreme values of k.

(6) $$\sum_{p>x} \frac{1}{p^\sigma} \leqq c_7 \qquad \left(\sigma \geqq 1+\frac{1}{2\log x}\right).$$

For, denoting x^{2^l} by x_l, the sum is

$$\sum_{l=1}^{\infty} \sum_{x_{l-1}<p\leqq x_l} \frac{1}{p^{\sigma-1}p} \leqq \sum_{l=1}^{\infty} x_{l-1}^{1-\sigma} \sum_{x_{l-1}<p\leqq x_l} \frac{1}{p} \leqq$$
$$\leqq c_8 \sum_{l=1}^{\infty} e^{-2^{l-1}\log x(\sigma-1)} \leqq c_8 \sum_{l=1}^{\infty} e^{-2^{l-2}} = c_7.$$

Here we have made use of the elementary

$$\sum_{x_{l-1}<p\leqq x_l} \frac{1}{p} = \log\log x_l - \log\log x_{l-1} + o(1) \sim \log 2.$$

(7) $$\sum_{p\leqq x_0} \left|\frac{1}{p}-\frac{1}{p^s}\right| \leqq c_9 \log\left(2+\frac{|s-1|}{\sigma_0-1}\right) \quad \left(\sigma_0 = 1+\frac{1}{\log x_0}, \quad \sigma = \operatorname{Re} s > 1\right)$$

and in particular

(8) $$\sum_{p\leqq x_0} \left(\frac{1}{p}-\frac{1}{p^\sigma}\right) \leqq c_{10} \quad \left(\sigma \leqq 1+\frac{3(\sigma_0-1)}{2}\right).$$

For

$$\sum_{p\leqq x_0} \frac{1}{p}\left|1-\frac{1}{p^{s-1}}\right| = \sum_{p\leqq x_0} \frac{1}{p}|1-e^{-(s-1)\log p}| \leqq$$
$$\leqq |s-1| \sum_{p\leqq x_0} \frac{\log p}{p} \leqq c_{11}|s-1|\log x_0 \leqq 2c_{11},$$

if $|s-1| \leqq 2(\sigma_0-1)$, say. If $|s-1| > 2(\sigma_0-1)$, the same way for $|s-1|\log p \leqq 1$, i.e. $p \leqq \exp\left\{\frac{1}{|s-1|}\right\} \overset{\text{def}}{=} a$

$$\sum_{p\leqq a} \frac{1}{p}\left|1-\frac{1}{p^{s-1}}\right| \leqq |s-1| \sum_{p\leqq a} \frac{\log p}{p} \leqq c_{11}|s-1|\log a = c_{11},$$

while for $p>a$ the trivial estimation $\left|1-\frac{1}{p^{s-1}}\right| \leqq 2$ is better:

$$\sum_{a<p\leqq x_0} \frac{2}{p} \leqq 2(\log\log x_0-\log\log a)+O(1) = 2\log(|s-1|\log x_0)+O(1) \leqq$$
$$\leqq c_{12}\log\frac{|s-1|}{\sigma_0-1}.$$

Combining this with the previous case we see that (7) holds in any case.

We define

$$\Phi(\sigma)=\sum_{n=1}^{\infty}\frac{|f(n)|}{n^\sigma}=e^{\sum_{n=1}^{\infty}\frac{\lambda(n)|f(n)|}{n^\sigma}} \tag{9}$$

for $|f(n)|$ is also completely multiplicative and so the exponential representation (2) holds. Taking logarithmic derivative and using (3)

$$-\frac{\Phi'}{\Phi}(\sigma)=\sum_{n=1}^{\infty}\frac{\Lambda(n)|f(n)|}{n^\sigma}=\sum_p\frac{\log p|f(p)|}{p^\sigma}+O(1)\leqq 2\sum_{n=1}^{\infty}\frac{\Lambda(n)}{n^\sigma}+O(1)=$$
$$=-2\frac{\zeta'}{\zeta}(\sigma)+O(1)\leqq\frac{2}{\sigma-1}+O(1) \tag{10}$$

and also

$$-\frac{\Phi'}{\Phi}(\sigma)\geqq\sum_p\frac{\log p|f(p)|}{p^\sigma}\geqq\delta\sum_{n=1}^{\infty}\frac{\Lambda(n)}{n^\sigma}+O(1)= \tag{11}$$
$$=-\delta\frac{\zeta'}{\zeta}(\sigma)+O(1)\geqq\frac{\delta}{\sigma-1}+O(1)$$

by elementary properties of $\zeta(\sigma)$. Here $\zeta(s)$ is the Riemann zeta function, i.e. $F(s)$ with $f(n)\equiv 1$. Integrating (10) and (11)

$$\frac{c_{13}}{(\sigma-1)^\delta}\leqq\Phi(\sigma)\leqq\frac{c_{14}}{(\sigma-1)^2}\qquad(1<\sigma\leqq 3). \tag{12}$$

Next we shall give various estimations of $F(s)$ in § 1 and § 3. In § 2 integration with respect to s will lead to Theorem 2 and in § 4 to Theorem 3. A straightforward integration with respect to z will lead to $N(m, x)$ (Theorem 1) in § 5.

§ 1. Let us compare $|F(s)|$ with the $\Phi(\sigma)$ of (9) by their exponential representations (2) and (9):

$$\frac{|F(s)|}{\Phi(\sigma)}=e^{-\sum_{n=1}^{\infty}\frac{\lambda(n)}{n^\sigma}(|f(n)|-\operatorname{Re}f(n)n^{-it})}. \tag{13}$$

The preliminary estimation (3) shows that we can drop terms with $n=p^k$, $k\geqq 2$ making only a bounded error. The sum left,

$$\sum_p\frac{|f(p)|(1-\operatorname{Re}e^{i\vartheta_p}p^{-it})}{p^\sigma}\geqq\delta\sum_p\frac{1-\operatorname{Re}e^{i\vartheta_p}p^{-it}}{p^\sigma}$$

has non-negative terms. When, however, p^{it} passes through the middle of the omitted neighbourhood of Theorem 2, $|p^{it}-e^{i\vartheta_0}| \leqq \frac{\delta}{2}$, say, then under the condition of that theorem

$$|p^{it}-e^{i\vartheta_p}| \geqq \frac{\delta}{2}, \quad |1-e^{i\vartheta_p}p^{-it}| \geqq \frac{\delta}{2}$$

and we have the sharper

$$1-\operatorname{Re} e^{i\vartheta_p}p^{-it} \geqq \frac{\delta^2}{8}.$$

Let us therefore construct a function $h(e^{i\vartheta})$ that vanishes for $|e^{i\vartheta}-e^{i\vartheta_0}| \geqq \delta/2$, is positive otherwise but never exceeds $\delta^2/8$. Then

$$1-\operatorname{Re} e^{i\vartheta_p}p^{-it} \geqq h(p^{it})$$

and we get for our sum

$$\sum_p \frac{1-\operatorname{Re} e^{i\vartheta_p}p^{-it}}{p^\sigma} \geqq \sum_p \frac{h(p^{it})}{p^\sigma}.$$

By readding terms with $n=p^k$, $k\geqq 2$ we complete the sum into

$$\sum_{n=1}^{\infty} \frac{\lambda(n)h(n^{it})}{n^\sigma}$$

making another $O(1)$ error. We now expand $h(e^{i\vartheta})$ into its Fourier series

$$h(e^{i\vartheta}) = \sum_{l=-\infty}^{\infty} a_l e^{-il\vartheta}.$$

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} h(e^{i\vartheta})\,d\vartheta$$

is a fixed positive quantity. All we need for the other coefficients is e.g.

$$a_l = O\left(\frac{1}{l^2}\right) \qquad (l \neq 1).$$

$h(e^{i\vartheta})$ can in fact be chosen to satisfy these requirements: the triangle shaped function with height $\delta^2/8$ built over the positivity interval is one possibility.

Substituting $h(e^{i\vartheta})$ by its Fourier series

$$\sum_{n=1}^{\infty} \frac{\lambda(n)h(n^{it})}{n^\sigma} = \sum_{n=1}^{\infty} \frac{\lambda(n)}{n^\sigma} \sum_{l=-\infty}^{\infty} a_l n^{-ilt} = \sum_{l=-\infty}^{\infty} a_l \sum_{n=1}^{\infty} \frac{\lambda(n)}{n^{\sigma+ilt}} = \tag{14}$$

$$= \sum_{l=-\infty}^{\infty} a_l \log\zeta(\sigma+ilt)$$

by the exponential form (2) for $\zeta(s)$. (The interchange of the summation signs is justified by absolute convergence.) Now, by the pole at $s=1$

$$\log\zeta(\sigma)=\log\frac{1}{\sigma-1}+O(1)$$

$$\log\zeta(\sigma+ilt)=\log\frac{1}{\sigma-1+ilt}+O(1)=-\log(\sigma-1+|lt|)+O(1)=$$

$$=-\log(\sigma-1+|t|)+O\big(\log(|l|+1)\big)\qquad(l\neq 0)$$

provided $|lt|\leqq 3$, say. If, however, we first let $|t|\leqq 3$, the last equality holds for $|lt|>3$ as well: the $O(\ )$ term dominates over the main term $\left(\text{since}\log|l|\geqq\log\frac{3}{|t|}\right)$ and owing to estimation of $\zeta(s)$ valid for large values of the argument we in fact have

$$\log\zeta(\sigma+ilt)=O(\log\log|lt|)=O(\log\log 3|l|)=O\big(\log(|l|+1)\big)$$

uniformly for $\sigma>1$. From the rapid convergence of a_l we see that the contribution of $O\big(\log(|l|+1)\big)$ is bounded in (14). The main terms give

$$a_0\log\frac{1}{\sigma-1}-\log(\sigma-1+|t|)\sum_{l\neq 0}a_l.$$

The sum here, completed by a_0, is just the value $h(e^{i0})=h(1)$ that is non-negative. The factor $\log(\sigma-1+|t|)$ is bounded above (if e.g. $\sigma\leqq 3$ for now $|t|\leqq 3$). Thus we get further

$$\text{(15)}\qquad \geqq a_0\log\frac{1}{\sigma-1}+a_0\log(\sigma-1+|t|)-h(1)\log 4=a_0\log\left(1+\frac{|t|}{\sigma-1}\right)+O(1)$$

For $|t|>3$

$$\log\zeta(\sigma+ilt)=O(\log\log|lt|)=O\big(\log\log(|l|+2)\big)+O(\log\log|t|)$$

for all $l\neq 0$, hence in this case (14) becomes

$$a_0\log\frac{1}{\sigma-1}+O(1)\sum_{l\neq 0}\frac{\log\log(|l|+2)}{l^2}+O(\log\log|t|)\sum_{l\neq 0}\frac{1}{l^2}=$$

$$=a_0\log\frac{1}{\sigma-1}+O(\log\log|t|).$$

Only $|t|\leqq\frac{1}{(\sigma-1)^6}$ will be needed. Then the last term on the right is negligible here and we see that the lower bound found in (15) is valid here, too, if we replace a_0 by $a_0/8$, say. Returning to (13) and taking into account the various bounded errors made in the meantime

$$\text{(16)}\qquad \frac{|F(s)|}{\Phi(\sigma)}\leqq e^{-c_{15}\log\left(1+\frac{|t|}{\sigma-1}\right)+O(1)}\qquad\left(|t|\leqq\frac{1}{(\sigma-1)^6}\right).$$

Our next aim is to get a uniform estimate in terms of $F(\sigma)$:

$$\frac{|F(s)|}{\Phi(\sigma)} \leqq c_{16}\left(\frac{|F(\sigma)|}{\Phi(\sigma)}\right)^{c_{17}} \qquad \left(|t| \leqq \frac{1}{(\sigma-1)^6}\right). \tag{17}$$

Consider

$$\left|\frac{F(s)}{F(\sigma)}\right| = e^{\operatorname{Re}\sum_{n=1}^{\infty}\lambda(n)\left(\frac{f(n)}{n^s}-\frac{f(n)}{n^\sigma}\right)}$$

and use (3) to drop terms $n=p^k$ with $k\geqq 2$, (6) to drop terms with $k=1$, $p>x$ $\left(\frac{1}{\log x}=\sigma-1\right)$ and (7), (8) to replace s, σ by 1. After the last step the sum vanishes and the errors made are $O(1)$ and once $O\left(\log\left(2+\frac{|s-1|}{\sigma-1}\right)\right)$. Here $|s-1| \leqq$ $\leqq \sigma-1+|t|$. Hence

$$\left|\frac{F(s)}{F(\sigma)}\right| \leqq c_{18}e^{c_{19}\log\left(2+\frac{|t|}{\sigma-1}\right)}.$$

Writing

$$\frac{|F(s)|}{\Phi(\sigma)} = \left|\frac{F(s)}{F(\sigma)}\right|\frac{|F(\sigma)|}{\Phi(\sigma)}$$

we have

$$\frac{|F(s)|}{\Phi(\sigma)} \leqq c_{18}e^{c_{19}\log\left(2+\frac{|t|}{\sigma-1}\right)-\log\frac{\Phi(\sigma)}{|F(\sigma)|}}.$$

If

$$c_{19}\log\left(2+\frac{|t|}{\sigma-1}\right) \leqq \frac{1}{2}\log\frac{\Phi(\sigma)}{|F(\sigma)|}$$

we obtain (17) with $c_{17} = 1/2$. In the alternative case (17) with $c_{17} = \frac{c_{15}}{2c_{19}}$ is weaker than (16) that we have already proved.

3. Proof of Theorem 2. Instead of an exact coefficient formula, as in [12], we consider

$$I(\sigma) = \int_{(\sigma)}\left|\frac{F'(s)}{s}\right|^2|ds| = \int_{-\infty}^{\infty}\left|\frac{F'(\sigma+it)}{\sigma+it}\right|^2 dt \qquad (\sigma>1) \tag{18}$$

Its relevance will soon be clear; we first estimate it.

The infinite part $|t|\geqq T$ with T appropriately chosen is to be estimated trivially:

$$|F'(s)| = \left|\sum_{n=1}^{\infty}\frac{f(n)\log n}{n^s}\right| \leqq \sum_{n=1}^{\infty}\frac{|f(n)|\log n}{n^\sigma} = -\Phi'(\sigma) = -\frac{\Phi'}{\Phi}(\sigma)\Phi(\sigma) \leqq c_{20}\frac{\Phi(\sigma)}{\sigma-1},$$

(see preliminary estimation (10)),

$$\int_{|t|\geqq T}\left|\frac{F'(s)}{s}\right|^2|ds| \leqq \left(c_{20}\frac{\Phi(\sigma)}{\sigma-1}\right)^2 2\int_T^{\infty}\frac{dt}{t^2} = 2c_{20}^2\left(\frac{\Phi(\sigma)}{\sigma-1}\right)^2\cdot\frac{1}{T}.$$

By (12) $\Phi(\sigma)$ is at most of order $\frac{1}{(\sigma-1)^2}$ and we can e.g. take $T = \frac{1}{(\sigma-1)^6}$ to make the contribution bounded. On the finite part we write $F' = \frac{F'}{F} \cdot F$ and factor out F:

$$\int\limits_{\substack{(\sigma)\\|t|\leqq T}} \left|\frac{F'(s)}{s}\right|^2 |ds| \leqq \max_{\substack{(\sigma)\\|t|\leqq T}} |F(s)|^2 \int\limits_{(\sigma)} \left|\frac{F'}{F}(s)\right|^2 \frac{|ds|}{|s|^2}.$$

Owing to the choice of T the uniform bound (17) is valid giving

$$\max_{\substack{(\sigma)\\|t|\leqq T}} |F(s)|^2 \leqq c_{21} \left[\Phi(\sigma)\left(\frac{|F(\sigma)|}{\Phi(\sigma)}\right)^{c_{17}}\right]^2. \tag{19}$$

Taking logarithmic derivative of the exponential form (2), the integrand,

$$\frac{1}{s}\frac{F'}{F}(s) = -\frac{1}{s}\sum_{n=1}^{\infty}\frac{\Lambda(n)f(n)}{n^s}$$

can be rewritten by partial integration, introducing

$$L(u) = \sum_{n\leqq u}\Lambda(n)f(n)$$

as

$$-\frac{1}{s}\int\limits_1^{\infty}\frac{dL(u)}{u^s} = -\int\limits_1^{\infty}\frac{L(u)}{u^{s+1}}\,du = -\int\limits_0^{\infty}L(e^u)e^{-us}\,du.$$

By Parseval's formula

$$\int\limits_{(\sigma)}\left|\frac{F'}{F}(s)\right|^2\frac{|ds|}{|s|^2} = 2\pi\int\limits_0^{\infty}|L(e^u)|^2e^{-2u\sigma}\,du. \tag{20}$$

Here

$$|L(u)| = \left|\sum_{n\leqq u}\Lambda(n)f(n)\right| \leqq 2\sum_{p\leqq u}\log p + \sum_{\substack{n=p^k\leqq u\\k\geqq 2}}\Lambda(n)|f(n)| \leqq c_{22}u$$

(for the second sum see (5)), implying

$$\int\limits_0^{\infty}|L(e^u)|^2e^{-2u\sigma}\,du \leqq c_{22}^2\int\limits_0^{\infty}e^{-2u(\sigma-1)}\,du = \frac{c_{23}}{\sigma-1}$$

We have thus found for our original integral in (18)

$$I(\sigma) \leqq c_{21}\left[\Phi(\sigma)\left(\frac{|F(\sigma)|}{\Phi(\sigma)}\right)^{c_{17}}\right]^2\frac{2\pi c_{23}}{\sigma-1} + O(1). \tag{21}$$

Now, the Parseval equation applied to F' the same way as to $\frac{F'}{F}$ in (20) relates $I(\sigma)$

to our problem on coefficient sum:

(22) $$I(\sigma) = 2\pi \int_0^\infty |N(e^u)|^2 e^{-2u\sigma}\, du = 2\pi \int_0^\infty \frac{|N(u)|^2}{u^{3+2(\sigma-1)}}\, du$$

where

$$N(u) = \sum_{n \leq u} f(n) \log n$$

is the coefficient sum of $F'(s)$. Our task is now to deduce the necessary information from the sharp bound of $I(\sigma)$.

We think x_0 large but temporarily fixed, $\sigma_0 - 1 = \dfrac{1}{\log x_0}$ and consider the variable $x \leq x_0$, i.e. with $\sigma - 1 = \dfrac{1}{\log x}$ $\sigma \geq \sigma_0$. For $\log u \leq \dfrac{1}{\sigma-1} = \log x$ $\dfrac{1}{u^{2(\sigma-1)}} \geq e^{-2}$ and keeping this range only,

$$\int_1^x \frac{|N(u)|^2}{u^3}\, du \leq \frac{e^2}{2\pi} I(\sigma).$$

Schwarz's inequality implies

(23) $$\int_1^x \frac{|N(u)|}{u^2}\, du \leq \sqrt{\int_1^x \frac{|N(u)|^2}{u^3}\, du \int_1^x \frac{du}{u}} \leq c_{24} \sqrt{\frac{I(\sigma)}{\sigma-1}}$$

We want to replace $N(x)$ by

$$M(x) = \sum_{n \leq x} f(n).$$

The relation is given by

(24) $$N(x) = \int_1^\infty \log u\, dM(u) = M(x) \log x - \int_1^x \frac{M(u)}{u}\, du.$$

We can save the actual estimation of terms of secondary importance like the last one by defining

(25) $$R = \max_{2 \leq x \leq x_0} \frac{|M(x)|}{x} \log^{1-\delta_1} x$$

Here and in that follows $\delta_1, \delta_2, \ldots$ and $K_1, K_2, \ldots$ will denote positive constants to be fixed later sufficiently small and sufficiently large, respectively. With R so defined we can write (24) as

(26) $$N(x) = M(x) \log x + O\left(R \int_2^x \log^{\delta_1 - 1} u\, du\right) = M(x) \log x + O(Rx \log^{\delta_1 - 1} x)$$

and (23) takes the form

$$\int_{x^{\delta_2}}^{x} \frac{|M(u)|}{u^2}\,du \leqq \frac{1}{\delta_2 \log x} \int_1^x \frac{|M(u)|\log u}{u^2}\,du \leqq$$

$$(27)\qquad = \frac{1}{\delta_2 \log x}\left(c_{24}\sqrt{\frac{I(\sigma)}{\sigma-1}} + O(R)\int_2^x \frac{\log^{\delta_1-1} u}{u}\,du\right) \leqq$$

$$\leqq \frac{c_{25}}{\delta_2 \log x}\left(\sqrt{\frac{I(\sigma)}{\sigma-1}} + \frac{R}{\delta_1}\log^{\delta_1} x\right).$$

To be able to use this rather weak average we make another essential use of multiplicativity:

$$N(u) = \sum_{n\leqq u} f(n)\log n = \sum_{n\leqq u} f(n) \sum_{d|n} \Lambda(n) = \sum_{d\leqq u} \Lambda(d) f(d) \sum_{k\leqq \frac{u}{d}} f(k) =$$

$$(28)\qquad = \sum_{d\leqq u} \Lambda(d) f(d) M\left(\frac{u}{d}\right),$$

$$(29)\qquad |N(u)| \leqq \sum \Lambda(d)\,|f(d)|\,\left|M\left(\frac{u}{d}\right)\right|.$$

(Here we omitted $d\leqq u$ since $M\left(\frac{u}{d}\right) = 0$ for $d > u$.)

We first show that to estimate $N(x)$ it suffices to estimate

$$\int_y^x |N(u)|\,du$$

with y sufficiently close to x. In fact, for the deviation $(y\leqq u\leqq x)$

$$(30)\qquad ||N(u)|-|N(x)|| \leqq |N(u)-N(x)| \leqq \sum_{y<n\leqq x} |f(n)|\log n \leqq \log x \sum_{y<n\leqq x} |f(n)|$$

by Hölder's inequality with $\frac{1}{p}+\frac{1}{q} = 1$, $\sigma = 1+\frac{1}{\log x}$ as usual,

$$\leqq \log x \cdot x^{\sigma/q} \sum_{y<n\leqq x} \frac{|f(n)|}{n^{\sigma/q}} \leqq \log x\cdot (ex)^{1/q}(x-y+1)^{1/p}\left(\sum_{n=1}^{\infty}\frac{|f(n)|^q}{n^\sigma}\right)^{1/q}.$$

$|f(n)|^q$ here is again multiplicative and determining q so as to satisfy $(2-\delta)^q < 2$ the sum is a function like $\Phi(\sigma)$ from (9) also of order $\frac{1}{(\sigma-1)^2}$ at most (see (12)). Choosing y with $x-y = \frac{x}{\log^{c_{26}} x}$ we get further

$$\leqq c_{27} x^{1/q} x^{1/p} (\log x)^{1-c_{26}/p+2/q} = c_{27} x,$$

the value of c_{26} to be determined by $1-c_{26}/p+2/q=0$. The bound obtained for (30) means for $y \leqq u \leqq x$

$$(31) \qquad |N(u)| = |N(x)|+O(x), \quad |N(x)|(x-y) \leqq \int_y^x |N(u)|\,du + c_{27}x(x-y).$$

Putting (29) into the integral

$$\int_y^x |N(u)|\,du \leqq \sum_d \Lambda(d)|f(d)| \int_y^x \left| M\left(\frac{u}{d}\right)\right| du = \sum_d \Lambda(d)|f(d)|\,d \int_{y/d}^{x/d} |M(u)|\,du =$$

$$(32) \qquad = \int_1^x |M(u)| \Bigg(\sum_{\frac{y}{u} \leqq d \leqq \frac{x}{u}} \Lambda(d)|f(d)|\,d \Bigg) du.$$

For the sum, using (5),

$$\sum_{\alpha \leqq d \leqq \beta} \Lambda(d)|f(d)|\,d \leqq \beta\Big(2 \sum_{\alpha \leqq p \leqq \beta} \log p + c_5 \beta^{1-c_6}\Big).$$

We need the prime number theorem with logarithmic remainder term to infer

$$\sum_{\alpha \leqq p \leqq \beta} \log p \leqq c_{28}(\beta-\alpha)$$

for $\beta-\alpha \geqq \dfrac{\beta}{\log^{K_1}\beta}$ with arbitrarily large but fixed K_1. If so,

$$\sum_{\alpha \leqq d \leqq \beta} \Lambda(d)|f(d)|\,d \leqq c_{29}\beta(\beta-\alpha).$$

In our case $\alpha = \dfrac{y}{u}$, $\beta = \dfrac{x}{u}$ and the condition for applicability becomes

$$\frac{x-y}{u} \geqq \frac{\frac{x}{u}}{\log^{K_1}\frac{x}{u}}, \quad \frac{x}{\log^{c_{26}}x} \geqq \frac{x}{\log^{K_1}\frac{x}{u}}, \quad u \leqq x e^{-\log^{c_{26}/K_1}x} \stackrel{\text{def}}{=} v.$$

Integration up to this v in (32) thus gives

$$\int_1^v |M(u)| \Bigg(\sum_{\frac{y}{u} \leqq d \leqq \frac{x}{u}} \Lambda(d)|f(d)|\,d \Bigg) du \leqq c_{29}x(x-y) \int_1^v \frac{|M(u)|}{u^2}\,du.$$

Here, for $x^{\delta_2} \leqq u \leqq v$ we can make use of our basic estimate (27), while for $u \leqq x^{\delta_2}$ we apply (25), i.e.

$$|M(u)| \leqq Ru\log^{\delta_1-1}:$$

$$\int_1^v \leqq c_{29}x(x-y)\left\{\frac{c_{25}}{\delta_2 \log x}\left(\sqrt{\frac{I(\sigma)}{\sigma-1}} + \frac{R}{\delta_2}\log^{\delta_1} x\right) + R\int_1^{x^{\delta_2}} \frac{\log^{\delta_1-1}}{u}\,du\right\} \leqq$$

$$(33) \qquad \leqq c_{30}x(x-y)\left\{\frac{1}{\delta_2 \log x}\left(\sqrt{\frac{I(\sigma)}{\sigma-1}} + \frac{R}{\delta_1}\log^{\delta_1} x\right) + \frac{R}{\delta_1}\delta_2^{\delta_1}\log^{\delta_1} x\right\}.$$

It remains $v \leqq u \leqq x$. We write our quantity in the last but one form in (32) and again estimate in terms of R:

$$\sum_d \Lambda(d)|f(d)|d \int\limits_{\max\left(v,\, \frac{y}{d}\right) \leqq u \leqq \frac{x}{d}} |M(u)|\, du \leqq \sum_{d \leqq \frac{x}{v}} \Lambda(d)|f(d)|d \frac{x-y}{d} R \frac{x}{d} \log^{\delta_1 - 1} v.$$

For $K_1 > c_{26}$ $v \geqq \sqrt{x}$, say, and we get further

$$\leqq Rx(x-y) \log^{\delta_1 - 1} \sqrt{x} \sum_{d \leqq e^{\log^{c_{26}/K_1} x}} \frac{\Lambda(d)|f(d)|}{d} \leqq$$

$$\text{(34)} \qquad \leqq c_{31} Rx(x-y) \log^{\delta_1 - 1} x \left(\sum_{d \leqq e^{\log^{c_{26}/K_1} x}} \frac{\log p}{p} + O(1) \right) \leqq$$

$$\leqq c_{32} Rx(x-y) \log^{\delta_1 - 1} x \log^{c_{26}/K_1} x,$$

taking into account (3). Putting our estimates (33) and (34) into (32) and then into (31) and at the same time expressing $N(x)$ again by $M(x)$ with the aid of (26), we get, after dividing by $x(x-y)\log^{\delta_1} x$,

$$\frac{|M(x)|}{x} \log^{1-\delta_1} x \leqq c_{33} \left(\frac{R}{\log x} + \frac{1}{\log^{\delta_1} x} + \frac{1}{\delta_2 \log^{1+\delta_1} x} \sqrt{\frac{I(\sigma)}{\sigma - 1}} + \right.$$

$$\left. + \frac{R}{\delta_1 \delta_2 \log x} + \frac{R}{\delta_1} \delta_2^{\delta_1} + R \log^{c_{26}/K_1 - 1} x \right)$$

for $2 \leqq x \leqq x_0$. Let us choose x to maximize the left hand side. The maximal value was called R in (25), and as we have anticipated, all the terms on the right containing R can be neglected compared to R on the left for large enough x and K_1. The only exception is the last but one. Here we can achieve the same result by choosing δ_2 sufficiently small once δ_1 to be determined later (depending only on δ) is fixed. Thus we have proved, rewriting $\log x$ everywhere as $\frac{1}{\sigma - 1}$,

$$\text{(35)} \qquad R \leqq c_{34} \sqrt{I(\sigma)(\sigma - 1)^{1+2\delta_1}} + O(1)$$

where the $O(1)$ is to take care of the possibility that the maximum is attained for $x \leqq c_{35}$ in which case the above argument is not applicable. Inserting here our bound for $I(\sigma)$ in (21), we have, combining the $O(1)$ terms together,

$$R \leqq c_{35} (\sigma - 1)^{\delta_1} \Phi(\sigma) \left(\frac{|F(\sigma)|}{\Phi(\sigma)} \right)^{c_{17}} + O(1)$$

for one particular $\sigma \geqq \sigma_0$. Suppose we can show that the first term on the right is decreasing in σ. Then we can replace σ by σ_0 and also the second term can be omitted. Hence

$$\frac{|M(x_0)|}{x_0}\log^{1-\delta_1} x_0 \leqq R \leqq c_{36}(\sigma_0-1)^{\delta_1}\Phi(\sigma_0)\left(\frac{|F(\sigma_0)|}{\Phi(\sigma_0)}\right)^{c_{17}},$$

$$(36)\qquad \begin{aligned}|M(x_0)| &\leqq c_{36}x_0(\sigma_0-1)\Phi(\sigma_0)\left(\frac{|F(\sigma_0)|}{\Phi(\sigma_0)}\right)^{c_{17}} \leqq \\ &\leqq c_{37}x_0\frac{\Phi(\sigma_0)}{\zeta(\sigma_0)}\left(\frac{|F(\sigma_0)|}{\Phi(\sigma_0)}\right)^{c_{17}}\end{aligned}$$

and since x_0 is arbitrary, Theorem 2 follows on observing that

$$\frac{\Phi(\sigma_0)}{\zeta(\sigma_0)} = e^{\sum_{n=1}^{\infty}\frac{\lambda(n)(|f(n)|-1)}{n^{\sigma_0}}} = e^{\sum_{p\leqq x_0}\frac{|f(p)|-1}{p}+O(1)}$$

$$\frac{|F(\sigma_0)|}{\Phi(\sigma_0)} = e^{-\sum_{n=1}^{\infty}\frac{\lambda(n)(|f(n)|-\operatorname{Re}f(n))}{n^{\sigma_0}}} = e^{-\sum_{p\leqq x_0}\frac{|f(p)|-\operatorname{Re}f(p)}{p}+O(1)},$$

by (3), (6) and (8) the usual way.

So it remains to show that

$$(\sigma-1)^{\delta_1}\Phi(\sigma)\left(\frac{|F(\sigma)|}{\Phi(\sigma)}\right)^{c_{17}}$$

is decreasing or, equivalently, that its logarithmic derivative is negative. By (11)

$$\frac{\Phi'}{\Phi}(\sigma) \leqq -\frac{\delta}{\sigma-1}+O(1),$$

by (10)

$$\frac{|F(\sigma)|'}{|F(\sigma)|} = -\sum_{n=1}^{\infty}\frac{\Lambda(n)\operatorname{Re}f(n)}{n^{\sigma}} \leqq \sum_{n=1}^{\infty}\frac{\Lambda(n)|f(n)|}{n^{\sigma}} \leqq \frac{2}{\sigma-1}+O(1)$$

and the logarithmic derivative in question,

$$\frac{\delta_1}{\sigma-1}+(1-c_{17})\frac{\Phi'}{\Phi}(\sigma)+c_{17}\frac{|F(\sigma)|'}{|F(\sigma)|} \leqq \frac{\delta_1}{\sigma-1}-\frac{(1-c_{17})\delta}{\sigma-1}+\frac{2c_{17}}{\sigma-1}+O(1)$$

is in fact negative for $\delta_1=\delta/2$, say, possibly by further decreasing the value of c_{17}, c_1 in Theorem 2, at least for small enough $\sigma-1$ that is all we actually need.

3. We continue our investigation of $F(s)$ started in § 1 aiming at a main ... We work with the assumption $|f(p)-1| \leqq \eta$ of Theorem 3 implying also ... condition of Theorem 2 so that all our previous results hold in this case as well. Consider

$$\frac{F(s)}{\zeta(s)} = e^{\sum_{n=1}^{\infty}\frac{\lambda(n)(f(n)-1)}{n^{s}}}$$

for $1+\frac{\sigma_0-1}{2} \leqq \operatorname{Re} s = \sigma \leqq 3$, fixing σ_0 as in § 2. Terms with $n=p^k$, $k \geqq 2$ contribute, according to (4), $O(\eta)$. Summation over $k=1$, $p>x_0 \left(\frac{1}{\log x_0} = \sigma_0 - 1\right)$, owing to the factor $|f(p)-1| \leqq \eta$ gives another $O(\eta)$, by (6). The remaining sum is

$$\sum_{p \leqq x_0} \frac{f(p)-1}{p^s}$$

where s when replaced by 1 causes an error $O\left(\eta \log\left(2+\frac{|s-1|}{\sigma_0-1}\right)\right)$ (see (7)). Defining therefore

$$A = e^{\sum\limits_{p \leqq x_0} \frac{f(p)-1}{p}},$$

we have shown

$$\frac{F(s)}{\zeta(s)} = Ae^{O\left(\eta+\eta\log\left(2+\frac{|s-1|}{\sigma_0-1}\right)\right)} = A\left\{1+O\left(\eta\log\left[2+\frac{|s-1|}{\sigma_0-1}\right]\right)\right\} \tag{37}$$

provided

$$\eta \log\left(2+\frac{|s-1|}{\sigma_0-1}\right) \leqq 3$$

or in an equivalent form, using

$$\zeta(s) = O\left(\frac{1}{|s-1|}\right) \qquad (|t| \leqq 2),$$

$$F(s) = A\zeta(s) + O\left(\eta |A| \frac{\log\left(2+\frac{|s-1|}{\sigma_0-1}\right)}{|s-1|}\right).$$

We have derived this uniformly in the strip $1+\frac{\sigma_0-1}{2} \leqq \sigma \leqq 3$ with a view to apply it to F' by the aid of Cauchy's estimate for the derivative. For fixed s the circle $|z-s| \leqq \frac{\sigma-1}{2}$ lies in the above strip and our remainder gives a uniform bound in the circle if we replace $|s-1|$ by its double and half, respectively, giving rise only to another constant in $O(\ )$. Thus

$$|F'(s) - A\zeta'(s)| \leqq \frac{1}{\frac{\sigma-1}{2}} O\left(\eta |A| \frac{\log\left(2+\frac{|s-1|}{\sigma_0-1}\right)}{|s-1|}\right) \tag{38}$$

under the conditions

$$\sigma_0 \leqq \operatorname{Re} s = \sigma \leqq 2, \quad \eta \log\left(2+\frac{|s-1|}{\sigma_0-1}\right) \leqq 2, \quad |t| \leqq 1.$$

§ 4. PROOF of Theorem 3. This goes essentially on the same lines as that of Theorem 2 in § 2 with the new definitions

$$M(x) = \sum_{n \leqq x} (f(n) - A),$$

$$N(x) = \sum_{n \leqq x} (f(n) - A) \log n,$$

$$I(\sigma) = \int_{(\sigma)} \frac{|F'(s) - A\zeta'(s)|^2}{|s|^2} |ds|.$$

Repeating first the elementary part (from (22) onwards), most of the relations connecting these quantities remain unchanged (e.g. (22), (24), etc.). We enumerate now the differences.

Writing down (28) for both $f(n)$ and $f(n) \equiv 1$, multiplying the second by A and substracting, we get for our new functions M, N

$$N(u) = \sum_{d \leqq u} \Lambda(d) f(d) \sum_{k \leqq \frac{u}{d}} f(k) - \sum_{d \leqq u} \Lambda(d) \sum_{k \leqq \frac{u}{d}} A =$$

$$= \sum_{d \leqq u} \Lambda(d) f(d) M\left(\frac{u}{d}\right) + A \sum_{d \leqq u} \Lambda(d)(f(d) - 1) \sum_{k \leqq \frac{u}{d}} 1$$

and this will make an additional term (using (4))

$$|A| \sum_{d \geqq u} \Lambda(d) |f(d) - 1| \frac{u}{d} \leqq |A| u \left(\eta \sum_{p \leqq u} \frac{\log p}{p} + O(\eta)\right) \leqq c_{38} \eta |A| u \log u$$

on the right of (29). Also, in estimating the deviation in (30), we have to take into account the contribution of the main term,

$$|A| \sum_{y < n \leqq x} \log n \leqq |A| (x - y + 1) \log x \leqq c_{39} |A| \frac{x}{\log^{c_{26} - 1} x}.$$

The proof then goes unaltered up to (35) where, to be precise, the $O(1)$ term is to be complemented by an $O(|A|)$. The above two changes yield then (35) in the form

$$(39) \quad \begin{aligned} R &\leqq c_{34} \sqrt{I(\sigma)(\sigma - 1)^{1 + 2\delta_1}} + O(1) + c_{38} \eta |A| \log^{1 - \delta_1} x + \frac{c_{39} |A|}{\log^{c_{26} + \delta_1 - 1} x} + \\ &+ O(|A|) \leqq c_{40} \left(\sqrt{I(\sigma)(\sigma - 1)^{1 + 2\delta_1}} + \eta |A| \log^{1 - \delta_1} x_0 + |A| + 1\right) \end{aligned}$$

for a $\sigma \geqq \sigma_0$.

Now we turn to the analytic part. We recall

$$I(\sigma) = \int_{(\sigma)} \frac{|F'(s) - A\zeta'(s)|^2}{|s|^2} |ds|.$$

If $\eta \log \left(2 + \frac{2}{\sigma_0 - 1}\right) \geqq 2$, we define τ by

$$\eta \log \left(2 + \frac{2\tau}{\sigma_0 - 1}\right) = 2;$$

in the alternative case set $\tau=1$. For $\sigma > 1+\tau$ the asymptotic relation of § 3 is about to break down and we handle the two terms of the integrand separately, using $|a-b|^2 \leqq 2(|a|^2+|b|^2)$. For the square integral of $\frac{F'(s)}{s}$ we found in (21), omitting a factor $\leqq 1$ then essential,

$$\int_{(\sigma)} \left| \frac{F'(s)}{s} \right|^2 |ds| \leqq c_{41} \frac{[\Phi(\sigma)]^2}{\sigma-1} + O(1) \leqq c_{42} \frac{[\Phi(\sigma)]^2}{\sigma-1}. \tag{40}$$

For the second term

$$\int_{(\sigma)} \left| \frac{A\zeta'(s)}{s} \right|^2 |ds| \leqq \frac{c_{43}|A|^2}{(\sigma-1)^3}. \tag{41}$$

For, by the second order pole of $\zeta'(s)$ at $s=1$ and the well-known estimate for large values of t,

$$\int_{(\sigma)} \left| \frac{\zeta'(s)}{s} \right|^2 |ds| \leqq c_{44} \left(\int_{(\sigma)} \frac{dt}{|s-1|^4} + \int_{|t| \geqq 2} \frac{\log^4 |t|}{|t|^2} dt \right) = \frac{c_{45}}{(\sigma-1)^3} + c_{46}.$$

Let now $\sigma \leqq 1+\tau$. Then by the definition of τ the asymptotic formula (38) holds up to $|s-1| \leqq 2\tau$, certainly for $|t| \leqq \tau$, giving

$$\int_{\substack{(\sigma) \\ |t| \leqq \tau}} \frac{|F'(s)-A\zeta'(s)|^2}{|s|^2} |ds| \leqq c_{47} \frac{\eta^2 |A|^2}{(\sigma-1)^2} \int_{(\sigma)} \frac{\log^2 \left(2+\frac{|s-1|}{\sigma_0-1}\right)}{|s-1|^2} dt \leqq$$

$$\leqq c_{48} \frac{\eta^2 |A|^2}{(\sigma-1)^3} \log^2 \left(2+\frac{2(\sigma-1)}{\sigma_0-1}\right). \tag{42}$$

In fact, the contribution in the last integral of $|s-1| \leqq 2(\sigma-1)$ is less than

$$\frac{(\sigma-1)\log^2\left(2+\frac{2(\sigma-1)}{\sigma_0-1}\right)}{(\sigma-1)^2}$$

and the contribution of $|s-1| \geqq 2(\sigma-1)$ is less than

$$2\int_{\sigma-1}^{\infty} \frac{\log^2\left(2+\frac{2t}{\sigma_0-1}\right)}{t^2} dt = \frac{2}{\sigma_0-1} \int_{\frac{\sigma-1}{\sigma_0-1}}^{\infty} \frac{\log^2(2+2t)}{t^2} dt \leqq \frac{c_{49}}{\sigma_0-1} \frac{\log^2\left(2+\frac{2(\sigma-1)}{\sigma_0-1}\right)}{\frac{\sigma-1}{\sigma_0-1}}$$

the same as above, proving the last step in (42). For $|t| \geqq \tau$ we again separate terms. The square integral of $\frac{F'(s)}{s}$ can be estimated the same way as in § 3 (see (18)—(21))

except that in place of (19) we only need $\tau \leqq |t| \leqq T$ and here we use (16) instead of (17). We get

$$\int_{\substack{(\sigma)\\ |t|\geqq\tau}} \left|\frac{F'(s)}{s}\right|^2 |ds| \leqq c_{50}\left[\Phi(\sigma)e^{-c_{15}\log\left(1+\frac{\tau}{\sigma-1}\right)}\right]^2 \frac{1}{\sigma-1}+O(1).$$

For the ζ-function as above

$$\int_{\substack{(\sigma)\\ |t|\geqq\tau}} \left|\frac{A\zeta'(s)}{s}\right|^2 |ds| \leqq c_{51}\frac{|A|^2}{\tau^3}.$$

(37) applied for $s=1+\tau$ shows, taking into account the definition of τ that

$$\frac{|A|^2}{\tau^3}=\frac{1}{\tau^3}\left|\frac{F(1+\tau)}{\zeta(1+\tau)}\right|^2 e^{O(1)} \leqq \frac{c_{52}}{\tau^3}[\Phi(1+\tau)\tau]^2 = c_{52}\frac{[\Phi(1+\tau)]^2}{\tau}$$

and we see that it is inferior to the previous bound. In fact, we shall show later that the latter one even when multiplied by a positive power of $\sigma-1$ is decreasing. Since now $\sigma \leqq 1+\tau$, its minial value, attained at $\sigma=1+\tau$ is $c_{50}[\Phi(1+\tau)e^{-c_{15}\log 2}]^2\frac{1}{\tau}$ and our statement follows.

Combining these estimates with the one for $|t|\leqq\tau$ in (42),

$$I(\sigma)=\int_{(\sigma)} \frac{|F'(s)-A\zeta'(s)|^2}{|s|^2}|ds| \leqq$$

(43)

$$\leqq c_{53}\left(\frac{\eta^2|A|^2}{(\sigma-1)^3}\log^2\left(2+\frac{2(\sigma-1)}{\sigma_0-1}\right)+\frac{[\Phi(\sigma)]^2}{\sigma-1}e^{-2c_{15}\log\left(1+\frac{\tau}{\sigma-1}\right)}+1\right)$$

for $\sigma \leqq 1+\tau$. But this, possibly with another c_{53} also holds for $1+\tau<\sigma\leqq 2$. In fact, as $\tau<1$ definitely, by its definition

$$\eta^2\log^2\left(2+\frac{2(\sigma-1)}{\sigma_0-1}\right) \geqq \eta^2\log^2\left(2+\frac{2\tau}{\sigma_0-1}\right) \geqq 4$$

and we see that (41) can be included in the first term on the right, while our first bound, (40) for the case $\sigma>1+\tau$ is smaller than the second term on the right, since $\log\left(1+\frac{\tau}{\sigma-1}\right) \leqq \log 2$. Putting (43) into (39), taking square roots term by term,

$$R \leqq c_{54}\left(\frac{\eta|A|}{(\sigma-1)^{1-\delta_1}}\log\left(2+\frac{\sigma-1}{\sigma_0-1}\right)+(\sigma-1)^{\delta_1}\Phi(\sigma)e^{-c_{15}\log\left(1+\frac{\tau}{\sigma-1}\right)}+\right.$$
$$\left.+(\sigma-1)^{\frac{1}{2}+\delta_1}+\eta|A|\log^{1-\delta_1}x+|A|+1\right)$$

for a particular $\sigma \geqq \sigma_0$.

The first term on the right can be seen directly to be decreasing for $\sigma \geqq \sigma_0$. The same is true for the second term. In fact, the logarithmic derivative, by (11),

$$\frac{\delta_1}{\sigma-1}+\frac{\Phi'}{\Phi}(\sigma)+\frac{c_{15}\tau}{(\sigma-1+\tau)(\sigma-1)} \leqq \frac{\delta_1}{\sigma-1}-\frac{\delta}{\sigma-1}+\frac{c_{15}}{\sigma-1}+O(1)$$

is negative for small enough δ_1, c_{15} and $\sigma-1$. We can therefore replace σ by σ_0 in the first two terms and noting $\tau \leqq 1$, $\sigma \leqq 2$ and also

$$|A| = \frac{|F(\sigma_0)|}{\zeta(\sigma_0)} e^{O(1)} \leqq c_{55} \frac{\Phi(\sigma_0)}{\log x_0},$$

by (37), the four other terms are superfluous. Recalling the definition of R we are left with

$$\frac{|M(x_0)|}{x_0} \log^{1-\sigma} x_0 \leqq R \leqq c_{56} \left(\eta |A| \log^{1-\delta_1} x_0 + \frac{\Phi(\sigma_0) e^{-c_{15} \log\left(1+\frac{\tau}{\sigma_0-1}\right)}}{\log^{\delta_1} x_0} \right).$$

For
$$\eta \log\left(2+\frac{2}{\sigma_0-1}\right) \geqq 2$$
τ is defined by

$$\eta \log\left(2+\frac{2\tau}{\sigma_0-1}\right) = 2, \quad \tau \geqq (\sigma_0-1)e^{1/\eta}$$

and we get

$$|M(x_0)| = \left| \sum_{n \leqq x_0} (f(n)-A) \right| \leqq c_{57} x_0 \left(\eta |A| + \frac{\Phi(\sigma_0)}{\log x_0} e^{-c_{58}/\eta} \right).$$

Otherwise $\tau = 1$ and there is an additional term $x_0 \dfrac{\Phi(\sigma_0)}{\log^{1+c_{59}} x_0}$, giving the three remainder terms in Theorem 3 using again

$$\frac{\Phi(\sigma_0)}{\log x_0} \leqq c_{60} \frac{\Phi(\sigma_0)}{\zeta(\sigma_0)} = c_{60} e^{\sum_{p \leqq x_0} \frac{|f(p)|-1}{p} + O(1)}.$$

§ 5. Proof of Theorem 1. As in the introduction we define the completely multiplicative $f(n) = z^{g(n)}$, $z = re^{i\vartheta}$, $|\vartheta| \leqq \pi$, i.e.

$$f(p) = \begin{cases} z & \text{if } p \in \mathfrak{P} \\ 1 & \text{if } p \notin \mathfrak{P}. \end{cases}$$

The quantities occuring in Theorem 2 and 3 take the form, recalling $E(x) = \sum_{\substack{p \leqq x \\ p \in \mathfrak{P}}} \frac{1}{p}$,

$$A = e^{\sum_{p \leqq x} \frac{f(p)-1}{p}} = e^{(z-1)E(x)},$$

$$e^{\sum_{p \leqq x} \frac{|f(p)|-1}{p}} = e^{(r-1)E(x)},$$

$$e^{\sum_{p \leqq x} \frac{|f(p)|-\operatorname{Re} f(p)}{p}} = e^{r(1-\cos\vartheta)E(x)}.$$

The conditions of Theorem 2 and 3 are satisfied if $\delta \leqq r \leqq 2-\delta$ and in the latter $\eta = |z-1| < 1$. They give

$$\frac{1}{x}\sum_{n \leqq x} z^{g(n)} = O(e^{(r-1)E(x)-c_1 r(1-\cos\vartheta)E(x)})$$

and/or

$$\frac{1}{x}\sum_{n \leqq x} z^{g(n)} = e^{(z-1)E(x)} + O\left(|z-1|\, e^{(\mathrm{Re}\, z-1)E(x)} + e^{(r-1)E(x)}\left[e^{-\frac{c_2}{|z-1|}} + \frac{1}{\log^{c_3} x}\right]\right).$$

We have

$$\frac{N(m,x)}{x} = \frac{1}{2\pi i}\int_{|z|=r} \frac{\frac{1}{x}\sum_{n \leqq x} z^{g(n)}}{z^{m+1}}\, dz = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{1}{x}\sum_{n \leqq x} z^{g(n)}\right) r^{-m} e^{-im\vartheta}\, d\vartheta.$$

Our estimations give for the integrand the maximal value $e^{(r-1)E(x)} r^{-m}$ and the optimal choice of r will be $r = \frac{m}{E(x)}$, hence the condition $\delta \leqq \frac{m}{E(x)} \leqq 2-\delta$ in Theorem 1.

For $|\vartheta| \leqq \vartheta_0 \leqq \delta/2$ we apply our formula with main term, for $|\vartheta| \geqq \vartheta_0$ just the upper bound. Integrating first the main term,

$$\frac{r^{-m}}{2\pi}\int_{-\vartheta_0}^{\vartheta_0} e^{(z-1)E(x)} e^{-im\vartheta}\, d\vartheta = \frac{r^{-m}}{2\pi}\int_{-\pi}^{\pi} e^{(z-1)E(x)} e^{-im\vartheta}\, d\vartheta +$$

$$+ O\left(r^{-m}\int_{|\vartheta| \geqq \vartheta_0} e^{(r\cos\vartheta-1)E(x)}\, d\vartheta\right).$$

The first integral is the mth Taylor coefficient of the function $e^{(z-1)E(x)}$, i.e. $\frac{E^m(x)}{m!}e^{-E(x)}$. The error, factoring out $r^{-m}e^{(r-1)E(x)} = \left(\frac{E(x)}{m}\right)^m e^{m-E(x)}$ from this and all the subsequent error estimations, is

$$\int_{|\vartheta| \geqq \vartheta_0} e^{r(\cos\vartheta-1)E(x)}\, d\vartheta \leqq 2\int_{\vartheta_0}^{\infty} e^{-c_{61}\vartheta^2 E(x)} =$$

$$= \frac{2}{\sqrt{E(x)}}\int_{\vartheta_0\sqrt{E(x)}}^{\infty} e^{-c_{61}u^2}\, du \leqq \frac{c_{62}}{\sqrt{E(x)}} e^{-c_{61}\vartheta_0^2 E(x)}.$$

Integration of the remainders give, as $|z-1| \leqq |r-1|+2\vartheta$,

$$\int_{|\vartheta|\leqq\vartheta_0} |z-1| e^{r(\cos\vartheta-1)E(x)}\,d\vartheta \leqq \int_{-\infty}^{\infty} (|r-1|+2|\vartheta|) e^{-c_{61}\vartheta^2 E(x)}\,d\vartheta =$$

$$= \left(c_{63}\frac{|r-1|}{\sqrt{E(x)}}+\frac{c_{64}}{E(x)}\right),$$

$$\int_{|\vartheta|\leqq\vartheta_0} e^{-\frac{c_2}{|z-1|}}\,d\vartheta \leqq \begin{cases} 2\vartheta_0 e^{-\frac{1}{2|r-1|}} & \text{if } |r-1| \geqq 2\vartheta_0 \\ 2\vartheta_0 e^{-\frac{1}{4\vartheta_0}} & \text{if } |r-1| \leqq 2\vartheta_0, \end{cases}$$

$$\int_{|\vartheta|\leqq\vartheta_0} \frac{1}{\log^{c_3} x}\,d\vartheta \leqq \frac{2}{\log^{c_3} x} \leqq \frac{c_{65}}{E(x)} \qquad (E(x) \ll \log\log x)$$

and the bound for $|\vartheta| \geqq \vartheta_0$ gives

$$\int_{|\vartheta|\geqq\vartheta_0} e^{-c_1 r(1-\cos\vartheta)E(x)}\,d\vartheta \leqq \frac{c_{66}}{\sqrt{E(x)}} e^{-c_{67}\vartheta_0^2 E(x)},$$

a similar quantity occured before. ϑ_0 is therefore to optimize the last and the last but two. A straightforward calculation suggests

$$\vartheta_0 = \sqrt{\frac{2}{|r-1|E(x)}} \quad \text{if} \quad |r-1| \geqq 2[E(x)]^{-\frac{1}{3}},$$

$$\vartheta_0 = [E(x)]^{-\frac{1}{3}} \quad \text{if} \quad |r-1| \leqq 2[E(x)]^{-\frac{1}{3}},$$

the optimal values being at most $\frac{c_{68}}{\sqrt{E(x)}} e^{-\frac{c_{69}}{|r-1|}}$ and $c_{70} e^{-c_{71}[E(x)]^{\frac{1}{3}}}$, respectively and both are well exceeded by other error terms. Hence we find

$$\frac{N(m,x)}{x} = \frac{E^m(x)}{m!} e^{-E(x)} + O\left(\left(\frac{E(x)}{m}\right)^m \frac{e^{m-E(x)}}{\sqrt{E(x)}} \left[\left|\frac{m}{E(x)}-1\right| + \frac{1}{\sqrt{E(x)}}\right]\right)$$

and Stirling's formula shows that this is only an alternative formulation of Theorem 1.

REFERENCE

[1] KUBILIUS, J.: *Probabilistic methods in the theory of numbers*, Translations of Mathematical Monographs, vol. 11, Amer. Math. Soc. 1964.

[2] LANDAU, E.: *Handbuch der Lehre von der Verteilung der Primzahlen*, Leipzig, Teubner 1909.

[3] SATHE, L. G.: On a problem of Hardy on the distribution of integers having a given number of prime factors, I., II., III., IV., *J. Indian Math. Soc.* **17** (1953), 63—82, 83—141; **18** (1954), 27—42, 43—81.

[4] SELBERG, A.: Note on a paper by L. G. Sathe, *J. Indian Math. Soc.* **18** (1954), 83—87.

[5] ERDŐS, P.: On the integers having exactly k prime factors, *Ann. Math.* **2**, 49 (1948), 53—66.

[6] HARDY, G. H. and RAMANUJAN, S.: The normal number of prime factors of a number n, *Quart. J. Math.*, **48** (1920), 76—92.

[7] HARDY, G. H.: *Ramanujan*, Chelsea, New York, N. Y.
[8] RÉNYI, A.: On the density of certain sequences of integers, *Publ. Inst. Math. Acad. Serbe Sci.* **8** (1955), 157—162.
[9] KUBILIUS, J.: On local theorems for additive number-theoretic functions, *Abhandlungen aus Zahlentheorie und Analysis zur Errinerung an Edmund Landau*, VEB Deutscher Verlag der Wissenschaften, Berlin 1968, pp. 175—191.
[10] WIRSING, E.: Das asymptotische Verhalten von Summen über multiplikative Funktionen II, *Acta Math. Acad. Sci. Hung.* **18** (1967), 411—467.
[11] DELANGE, H.: Un théorème sur les fonctions arithmetiques multiplicatives et ses applications, *Annal. Sci. Éc. Norm. Sup.*, 3 ser. **78** (1961), 1—29.
[12] HALÁSZ, G.: Über die Mittelwerte multiplikativer zahlentheoretischer Funktionen, *Acta Math. Acad. Sci. Hung.* **19** (1968), 365—403.
[13] DELANGE, H.: A theorem on multiplicative arithmetic functions, *Proc. Amer. Math. Soc.* **18**, No. 4 (1967), 743—749.
[14] LEVIN, B. V. and FAINLEIB, A. S.: Application of certain integral equations to questions of the theory of numbers, *Uspehi Mat. Nauk* **22** (1967), No. 3 (135), 119—197 (Russian). English translation: *Russian Math. Survey* **22** (1967), No. 3, 119—204.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received March 22, 1970; Revised March 15, 1971)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 235—241.

PROBABILITY DISTRIBUTIONS IN THE GEOMETRY OF CLUSTERS

by

R. V. AMBARTZUMIAN

Abstract. The term "cluster" is used in the present paper instead of more usual "finite set of points". Only the planar case is being discussed.

The probability distributions, to which we arrive, arise, when a fixed cluster $\mathfrak{M}$ is intersected by a random (oriented) straight line. But instead of defining the latter directly as a random point in appropriate phase space (in fact the "invariant measure" [1] of straight lines on plane is involved) we rather consider it as a limit of certain random circle, when its radius tends to ∞. This is the so called "invariant imbedding" technics, which has proved itself to be quite useful in far deeper problems of integral geometry, [2].

Independently of their probabilistic interpretation, some results of the paper seem to be new in the cluster geometry. For example, it is shown, that for centrally-symmetrical "rich" clusters (n, the number of points in $\mathfrak{M}$, tends to ∞, limiting quantities are considered)

$$h_{m+2} \geqq 4r\left(1-\frac{1}{2^{m+1}}\right)$$

where h_m is the mean length of the perimeter of minimal convex hull of a randomly chosen m-subset of the cluster, r is the mean distance of the points, belonging to $\mathfrak{M}$ from 0, the center of symmetry. This result holds under the condition, that r and $H=h_n$ do not increase with n.

Another result for "rich" clusters states, that

$$\frac{\varrho}{H} \leqq \frac{1}{4}$$

where ϱ is the mean distance between the pairs of points in $\mathfrak{M}$.

In fact, these inequalities are derived from their analogs for clusters with finite n. The cases when equalities hold are indicated.

*

Denote the points of the planar cluster $\mathfrak{M}$ by $\mathscr{P}_1, \mathscr{P}_2, \ldots, \mathscr{P}_n$.

Definition 1. The random circle $C(r)$ has a constant radius r and its center is distributed uniformly in the interior of the "basic circle" of radius R centered in the origin. It is assumed, that R is taken large enough, so that $\mathfrak{M}$ lies in the basic circle and the distance of $\mathfrak{M}$ from the boundary of the "basic circle" exceeds r.

We consider the probability $P_k(r)$ of finding k points from $\mathfrak{M}$ in $C(r)$.

Inside the random circle $C(r+h)$ (h is assumed to be small) we consider a concentric circle D of radius r. Obviously, the distribution of the number of points in D coincides with that for $C(r)$. We denote the annulus formed by the boundaries of $C(r+h)$ and D by K.

We introduce the probability $P\begin{pmatrix} D & K \\ k & ; & i \end{pmatrix}$ of an event, which occurs, when k points from $\mathfrak{M}$ are found in D, and i points from $\mathfrak{M}$ are found in K. The probability of a point from $\mathfrak{M}$ to be found on the boundary of D is equal to zero, so we have

$$(1) \qquad P_k(r+h) = \sum_{i=0}^{k} P\begin{pmatrix} D & & K \\ k-i & ; & i \end{pmatrix}$$

Now, the probability of finding two or more points from $\mathfrak{M}$ in K is $o(h)$ when $h \to 0$. We will refer to this fact as to Remark 1 (R1).

According to R1, the probabilities $P\begin{pmatrix} D & & K \\ k & ; & i \end{pmatrix}$ for $i > 1$ are $o(h)$ when $h \to 0$. Thus (1) is simplified to

$$(2) \qquad P_k(r+h) = P\begin{pmatrix} D & & K \\ k & ; & 0 \end{pmatrix} + P\begin{pmatrix} D & & K \\ k-1 & ; & 1 \end{pmatrix} + o(h)$$

Here, and in the sequel, the probabilities, invonving negative numbers of points from $\mathfrak{M}$ should be replaced by zeros.

On the other hand, R1 implies that

$$(3) \qquad P_k(r) = P\begin{pmatrix} D & & K \\ k & ; & 0 \end{pmatrix} + P\begin{pmatrix} D & & K \\ k & ; & 1 \end{pmatrix} + o(h)$$

Subtracting (3) from (2) we find, that

$$(4) \qquad P_k(r+h) - P_k(r) = P\begin{pmatrix} D & & K \\ k-1 & ; & 1 \end{pmatrix} - P\begin{pmatrix} D & & K \\ k & ; & 1 \end{pmatrix} + o(h)$$

Taking account of R1 again, we establish

$$(5) \qquad P\begin{pmatrix} D & & K \\ k & ; & 1 \end{pmatrix} = \sum_{i=0}^{n} P\begin{pmatrix} D & & K \\ k & ; & \mathscr{P}_i \end{pmatrix} + o(h)$$

where $P\begin{pmatrix} D & & K \\ k & ; & \mathscr{P}_i \end{pmatrix}$ is the probability of an event, which occurs when k points from $\mathfrak{M}$ are found in D and $\mathscr{P}_i \in \mathfrak{M}$ is found in K.

Definition 2. The random circle $C(r; \mathscr{P}_i)$ has a constant radius r, and the center of $C(r; \mathscr{P}_i)$ is distributed uniformly on a circle of radius r with its center in the point $\mathscr{P}_i \in \mathfrak{M}$. Thus, the point $\mathscr{P}_i \in \mathfrak{M}$ is found on the boundary of $C(r; \mathscr{P}_i)$ with probability 1.

Denote by $\pi_k(r; \mathscr{P}_i)$ the probability of finding k points from $\mathfrak{M}$ in $C(r; \mathscr{P}_i)$. It causes no difficulty to show the validity of

LEMMA 1.

$$P\binom{D}{k};\binom{K}{\mathscr{P}_i} = \frac{2rh}{R^2}\pi_k(r;\mathscr{P}_i)+o(h), \qquad h\to 0.$$

From Lemma 1 and the equations (3) and (4), it follows, that

$$\frac{dP_k(r)}{dr} = \frac{2r}{R^2}\sum_{i=1}^{k}[\pi_{k-1}(r;\mathscr{P}_i)-\pi_k(r;\mathscr{P}_i)] \tag{6}$$

The next step consists in analysing the probabilities $\pi_k(r;\mathscr{P}_i)$ by means of variation of r. In order to do this we make an additional

Assumption. No three points from $\mathfrak{M}$ lie on the same straight line. We shall refer to this Assumption as to A.

Inside the random circle $C(r+h;\mathscr{P}_i)$ we consider a circle $\tilde{D}$, which has radius r, and has a common tangent with $C(r+h;\mathscr{P}_i)$ in $\mathscr{P}_i$. The distribution of the number of the points from $\mathfrak{M}$ found in $\tilde{D}$ coincides with $\pi_k(r;\mathscr{P}_i)$. Denote by $\tilde{K}$ the part of $C(r+h;\mathscr{P}_i)$ which lies outside of $\tilde{D}$.

As a matter of a fact, A implies, that for larger values of r the probabilities $\Pi_i\binom{\tilde{K}}{j}$ of finding j points from $\mathfrak{M}$ in $\tilde{K}$ (Π_i refers to probabilities, associated with the random circle $C(r+h;\mathscr{P}_i)$) are $o(h)$ when $h\to 0$ for $i\geqq 2$. We shall refer to this fact as to R2.

According to R2 the equations

$$\begin{aligned}\pi_k(r+h;\mathscr{P}_i) &= \Pi_i\binom{\tilde{D}}{k};\binom{\tilde{K}}{0}+\Pi_i\binom{\tilde{D}}{k-1};\binom{\tilde{K}}{1}+o(h)\\ \pi_k(r;\mathscr{P}_i) &= \Pi_i\binom{\tilde{D}}{k};\binom{\tilde{K}}{0}+\Pi_i\binom{\tilde{D}}{k};\binom{\tilde{K}}{1}+o(h)\end{aligned} \tag{7}$$

are valid. The notations in (7) are analogous to ones in (2) and (3).

Hence, for larger values of r

$$\pi_k(r+h;\mathscr{P}_i)-\pi_k(r;\mathscr{P}_i) = \Pi_i\binom{\tilde{D}}{k-1};\binom{\tilde{K}}{1}-\Pi_i\binom{\tilde{D}}{k};\binom{\tilde{K}}{1}+o(h) \tag{8}$$

Again, from R2 we deduce, that for larger values of r

$$\Pi_i\binom{\tilde{D}}{k};\binom{\tilde{K}}{1} = \sum_{j\neq i}\Pi_i\binom{\tilde{D}}{k};\binom{\tilde{K}}{\mathscr{P}_j}+o(h) \tag{9}$$

The notations in (9) are analogous to ones in (5). An easy geometrical reasoning ensures the validity of

LEMMA 2. *Under assumption A for larger values of r*

$$\Pi_i\binom{\tilde{D}}{k};\binom{\tilde{K}}{\mathscr{P}_j} = \frac{h}{2\pi}\,\frac{\varrho_{ij}\nu_k(\mathscr{P}_i,\mathscr{P}_j)}{r\sqrt{4r^2-\varrho_{ij}^2}}+o(h)$$

where ϱ_{ij} is the distance between $\mathcal{P}_i \in \mathfrak{M}$ and $\mathcal{P}_j \in \mathfrak{M}$ and

$$v_k(\mathcal{P}_i, \mathcal{P}_j) = \begin{cases} 0 & \text{if neither of the two halfplanes, in which the plane is divided by the straight line, passing through } \mathcal{P}_i \in \mathfrak{M} \text{ and } \mathcal{P}_j \in \mathfrak{M} \text{ does not contain } k \text{ points from } \mathfrak{M}. \\ 1 & \text{if only one of those halfplanes contains } k \text{ points from } \mathfrak{M}. \\ 2 & \text{if both halfplanes contain } k \text{ points from } \mathfrak{M},\ k = \frac{n-2}{2} \end{cases}$$

The equations (8), (9) and Lemma 2 lead to the differential equations, valid for larger values of r

$$\text{(10)} \qquad \frac{d\pi_k(r; \mathcal{P}_i)}{dr} = \sum_j \frac{1}{2\pi} \frac{\varrho_{ij}}{r\sqrt{4r^2 - \varrho_{ij}^2}} [v_{k-1}(\mathcal{P}_i, \mathcal{P}_j) - v_k(\mathcal{P}_i, \mathcal{P}_j)]$$

The equation (10) taken together with (6) permits to find the asimptotic ($r \to \infty$) behaviour of the probabilities $P_k(r)$. A strightforward calculation shows, that

$$P_0(r) = 1 - \frac{r^2}{R^2} - \frac{r}{\pi R^2} \sum_{i<j} \varrho_{ij} v_0(\mathcal{P}_i, \mathcal{P}_j) + \frac{o(r)}{R^2}$$

(11)

$$P_k(r) = -\frac{r}{\pi R^2} \sum_{i<j} \varrho_{ij} [v_k(\mathcal{P}_i, \mathcal{P}_j) - 2v_{k-1}(\mathcal{P}_i, \mathcal{P}_j) + v_{k-2}(\mathcal{P}_i, \mathcal{P}_j)] + \frac{o(r)}{R^2}, \quad 0 < k < n,$$

$$P_n(r) = \frac{r^2}{R^2} - \frac{r}{\pi R^2} \sum_{i<j} \varrho_{ij} v_{n-2}(\mathcal{P}_i, \mathcal{P}_j) + \frac{o(r)}{R^2}.$$

For any cluster $\mathfrak{M}$, sattisfying A the following function is introduced

$$f_m = \sum_{i<j} v_m(\mathcal{P}_i, \mathcal{P}_j) \varrho_{ij}, \qquad m = 0, \ldots, n-2$$

As seen from the definition of $v_m(\mathcal{P}_i, \mathcal{P}_j)$ and (11), it possesses the following properties

a) $f_0 = f_{n-2} = H$, H is the length of the perimeter of minimal convex hull of $\mathfrak{M}$.

b) $f_k = f_{n-2-k}$.

c) $2f_{k-1} - f_k - f_{k-2} > 0$.

d) $\sum_{i=0}^{n-2} f_i = 2 \sum_{i<j} \varrho_{ij}$

c) is easilly seen to be true for several special cases, but as a general result it seems to be new.

The ratios $\frac{P_k(r)}{1-P_0(r)-P_n(r)}$, which may be interpreted as conditional probabilities to find k points from $\mathfrak{M}$ in $C(r)$, under the condition, that the border of $C(r)$ intersects $\mathfrak{M}$, do not depend on R, and from (11) we easily find the limits

$$(12) \qquad \mu_k = \lim_{r\to\infty} \frac{P_k(r)}{1-P_0(r)-P_n(r)} = \frac{2f_{k-1}-f_k-f_{k-2}}{2H}, \qquad k = 1, \dots, n-1.$$

The feeling that the distribution μ_k, as introduced in (12) should arise in the context involving random straight lines is justified by the Theorem 1, which follows. Writing the equation of a straight line on (x, y) plane in the form

$$(13) \qquad x\cos\varphi + y\sin\varphi = p, \quad p>0, \quad 0\leqq\varphi\leqq 2\pi$$

we denote by F the domain in (φ, p) stripe composed from all the (φ, p) points, for which the straight lines, given by (13) intersects $\mathfrak{M}$.

Definition 3. M is a random oriented straight line, which corresponds to the random point with uniform distribution in F. The orientation of M is choosen at random, with probability $^1/_2$, independent of (φ, p).

THEOREM 1. *The probability, that k points from $\mathfrak{M}$ are found on the right of M is equal to μ_k in* (12).

We omit the proof.

In calculating the moments of the distribution μ_k is important the symmetry

$$(14) \qquad \mu_k = \mu_{n-k}, \qquad k = 1, \dots, n-1$$

It follows from (14), that

$$\mathbf{E}\zeta = \frac{n}{2}$$

(ζ stands for the random number of points on the right of M).

For the variance of ζ one gets the expression

$$\mathbf{D}\zeta = \frac{n^2}{4} - 2\frac{\sum\limits_{i<j}\varrho_{ij}}{H}$$

This gives a rough estimate

$$\frac{\varrho}{H} < \frac{n}{4(n-1)}, \qquad \varrho = \frac{2\sum\limits_{i<j}\varrho_{ij}}{n(n-1)}$$

This bound of the ratio ϱ/H coincides with the sup ϱ/H, only in the case when n is even.

Indeed, if n is even, $\mathfrak{M}$ may be taken to consist of two subclusters with numbers of points equal to $n/2$, placed in distant circles of radius ε. Denote by $\mathfrak{M}_0$ the limiting cluster when $\varepsilon\to 0$. For $\mathfrak{M}_0$, $\zeta = \frac{n}{2}$ with probability 1 and hence $\mathbf{D}\zeta = 0$.

$$\sup\frac{\varrho}{H} = \frac{n}{4(n-1)} \quad \text{if } n \text{ is even.}$$

On the other hand, when n is odd, $\mathbf{D}\zeta$ can not be made arbitrarily small, because $\frac{n}{2}$ is not integer. One can see from (14), that in this case $\mathbf{D}\zeta > \frac{1}{4}$. Assuming, that $\mathfrak{M}$ consists of two subclusters, with $\frac{n-1}{2}$ points in one and $\frac{n+1}{2}$ points in another, placed in two distant circles of radius ε, tending to zero, we find, that

$$\inf \mathbf{D}\zeta = \frac{1}{4}$$

This means, that

$$\sup \frac{\varrho}{H} = \frac{n+1}{4n} \quad \text{if} \quad n \quad \text{is odd.}$$

The $\sup \mathbf{D}\zeta = \frac{(n-2)^2}{4}$ is approched, when *one* point of $\mathfrak{M}$ goes to ∞. This assures, that

$$\inf \frac{\varrho}{H} = \frac{1}{n}$$

c) finds another application in the problem of evaluating the mean length h_m of the perimeter of minimal convex hull of a randomly chosen m-subset of $\mathfrak{M}$. For instance, if n is even

$$h_{m+2} = \frac{1}{\binom{n}{m+2}} \left[\sum_{K=0}^{\frac{n}{2}-2} \left(\binom{K}{m} + \binom{n-2-K}{m} \right) f_K + \binom{\frac{n}{2}-1}{m} f_{\frac{n}{2}-1} \right]$$

since each pair $\mathscr{P}_i \mathscr{P}_j$ belongs to the perimeter of convex hull of exactly $\binom{K}{m} + \binom{n-2-K}{m}$ different $m+2$-subsets of $\mathfrak{M}$, where $K \leqq \frac{n}{2}-1$ is defined by the condition $v_k(\mathscr{P}_i, \mathscr{P}_j) > 0$.

c) implies that

$$f_K \geqq \frac{2}{n-2} \left(f_{\frac{n}{2}-1} - H \right) K + H, \qquad 0 \leqq K \leqq \frac{n}{2} - 1$$

As it follows from the Theorem, the cluster $\mathfrak{M}_0$ here again gives the equality. Thus we find, that

$$(15) \qquad h_{m+2} \geqq \frac{1}{\binom{n}{m+2}} \left[\frac{2A(n,m)}{n-2} \left(f_{\frac{n}{2}-1} - H \right) + B(n,m) H + \binom{\frac{n}{2}-1}{m} f_{\frac{n}{2}-1} \right]$$

The functions A and B depend on $\mathfrak{M}$ only through n, and may be written in the following compact form

$$A(n, m) = \binom{n-1}{m+2} - \binom{\frac{n}{2}+1}{m+2} - \binom{\frac{n}{2}-1}{m+2} - \left(\frac{n}{2}-2\right)\binom{\frac{n}{2}}{m+1} + \left(\frac{n}{2}-2\right)\binom{\frac{n}{2}-1}{m+1},$$

$$B(n, m) = \binom{\frac{n}{2}-1}{m+1} + \binom{n-1}{m+1} - \binom{\frac{n}{2}}{m+1}.$$

It should be noted, that $f_{\frac{n}{2}-1}$ (n is even) permits nice interpretation in the case of centrally-symmetrical clusters (0-the center of symmetry) as

$$f_{\frac{n}{2}-1} = 2\sum r_i$$

r_i being the distance from $\mathcal{P}_i$ to 0.

For such clusters denote by r the mean distance of the points belonging to $\mathfrak{M}$ from 0, $r = \frac{1}{n}\Sigma r_i$, and let $n \to \infty$ in such a way, that r and H tend to finite limiting values (this means, that a sequence of centrally-symmetrical clusters is considered). In the limit (15) simplifies to

$$h_{m+2} \geqq 4r\left(1 - \frac{1}{2^{m+1}}\right) \tag{16}$$

Obviously, on the right hand side stands the limiting ($n \to \infty$) value of h_{m+2} for $\mathfrak{M}_0$ cluster

Acknowledgement. The paper was discussed with late professor A. RÉNYI.

LITERATURE

[1] BLASCHKE, W.: *Vorlesungen über Integralgeometrie*. Chelsea, New York, 1949.

[2] Амбарцумян Р. В.: Метод инвариантного вложения в теории случайных прямых. Изв. АН Арм. ССР, Математика, т. V,, № 3, 1970 г.

Institute of Mathematics and Mechanics Erevan, Armenia, USSR.

(Received December 20, 1968.)

INDEX

Printed in Hungary

A kiadásért felel az Akadémiai Kiadó igazgatója – Műszaki szerkesztő: Várhelyi Tamás
A kézirat a nyomdába érkezett: 1971. I. 8. – Terjedelem: 21,5 (A/5) ív, 9 ábra

71-61 – 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 Reaktion: 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 Mathematischen Instituts (Budapest V., Reáltanoda u. 13—15) getroffen werden.

Die zur Veröffentlichung bestimmten Manuskripte sind in zwei Exemplaren an die Redaktion zu schicken.

Studia Scientiarum Mathematicarum Hungarica est une revue biannuelle de l'Académie Hongroise des Sciences publiant des essais originaux, en français, anglais, allemand ou russe, du domaine des mathématiques.

Rédaction: 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 Mathématique (Budapest V., Reáltanoda u. 13—15).

On est prié d'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 немеском, английском, французском и русском языках. Отдельвные вышуски составляют ежегодно один том.

Адрес редакции: Budapest V., Reáltanoda u. 13—15, Венгрия.
Технический редактор: Gy. Petruska.

Подпусная цена на год (за один том): $ 16.00. Подписка на журнал принимается Внешнеторговым предприятием „Культура" (Budapest 62, P. O. B. 149) или его представительствами за границей.

По поводу отношения обмена просим обрашаться к Библиотеке Института Математики (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 Librairie
48, rue Gay Lussac
Paris 5

GERMAN DEMOCRATIC REPUBLIC

Deutscher Buchexport und Import
Leninstraße 16.
Leipzig C. I.
Zeitungsvertriebsamt
Clara Zetkin Straße 62.
Berlin N. W.

GERMAN FEDERAL REPUBLIC

Kunst und Wissen
Erich 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-chome
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 Grundt 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

29. II. 1972. **Index: 26 748**

Studia Scientiarum Mathematicarum Hungarica

AUXILIO
CONSILII INSTITUTI MATHEMATICI
ACADEMIAE SCIENTIARUM HUNGARICAE

TOMUS VI.
FASC. 3–4.
1971

AKADÉMIAI KIADÓ, BUDAPEST

Studia Scientiarum Mathematicarum Hungarica

A Magyar Tudományos Akadémia matematikai folyóirata

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

Technikai szerkesztő: Petruska Gy.

Kiadja az Akadémiai Kiadó, 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 Hírlap Külkereskedelmi Vállalatnál (Budapest II., Fő u. 32).

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

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: 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 and 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 (Budapest V., Reáltanoda u. 13—15.)

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

STUDIA SCIENTIARUM MATHEMATICARUM HUNGARICA

TOMUS VI

INDEX

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 243—249.

CONGRUENCES DE THUE ET t-LANGAGES

par

M. NIVAT

Résumé: On introduit la classe des t-langages, sous-classe de celle des langages de CHOMSKY. Tout t-langage est susceptible d'une définition algébrique puisqu'il s'agit de l'image homomorphe inverse d'un élément d'un groupe libre et d'une définition combinatoire au moyen d'une congruence de Thue que l'on construit explicitement à partir de la définition algébrique. Il en résulte la décidabilité du problème de l'équivalence de deux t-langages. La première partie de cet article contient un lemme combinatoire qui est à la base de nos résultats.

I — Un lemme combinatoire

Soit $Z=\{y_i^\varepsilon | i=1, \dots, n, \ \varepsilon=\pm 1\}$ un alphabet fini. Il est classique de considérer l'application ϱ de Z^* (monoïde libre engendré par Z) dans Z^* définie par

— $\varrho(e)=e$ (e désigne le mot vide dans ce qui suit)

— Pour tout $f \in Z^*: \varrho(fy_i^\varepsilon)\begin{cases} = \varrho(f)y_i^\varepsilon & \text{si} \quad \varrho(f) \notin Z^* y_i^{-\varepsilon} \\ = f_1 & \text{si} \quad \varrho(f)=f_1 y_i^{-\varepsilon}. \end{cases}$

La relation d'équivalence ϱ définie par $f\varrho g \Leftrightarrow \varrho(f)=\varrho(g)$ est une congruence sur Z^*. Le quotient de Z^* par ϱ n'est autre que le groupe libre à n générateurs. Nous utiliserons dans la suite toutes les propriétés de ϱ (voir par exemple MAGNUS, KARASS et SOLITAR [1]).

Nous noterons $|f|$ la longueur du mot f et si $f=y_{i_1}^{\varepsilon_1} y_{i_2}^{\varepsilon_2} \dots y_{i_p}^{\varepsilon_p}$, nous désignerons par f^{-1} le mot $f^{-1}=y_{i_p}^{-\varepsilon_p} y_{i_{p-1}}^{-\varepsilon_{p-1}} \dots y_{i_2}^{-\varepsilon_2} y_{i_1}^{-\varepsilon_1}$

Le lemme essentiel dans ce travail est le suivant:

LEMME 1. *Soient $f_1, \dots, f_p$ des mots de Z^* satisfaisant $\varrho(f_1, \dots, f_p)=e$. Il existe un entier l, $1 \leq l < p$*

$$|\varrho(f_l f_{l+1})| \leq \max_{i=1, \dots, p} |\varrho(f_i)|$$

DÉMONSTRATION. Si pour tout $i=1, \dots, p$ $\varrho(f_i)=e$ le lemme est trivialement vérifié. Sinon il existe certainement un entier j, $1 \leq j < p$, tel que

$$|\varrho(f_1 \dots f_j)| > 0.$$

Comme $|\varrho(f_1, \dots, f_p)|=0$ il existe un entier l, $1 \leq l < p$ tel que

$$|\varrho(f_1 \dots f_{l-1})| \leq |\varrho(f_1 \dots f_l)|$$

$$|\varrho(f_1 \dots f_l)| > |\varrho(f_1 \dots f_{l+1})|.$$

Il se présente alors deux cas:

1) $$\varrho(f_1 \dots f_{l-1}) = g_1 g_2$$
$$\varrho(f_1 \dots f_l) = g_1 g_3 \quad \text{où} \quad |g_3| \geqq |g_2|$$

ce qui suppose $\varrho(f_l) = g_2^{-1} g_3$ $\varrho(f_1 \dots f_{l+1}) = g_1 g_4 g_6$ avec $g_3 = g_4 g_5$, $|g_5| > |g_6|$ ce qui suppose $\varrho(f_{l+1}) = g_5^{-1} g_6$.

Finalement $\varrho(f_l f_{l+1}) = \varrho(g_2^{-1} g_4 g_5 g_5^{-1} g_6)$ $\varrho(f_l f_{l+1}) = \varrho(g_2^{-1} g_4 g_6)$ et comme $|g_5| > |g_6|$

$$|\varrho(f_l f_{l+1})| = |\varrho(g_2^{-1} g_4 g_6)| \leqq |g_2| + |g_4| + |g_6| <$$
$$< |g_2| + |g_4| + |g_5| = |\varrho(f_l)|.$$

2) $\varrho(f_1 \dots f_{l-1}) = g_1 g_2 g_3$

$$\varrho(f_1 \dots f_l) = g_1 g_2 g_4 \quad \text{avec} \quad |g_4| \geqq |g_3|$$

ce qui suppose $\varrho(f_l) = g_3^{-1} g_4$

$$(f_1 \dots f_{l+1}) = g_1 g_5 \quad \text{avec} \quad |g_5| < |g_2| + |g_4|$$

ce qui suppose $\varrho(f_{l+1}) = g_4^{-1} g_2^{-1} g_5$.

Nous avons alors

$$\varrho(f_l f_{l+1}) = \varrho(g_3^{-1} g_2^{-1} g_5)$$

soit

$$|\varrho(f_l f_{l+1})| \leqq |g_3| + |g_2| + |g_5| \leqq |g_4| + |g_2| + |g_5| = |\varrho(f_{l+1})|.$$

Avant d'établir les lemmes suivants nous avons besoin de quelques définitions que nous empruntons à [1] page 288.

Définition 1. Les schémas binaires de parenthèses sont des mots sur l'alphabet $\Delta = \{(\ ,\ ),\ *\}$ que l'on définit récursivement comme suit:

— L'ensemble S_1 des schémas binaires de parenthèses d'ordre 1 est $S_1 = \{(*)\}$.

— L'ensemble S_n des schémas binaires de parenthèses d'ordre n est

$$S_n = \{(\beta^k \cdot \beta^l) \mid \beta^k \in S_k,\ \beta^l \in S_l,\ k + l = n\}$$

Nous appelerons sous schéma du schéma de parenthèses β tout triple $[\beta_1, \beta', \beta_2]$ où β' est un schéma de parenthèse, β_1 et β_2 sont des mots de Δ^* et $\beta = \beta_1 \beta' \beta_2$.

Les propriétés suivantes sont immédiates.

1) Si $[\beta_1, \beta', \beta_2]$ est un sous-schéma de $\beta \in S_n$ avec $\beta' \in S_k$ et si $\beta'' \in S_l$ $\beta_1 \beta'' \beta_2$ est un schéma de parenthèses d'ordre $n - k + l$.

2) Si $[\beta_1, \beta', \beta_2]$ et $[\gamma_1, \gamma', \gamma_2]$ sont deux sous schémas de β, l'un des deux cas suivants se produit:

a) ce sont des sous schémas disjoints, c'est-à-dire:

$$\text{soit } \beta_1 = \gamma_1 \gamma' \gamma_1', \quad \text{soit } \gamma_1 = \beta_1 \beta' \beta_1'.$$

b) l'un des sous schémas est intérieur à l'autre, c'est-à-dire:

soit $\beta_1 = \gamma_1 \gamma_1'$, $\beta_2 = \gamma_2' \gamma_2$ et $[\gamma_1', \beta', \gamma_2']$ est un sous schéma de γ'

soit $\gamma_1 = \beta_1 \beta_1'$, $\gamma_2 = \beta_2' \beta_2$ et $[\beta_1', \gamma', \beta_2']$ est un sous schéma de β'

Définition 2. Si $f_1, \ldots, f_n$ est une suite d'éléments d'un monoïde multiplicatif le produit $f_1 \ldots f_n$ parenthésé par le schéma binaire de parenthèses $\beta \in S_n$ est défini récursivement comme suit:

pour $n=1, \quad \beta=(*)\colon \quad \beta(f_1)=(f_1)$

pour $n>1, \quad \beta=(\beta^k\beta^l)\colon$

$$\beta(f_1 \cdots f_n) = \big(\beta^k(f_1 \cdots f_k)\beta^l(f_{k+1} \cdots f_n)\big)$$

Nous appelerons segment de $f_1 \ldots f_n$ parenthésé par β tout produit $f_i \ldots f_{i+h'}$ facteur de $f_1 \ldots f_n$ tel qu'il existe un sous-schéma $[\beta_1, \beta', \beta_2]$ de β avec $\beta(f_1 \ldots f_n)=g_1 \beta'(f_i \ldots f_{i+h})g_2$.

Il est clair que la correspondance entre segment et sous schéma est biunivoque.

LEMME 2. *Soient $f_1 \ldots f_n$ des mots de Z^* qui vérifient $\varrho(f_1 \ldots f_n)=e$. Il existe un schéma binaire de parenthèses β tel que tout segment f' de $f_1 \ldots f_n$ parenthésé par β satisfait*

$$|\varrho(f')| \leqq \max_{i=1, \ldots, n} |\varrho(f_i)|$$

DÉMONSTRATION.

Nous faisons une récurrence sur l'entier n.

Pour $n=1$ c'est immédiat. Supposons la propriété vérifiée pour tout $n<p$ et considérons $f_1, \ldots, f_p$ tels que $\varrho(f_1 \ldots f_p)=e$.

D'après le lemme 1 il existe un entier l tel que

$$|\varrho(f_l f_{l+1})| \leqq \max_{i=1, \ldots, p} |\varrho(f_i)| = \lambda.$$

Posons $f_l f_{l+1}=g$. D'après l'hypothèse de récurrence, il existe un schéma de parenthèse β satisfaisant les conditions du lemme pour les $p-1$ mots $f_1 \ldots f_{l+1} g f_{l+1} \ldots \ldots f_p$.

Or $\max\,(|\varrho(f_i)|, |\varrho(g)| : i = 1, \ldots, l-1, l+1, \ldots, p) \leqq \lambda$.

Soit $[\beta_1, \beta', \beta_2]$ le sous-schéma de β correspondant au segment g $(\beta'=(*))$ et considérons le schéma de parenthèses $\gamma=\beta_1\big((*)(*)\big)\beta_2$. Ce schéma satisfait les conditions du lemme pour $f_1 \ldots f_p$. En effet, soit $[\gamma_1, \gamma', \gamma_2]$ un sous-schéma de γ, auquel correspond le segment f'. Distinguons les trois cas:

— les sous-schémas $[\gamma_1, \gamma', \gamma_2]$ et $[\beta_1, \gamma'', \beta_2]$ de γ sont disjoints. Alors f' est un segment de $f_1 \ldots f_{l-1} g f_{l+1} \ldots f_p$ parenthésé par β d'où $|\varrho(f')| \leqq \lambda$,

— le sous-schéma $[\gamma_1, \gamma', \gamma_2]$ est intérieur à $[\beta_1, \gamma'', \beta_2]$. En ce cas, f' est soit f_l soit f_{l+1} soit $f_l f_{l+1}$ et $|\varrho(f')| \leqq \lambda$,

— le sous-schéma $[\beta_1, \gamma'', \beta_2]$ est intérieur à $[\gamma_1, \gamma', \gamma_2]$: alors f' est de la forme $f_i \ldots f_l f_{l+1} \ldots f_{l+4}$ et comme $f_i \ldots f_{l-1} g f_{l+2} \ldots f_{l+4}$ est un segment de $f_1 \ldots f_{l-1} g f_{l+1} \ldots f$ parenthésé par β on a bien $|\varrho(f')| \leqq \lambda$.

Remarque 2. Nous pouvons énoncer un lemme un peu plus fort que le lemma 2, dont la démonstration est presque identique:

LEMME 2': *Soient $f_1, \ldots, f_n$ tels que $\varrho(f_1, \ldots, f_n)=e$, $f'=f_i \ldots f_{i+h}$ un facteur de $f=f_1 \ldots f_n$ et β' un schéma de parenthèses tel que pour tout segment f'' de f' paran-*

thésé par β' on ait $|\varrho(f')| \leqq \lambda$. Il existe alors un schéma de parenthèses $\beta = \beta_1 \beta' \beta_2$ tel que:

a) *f' soit le segment de f parenthésé par β, associé au sous-schéma $[\beta_1, \beta' \beta_2]$ de β.*

b) *pour tout segment f''' de f parenthésé par β on ait $|\varrho(f''')| \leqq \lambda$.*

II — Algorithme

Soit φ un homomorphisme de X^* dans Z^* où Z^* est comme au paragraphe précédent, muni de l'application ϱ. Posons $\lambda = \max_{x \in X} |\varrho(\varphi(x))|$. Construisons récursivement les ensembles A_k, B_k, R_k, $k = 1, 2, \dots, n, \dots$. Les A_k et B_k sont des ensembles de mots de X^*, les R_k des ensembles de relation entre mots de X^* que nous noterons $f \equiv g$.

Si R est un tel ensemble de relations, $\bar{R}$ désigne l'ensemble de leurs membres gauches.

$$B_1 = X$$

$$R_1 = \{x \equiv e \mid \varrho(\varphi(x)) = e\}$$

$$A_1 = X \setminus \bar{R}_1$$

$$B_{n+1} = \left\{ fg \mid f \in A_k, g \in A_l, k+l = n+1 \; |\varrho(\varphi(fg))| \leqq \lambda, fg \notin X^* \left(\bigcup_{i=1}^{i=n'} \bar{R}_i \right) X^* \right\}.$$

$$R_{n+1} = \{f \equiv e \mid f \in B_{n+1}, \varrho(\varphi(f)) = e\}$$

$$\cup \left\{ f \equiv g \mid f \in B_{n+1}, g \in \bigcup_{i=1}^{i=n} A_i, \varrho(\varphi(f)) = \varrho(\varphi(g)) \right\}$$

$$A_{n+1} = B_{n+1} \setminus \bar{R}_{n+1}$$

Remarquons que cet algorithme se termine.

En effet, il ne peut figurer de mots f et g tels que $\varrho(\varphi(f)) = \varrho(\varphi(g))$, $|\varrho(\varphi(f))| \leqq \lambda$ dans deux ensembles A_k et A_l distincts. Il y a donc au plus card $\{f \in Z^* | \varrho(f) = f, |f| \leqq \lambda\}$ valeurs de l'indice i pour lequel A_i est non vide. D'autre part si A_k est vide pour tout k compris entre n et $2n$, B_k donc R_k et A_k sont vides pour tout k supérieur à $2n$. Il existe donc certainement un N tel que pour tout $k > N$. R_k est vide.

Posons $R = \bigcup_{i=1}^{\infty} R_i$.

Etablissons d'autre part la propriété:

LEMME 3. *Le mot f appartient à B_{n+1} si et seulement si les deux conditions suivantes sont satisfaites:*

— *aucun facteur de f n'appartient à $\left(\bigcup_{i=1}^{i=n} \bar{R}_i \right)$*

— *il existe un schéma binaire de parenthèses β tel que tout segment f' de f parenthésé par β vérifie $|\varrho(\varphi(f'))| \leqq \lambda$.*

DEMONSTRATION. Nous faisons une récurrence sur n.

Soit $f \in B_{n+1}$: $f = g_1 g_2$, $g_1 \in A_k$, $g_2 \in A_l$, $k+l = n+1$. D'après l'hypothèse de récurrence il existe deaux schémas de parenthèses pour g_1 et g_2 soient β_1 et β_2 satisfaisant la condition du lemme. On prendra pour β le schéma $(\beta_1 \beta_2)$. D'autre part f ne contient pas de facteur dans $\bigcup_{i=1}^{i=n} \bar{R}_i$ par définition.

Réciproquement suppons que f satisfasse les deux conditions du lemme avec $\beta = (\beta_1 \beta_2)$, $\beta_1 \in S_k$, $\beta_2 \in S_l$ $k+l = n+1$. Par récurrence les segments de f correspondants aux sous schémas $[(\ , \beta_1, \beta_2)]$ et $[(\beta_1, \beta_2, \ )]$ de β et satisfaisant les conditions du lemme sont respectivement dans B_k et B_l, mais aussi dans A_k et A_l sans quoi f aurait des facteurs dans $\bigcup_{i=1}^{i=n} \bar{R}_i$. Finalement $f \in B_{n+1}$ QED.

Interprétation de l'algorithme.

Soit R l'ensemble de relations $\{f_i \equiv g_i | f_i, g_i \in X^*, \ i \in I\}$. Notons $f \equiv g(R)$ le fait que f et g sont congrus modulo la congruence σ la plus grossière sur X^* telle que $\forall i \in I$: $f_i \sigma g_i$.

Une telle congruence est une congruence de THUE: c'est la congruence de THUE engendrée par R.

Nous pouvons énoncer: en posant $R = \bigcup_{i=1}^{\infty} R_i$.

LEMME 4. *Les deux conditions suivantes sont équivalentes:*

1) $f \equiv e(R)$,

2) $\varrho(\varphi(f)) = e$.

L'implication $f \equiv e(R) \Leftrightarrow \varrho(\varphi(f)) = e$ est immédiate puisque par définition si $f' \equiv g' \in R$ on a $\varrho(\varphi(f')) = \varrho(\varphi(g'))$.

Réciproquement faisons une récurrence.

$\varrho(\varphi(f)) = e \Leftrightarrow f \equiv e(R)$ est vrai pour $|f| = 1$ par définition de R_1.

Supposons le donc vrai pour tout f': $|f'| < |f| = n+1$.

Deux cas se présentent:

— ou f contient comme facteur propre un membre gauche de relation de R. Ainsi si $f_2 \equiv f_2' \in R$

$$f = f_1 f_2 f_3 \equiv f_1 f_2' f_3 (R)$$

$$\varrho(\varphi(f)) = \varrho(\varphi(f_1 f_2' f_3)) = e$$

Or $f_1 f_2' f_3$ étant plus court que f, par récurrence $f_1 f_2' f_3 \equiv e(R)$ d'où $f \equiv e(R)$,

— ou f ne contient pas de tel facteur. D'après le lemme 2 il existe un schéma de parenthèse β tel que pour tout segment f' de f parenthésé par β: $|\varrho(\varphi(f'))| \leq \lambda$.

f satisfait alors aux conditions du lemme 3. D'où $f \in B_{n+1}$ et comme $\varrho(\varphi(f)) = e$, la relation $f \equiv e$ appartient à R_{n+1}. QED.

Remarque 1. La remarque suivante est fondamentale dans la suite: il est certain que si $f \equiv g$ est une relation de R tel que pour tout $h_1, h_2 \in X^*$: $\varrho(\varphi(h_1, gh_2)) \neq e$ on peut supprimer cette relation, dite inutile, puisqu'il est impossible que f ou g soient facteurs d'un mot congru à e modulo (R).

Or nous avons un algorithme pour décider si une relation dans est inutile. D'après le lemme 2, en effet, pour que $f \equiv g$ soit utile il faut et il suffit qu'il existe une suite de mots $g_1, \ldots, g_p$ de X^* tels que:

1) $g_1 = g$

2) $g_{i+1} = ag_i$ ou $g_i a$ avec $a \in \bigcup_{i=1}^{\infty} A_k$

3) $|\varrho(g_i)| \leqq \lambda$ et $i \neq j \Rightarrow \varrho(g_i) \neq \varrho(g_i)$

4) $\varrho(g_p) = e$.

Nous supposerons toujours désormais que R ne contient que des relations utiles.

Remarque 2. Les résultats des deux paragraphes précédents peuvent se résumer en le

THEORÈME 1. *Pour tout homomorphisme φ de X^* dans Z^* où $Z = \{y_i^\varepsilon\}$ et Z^* est muni de l'application ϱ, l'énsemble des mots f de X^x tel que $\varrho(\varphi(f))$ soit le mot vide de Z^* est égal à la classe d'équivalence du mot vide de X^* pour une congruence de Thue, sur X^*, finiment engendrée.*

III — *t*-langages

Définition 3. Le langage $L \subset X^*$ est un *t*-langage si et seulement si il existe un alphabet $Z = \{y_i^\varepsilon\}$ un homomorphisme φ de X^* dans Z^*, muni de l'application ϱ, et un mot $\gamma \in Z^*$ tels que $L = \{f \in X^* | \varrho(\gamma\varphi(f)) = e\}$.

Nous avons le théorème qui justifie leur considération en ce lieu.

THÉORÈME 2. *Le problème de l'équivalence est décidable pour deux t-langages.* Ce qui signifie que étant donné deux *t*-langages L et L' sur X^* respectivement définis par l'homomorphisme φ de X^* dans Z^* et γ dans Z^* et l'homomorphisme φ' de X'^* et γ' dans Z'^* il existe un algorithme qui permet en un nombre fini d'étapes de savoir si $L = L'$.

Considérons alors $L = \{f \in X^* | \varrho(\gamma\varphi(f)) = e\}$.

Soit $\omega \notin X$ et $X' = X \cup \{\omega\}$.

Posons $\varphi(\omega) = \gamma$ étendant ainsi l'homomorphisme $\varphi: X^* \to Z^*$ en un homomorphisme $\varphi: X'^* \to Z^*$. L'algorithme du paragraphe II nous permet de construire l'ensemble fini S de relations utiles tel que

$$\forall g \in X'^*: \varrho(\varphi(g)) = e \Leftrightarrow g \equiv e(S)$$

S est aussi tel que

$$\forall f \in X^*: \varrho(\gamma\varphi(f)) = e \Leftrightarrow \omega f \equiv e(S)$$

De la même façon nous pouvons construire un ensemble S' de relations utiles sur X'^* telles que

$$\forall f \in X^* \quad \varrho(\gamma'\varphi'(f)) = e \Leftrightarrow \omega f \equiv e(S')$$

Des conditions nécessaires et suffisantes pour que $L \subset L'$ sont alors:

1) $(f \equiv g) \in S$, $f \in X^*$, $g \in X^* \Rightarrow \varrho(\varphi'(f)) = \varrho(\varphi'(g))$.

2) $(\omega f \equiv g) \in S$, $f \in X^*$, $g \in X^* \Rightarrow \varrho(\gamma'\varphi'(f)) = \varrho(\varphi'(g))$.

3) $(\omega f = \omega g) \in S$, $f \in X^*$, $g \in X^* \Rightarrow \varrho(\gamma'\varphi'(f)) = \varrho(\gamma'\varphi'(g))$.

Ces conditions sont suffisantes puisque si elles son vérifiées et si $f \in L$ nous avons $\omega f \equiv e(S)$ donc il existe une suite $f_1 = \omega f, f_2, \ldots, f_p = e$ de mots de X'^* tels que f_{i+1} se déduit de f_i par substitution d'un membre à l'autre d'une relation de S ayant l'une des formes $f \equiv g$, $\omega f \equiv g$ ou $\omega f \equiv g$. Réciproquement, supposons l'une de ces trois conditions au moins non vérifiée, soit par exemple:

$$f \equiv g(S), \quad \varrho'(\varphi'(f)) \neq \varrho'(\varphi'(g)).$$

Il existe alors, puisque $f \equiv g$ est une relation utile, $h_1, h_2 \in X^*$ tels que $\varrho(\varphi(h_1 g h_2)) = e$, donc $h_1 g h_2 \in L$. Le mot $h_1 g h_2$ congru à $h_1 g h_2$ modulo (S) est aussi dans L. Or $h_1 f h_2$ et $h_1 g h_2$ ne peuvent être tous les deux dans L' puisque d'après notre hypothèse $\varrho(\varphi'(h_1 f h_2)) \neq \varrho(\varphi'(h_1 g h_2))$ et donc au plus un des mots

$$\varrho(\gamma'\varphi'(h_1 f h_2)) \quad \text{et} \quad \varrho(\gamma'(h_1 g h_2)) \quad \text{est égal au mot vide.}$$

D'où la nécessité et ainsi un algorithme permettant de décider si $L \subset L'$. En renversant les rôles, nous obtenons un algorithme pour décider si $L' \subset L$ et finalement le théorème.

BIBLIOGRAPHIE

[1] MAGNUS, KARASS, SOLITAR: *Combinatorial Group Theory*, Interscience Publisher, 1966.
[2] NIVAT, M.: Transductions des langages de Chomsky. *Annales del'Institut Fourier*, **18** (1968) 339—456.

Université de Rennes, Faculté des Sciences
Adresse actuelle: Université de Paris 7, 2 place Jussieu, Paris 5, France

(Reçu le 14 Mai 1968)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 251—261.

SOME INEQUALITIES CONCERNING POLYNOMIALS HAVING ONLY REAL ZEROS

by

Á. ELBERT

Throughout this paper, the symbol f denotes a polynomial

$$(1)\qquad f(x)=(x+1)^{n_1}(x-1)^{n_2}\prod_{i=1}^{n_3}(x-x_i)\qquad(-1\leqq x_1\leqq x_2\leqq\cdots\leqq x_{n_3}\leqq 1),$$

where n_1, n_2, n_3 are nonnegative integers and $n_1+n_2+n_3=n>0$, x real. The symbol $\|f\|$ denotes the norm of f:

$$\|f\|=\max_{-1\leqq x\leqq 1}|f(x)|.$$

We are concerned with the connection of quantities $\|f\|$, n_1, n_2 and the centroid of the roots

$$(2)\qquad \bar{x}=\frac{-n_1+n_2+\sum_{i=1}^{n_3}x_i}{n}.$$

The quantities n_1, n_2 and $\bar{x}$ are not completely independent, because by the aid of the relation $|x_i|\leqq 1$ we have the inequality

$$(3)\qquad -1+\frac{2n_2}{n}\leqq\bar{x}\leqq 1-\frac{2n_1}{n}.$$

Introducing the function $\mu(\alpha,\beta)$ by

$$(4)\qquad \mu(\alpha,\beta)=(1+\alpha+\beta)^{1+\alpha+\beta}(1+\alpha-\beta)^{1+\alpha-\beta}(1-\alpha+\beta)^{1-\alpha+\beta}(1-\alpha-\beta)^{1-\alpha-\beta}$$

for $0\leqq\alpha,\beta,\alpha+\beta\leqq 1$*, we can formulate our statements as follows.

THEOREM 1. *For a polynomial* (1) *the inequality*

$$\|f\|\geqq\frac{1}{2^n}\left\{\mu\left(\frac{n_1}{n},\frac{n_2}{n}\right)\right\}^{\frac{n}{2}}$$

holds, where the the equality holds if and only if $n_3=0$.**

* If it is needed, the relation $0^0=1$ must be taken into account.

** This theorem can be extended to those polynomials, in which the factor $\prod_{i=1}^{n_3}(x-x_i)$ is replaced by $x^{n_3}+a_1x^{n_3-1}+\ldots+a_{n_3}$.

THEOREM 2. *If* $|\bar{x}| \geqq \alpha_0^2$, *where* $0 \leqq \alpha_0 \leqq 1$, *then*

$$\|f\| \geqq \frac{1}{2^n} \{\mu(\alpha_0, 0)\}^{\frac{n}{2}},$$

and the equality holds if and only if $\alpha_0 = 1$, *i.e. either* $n_1 = n$ *or* $n_2 = n$.

THEOREM 3. *Let the quantities* n_1, n_2 *and* $|\bar{x}| = \alpha_0^2$ *be prescribed, and let us suppose that* $n_1 \geqq n_2$, *then*

$$\|f\| \geqq \begin{cases} \dfrac{1}{2^n}\left\{\mu\left(\sqrt{\alpha_0^2 + \left(\dfrac{n_2}{n}\right)^2}, \dfrac{n_2}{n}\right)\right\}^{\frac{n}{2}} & \text{if } n_1^2 - n_2^2 \leqq \alpha_0^2 n^2 \\ \dfrac{1}{2^n}\left\{\mu\left(\dfrac{n_1}{n}, \sqrt{\left(\dfrac{n_1}{n}\right)^2 - \alpha_0^2}\right)\right\}^{\frac{n}{2}} & \text{if } n_1^2 - n_2^2 > \alpha_0^2 n^2, \end{cases}$$

and the equality holds if and only if $\alpha_0^2 = 1 - \dfrac{2n_2}{n}$.* *In the cases* $n_1 < n_2$ *the roles of* n_1 *and* n_2 *here are to be interchanged.*

Remark 1. It is well known, that

$$\|f\| \geqq 2^{-n+1} \tag{5}$$

and this minimum is attained only by the Tshebysheff polynomials $T_n(x) = 2^{-n+1} \cos n\,(\arccos x)$. For these polynomials we have by Theorem 1 $\|T_n\| > 2^{-n}\{\mu(0, 0)\}^{n/2} = 2^{-n}$.

Remark 2. It is easy to verify that if $n_3 = 0$, i.e. all zeros are lying at the end-points $x = \pm 1$ then in the theorems above the equality sign holds.

Remark 3. The polynomial (1) can be written by the aid of (2) in the form $x^n - n\bar{x} \cdot x^{n-1} + \ldots$, too, ZOLOTAREFF [1] considered the quantities $\min \|x^n - \sigma x^{n-1} + \ldots\|$ for fixed σ with no other restrictions concerning the zeros and he expressed them by means of elliptic functions. The value of this minimum is lying in the interval $[2^{-n+1}, 2^{-n+2})$; we see from our theorem that by the conditions $|x_i| < 1$ the minima are increased.

Remark 4. A consequence the the Theorem 1 is already proved in our paper [2].

In order to prove of theorems above, first we consider the quantity

$$M_{n,\alpha_0}(n_1, n_2) = \min_{\substack{f \\ |\bar{x}| = \alpha_0^2}} \|f\|. \tag{6}$$

It is clear that for every fixed n_1, n_2 and for all admissible ** α_0 there is an extremal

* Using (3) it can be shown that the sum of the two arguments of μ here is not greater than 1.

** W ecall an α_0 admissible if there is at least one polynomial (1) with centroid $\bar{x}$ fulfilling the requirement $|\bar{x}| = \alpha_0^2$ and the relation (3).

polynomial f_0 satisfying the relations

$$M_{n,\alpha_0}(n_1, n_2) = \|f_0\|$$

$$(7) \qquad f_0(x) = (x+1)^{n_1^{(0)}}(x-1)^{n_2^{(0)}} \prod_{i=1}^{n_3^{(0)}} (x - x_i^{(0)}) \qquad |x_i^{(0)}| < 1$$

$$\alpha_0^2 = \left| -\frac{n_1^{(0)}}{n} + \frac{n_2^{(0)}}{n} + \frac{1}{n} \sum_{i=1}^{n_3^{(0)}} x_i^{(0)} \right|$$

$$n_1^{(0)} \geqq n_1, \quad n_2^{(0)} \geqq n_2.$$

Let us now consider the polynomial f_0^2. This polynomial is of degree $2n$ and has the form (1) with $n_1 = 2n_1^{(0)}$ and $n_2 = 2n_2^{(0)}$ and its centroid is the same as of f_0, therefore by (6)

$$(8) \qquad M_{2n,\alpha_0}(2n_1^{(0)}, 2n_2^{(0)}) \leqq \|f_0\|^2 = M_{n,\alpha_0}^2(n_1, n_2).$$

We have again an extremal polynomial f_1 of degree $2n$ which fulfils the relation

$$M_{2n,\alpha_0}(2n_1^{(0)}, 2n_2^{(0)}) = \|f_1\|.$$

By the aid of this polynomial f_1 we have similarly a polynomial f_2 of degree $4n$, and so on. We have arrived on this way to the polynomial f_ν of degree $n^{(\nu)} = 2^\nu n$ $(\nu = 1, 2, \ldots)$ which has the form

$$(9) \quad f_\nu(x) = (x+1)^{n_1^{(\nu)}}(x-1)^{n_2^{(\nu)}} \prod_{i=1}^{n_3^{(\nu)}} (x - x_i^{(\nu)}) \qquad |x_i^{(\nu)}| < 1 \ \ (i = 1, \ldots, n_3^{(\nu)}),$$

where

$$(10) \qquad \alpha_0^2 = \left| -\frac{n_1^{(\nu)}}{n^{(\nu)}} + \frac{n_2^{(\nu)}}{n^{(\nu)}} + \frac{1}{n^{(\nu)}} \sum_{i=1}^{n_3^{(\nu)}} x_i^{(\nu)} \right|, \qquad n_1^{(\nu)} + n_2^{(\nu)} + n_3^{(\nu)} = n^{(\nu)}$$

and

$$(11) \qquad M_{2^\nu n,\alpha_0}(2n_1^{(\nu-1)}, 2n_2^{(\nu-1)}) = \|f_\nu\| \leqq \|f_{\nu-1}\|^2.$$

Concerning the zeros $\{x_i^{(\nu)}\}$ we state here that they are, if they exist at all, simple (henceforth we assume that they are enumerated in increasing order) and there are points $\xi_i^{(\nu)}$ such that

$$(12) \qquad |f_\nu(\xi_i^{(\nu)})| = \|f_\nu\| \qquad x_i^{(\nu)} < \xi_i^{(\nu)} < x_{i+1}^{(\nu)}, \ i = 1, 2, \ldots, n_3^{(\nu)} - 1.$$

Supposing the contrary we would have $|f_\nu(x)| < \|f_\nu\|$ in $[x_k^{(\nu)}, x_{k+1}^{(\nu)}]$ for some $1 \leqq k \leqq n_3^{(\nu)} - 1$. Let us consider the polynomial $f_{\nu,\varepsilon}(x) = f_\nu(x) \cdot (x - x_k^{(\nu)} + \varepsilon)(x - x_{k+1}^{(\nu)} - \varepsilon) / [(x_k - x_k^{(\nu)})(x - x_{k+1}^{(\nu)})]$ for sufficiently small $\varepsilon > 0$, then all zeros of $f_{\nu,\varepsilon}$ are in $[-1, 1]$, the centroid of its zeros is the same as those of f_ν, the multiplicities of the zeros at the endpoints are unchanged and finally $f_{\nu,\varepsilon}(x)/f_\nu(x) < 1$ for $x \notin (x_k^{(\nu)} - \varepsilon, x_{k+1}^{(\nu)} + \varepsilon)$. This and the continuous dependence on ε of $f_{\nu,\varepsilon}$ in $[x_k^{(\nu)} - \varepsilon, x_{k+1}^{(\nu)} + \varepsilon]$ supply the existence at least one value of ε for which $\|f_{\nu,\varepsilon}\| < \|f_\nu\|$, in contradiction to the extremality of f_ν.

A consequence of (12) is that if $n_3^{(\nu-1)} > 0$ then the polynomial $f_{\nu-1}^2$ can not be extremal for $M_{2^\nu n,\alpha_0}(2n_1^{(\nu-1)}, 2n_2^{(\nu-1)})$, hence the relation (11) can be completed by

$$(11^*) \qquad \|f_{\nu-1}\|^2 \underset{(=)}{>} \|f_\nu\| \quad \text{if} \quad n_3^{(\nu-1)} \underset{(=)}{>} 0.$$

Let us introduce the quantities $\alpha^{(\nu)}$, $\beta^{(\nu)}$, $\gamma^{(\nu)}$ and M_ν by

$$(13)\qquad \alpha^{(\nu)} = \frac{n_1^{(\nu)}}{n^{(\nu)}},\quad \beta^{(\nu)} = \frac{n_2^{(\nu)}}{n^{(\nu)}},\quad \gamma^{(\nu)} = \frac{n_3^{(\nu)}}{n^{(\nu)}},\quad M_\nu = \|f_\nu\|^{\frac{1}{n^{(\nu)}}}.$$

By (11) we have $\alpha^{(\nu)} \geqq \alpha^{(\nu-1)}$, $\beta^{(\nu)} \geqq \beta^{(\nu-1)}$, $M_\nu \leqq M_{\nu-1}$, therefore the following limits exist:

$$(14)\qquad \alpha = \lim_{\nu\to\infty} \alpha^{(\nu)} \geqq \alpha^{(0)},\quad \beta = \lim_{\nu\to\infty} \beta^{(\nu)} \geqq \beta^{(0)},\quad \gamma = \lim_{\nu\to\infty} \gamma^{(\nu)} = 1-\alpha-\beta,$$

$$(15)\qquad M = \lim_{\nu\to\infty} M_\nu \leqq M_0.$$

By (5) we have $M \geqq 1/2$. We are going to establish the relation

$$(16)\qquad M = \frac{1}{2}\sqrt{\mu(\alpha, \beta)}.$$

At first we shall prove this in the case $\gamma = 0$. If there is an $n_3^{(\nu-1)} = 0$ then by (11) $n_3^{(\lambda)} = 0$, $\alpha^{(\lambda)} = \alpha$, $\beta^{(\lambda)} = \beta$ and by (9) $|f_\lambda|^{1/n^{(\lambda)}} = |x+1|^\alpha |x-1|^\beta$ for all $\lambda \geqq \nu-1$, hence $M_\lambda = M = \|(1+x)^\alpha(1-x)^\beta\| = \frac{1}{2}\sqrt{\mu(\alpha, \beta)}$, which proves (16) in this case.

Now we can assume that $\gamma^{(\nu)} > 0$ and $\gamma^{(\nu)} \to 0$. Let

$$\xi^{(\nu)} = \alpha^{(\nu)} - \beta^{(\nu)} \quad\text{and}\quad \xi = \alpha - \beta$$

then $\|(1+x)^\alpha(1-x)^\beta\| = (1+\xi)^\alpha(1-\xi)^\beta = \sqrt{\mu(\alpha, \beta)}/2$. From (9) and (13) it follows

$$M_\nu = \left\| (1+x)^{\alpha^{(\nu)}}(1-x)^{\beta^{(\nu)}} \left[\prod_{i=1}^{n_3^{(\nu)}} (x - x_i^{(\nu)})\right]^{\frac{1}{n^{(\nu)}}} \right\| < \|(1+x)^{\alpha^{(\nu)}}(1-x)^{\beta^{(\nu)}}\| \cdot 2^{\gamma^{(\nu)}}$$

hence by (14) and (15)

$$(17)\qquad M \leqq \frac{1}{2}\sqrt{\mu(\alpha, \beta)}.$$

According to a result of G. PÓLYA [3] for a polynomial

$$g(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0 \quad (a_i \text{ are real, } n \geqq 1)$$

the inequality $|g(x)| > 1$ holds except a set being composed of intervals which has measure less than 4. As a consequence of this result we have

$$(18)\qquad \operatorname{Mes}\{x; |g(x)| \leqq \varepsilon^n, \varepsilon > 0\} < 4\varepsilon.$$

Hence we have a point ζ_ν in $(\xi^{(\nu)} - \gamma^{(\nu)}, \xi^{(\nu)} + \gamma^{(\nu)}) \subset [-1, 1]$ for which

$$\left|\prod_{i=1}^{n_3^{(\nu)}} (\zeta_\nu - x_i^{(\nu)})\right| > \left[\frac{\gamma^{(\nu)}}{2}\right]^{n_3^{(\nu)}},$$

therefore by (9) and (13)

$$M_\nu > (\zeta_\nu + 1)^{\alpha^{(\nu)}}(1-\zeta_\nu)^{\beta^{(\nu)}} \left[\frac{\gamma^{(\nu)}}{2}\right]^{\gamma^{(\nu)}},$$

hence by (14) and (15) and

$$M \geqq (1+\xi)^{\alpha}(1-\xi)^{\beta} = \frac{1}{2}\sqrt{\mu(\alpha,\beta)}\,,$$

which completes the proof (16) by (17) in the case $\gamma=0$.

Now it remains the case $\gamma>0$. Since the sequence $\{\gamma^{(\nu)}\}$ is nonincreasing therefore $n_3^{(\nu)}=\gamma^{(\nu)}n^{(\nu)}\geqq\gamma n^{(\nu)}$. Let $n_3^{(\nu)}\geqq 2$. By an inequality due to BERNSTEIN [4] we have with our notations $|f_\nu'| \leqq n^{(\nu)}\|f_\nu\|/\sqrt{1-x^2}$ and by (12)

$$|f_\nu(\xi_i^{(\nu)})| \leqq n^{(\nu)}\|f_\nu\|\, I(x_i^{(\nu)},\xi_i^{(\nu)}) \quad \text{and} \quad |f_\nu(\xi_i^{(\nu)})| \leqq n^{(\nu)}\|f_\nu\|\, I(\xi_i^{(\nu)},x_{i+1}^{(\nu)}),$$

where $I(x,y) = \int\limits_x^y dt/\sqrt{1-t^2}$, hence

$$I(x_i^{(\nu)},\xi_i^{(\nu)}) \geqq \frac{1}{n^{(\nu)}}, \quad I(\xi_i^{(\nu)},x_{i+1}^{(\nu)}) \geqq \frac{1}{n^{(\nu)}} \qquad (i=1,2,\dots,n_3^{(\nu)}-1). \tag{19}$$

Let us introduce the functions

$$\varrho_\nu(x) = \begin{cases} \dfrac{1}{n^{(\nu)}(x_{i+1}^{(\nu)}-x_i^{(\nu)})} & x_i^{(\nu)}\leqq x < x_{i+1}^{(\nu)} \qquad i=1,\dots,n_1^{(\nu)}-1 \\ 0 & x<x_1^{(\nu)},\quad x\geqq x_{n_3^{(\nu)}}^{(\nu)} \end{cases}$$

(20)

$$\Gamma_\nu(x) = \int\limits_{-1}^{x} \varrho_\nu(t)\,dt.$$

As a consequence of this definition we have

$$\Gamma_\nu(x_i^{(\nu)}) = \frac{i-1}{n^{(\nu)}} \qquad (i=1,\dots,n_3^{(\nu)}), \tag{21}$$

moreover the function $\varrho_\nu(x)$ fulfils the relation

$$\varrho_\nu(x) < \frac{\pi}{2\sqrt{1-x^2}} \qquad (-1<x<1), \tag{22}$$

because by (19) we have $2/n^{(\nu)} \leqq I(x_i^{(\nu)},x_{i+1}^{(\nu)})$, and putting $x=\cos\theta$, $x_i^{(\nu)}=\cos\theta_i$ $(\pi>\theta_1>\theta_2>\dots>\theta_{n_3^{(\nu)}}^{(\nu)}>0)$ it is sufficient to show by (20) that

$$\frac{\theta_i-\theta_{i+1}}{2(\cos\theta_{i+1}-\cos\theta_i)} < \frac{\pi}{\sin\theta} \qquad (\theta_{i+1}\leqq\theta\leqq\theta_i).$$

But

$$\max_{0<\theta_{i+1}\leqq\theta\leqq\theta_i<\pi} \frac{\sin\theta}{\sin\dfrac{\theta_i+\theta_{i+1}}{2}}\;\frac{\dfrac{\theta_i-\theta_{i+1}}{2}}{\sin\dfrac{\theta_i-\theta_{i+1}}{2}} < 2\cdot\frac{\pi}{2},$$

hence (22) is true.

By (20) and (22) the sequence $\{\Gamma_\nu(x)\}_{\nu=0}^{\infty}$ is equicontinuous, hence by ARZELA's theorem there is a convergent subsequence:

(23) $$\lim_{s\to\infty} \Gamma_{\nu_s}(x) = \Gamma(x) \qquad (-1 \leqq x \leqq 1).$$

The limit function $\Gamma(x)$ is continuous, nondecreasing and fulfils the relation

(24) $$\Gamma(x'') - \Gamma(x') \leqq \frac{\pi}{2} \int_{x'}^{x''} \frac{dt}{\sqrt{1-t^2}} \qquad (-1 \leqq x' \leqq x'' \leqq 1).$$

From (20) and (21) we have $\Gamma(-1) = 0$, $\Gamma(1) = \gamma$. We define

(25) $$a = \max\{x; \Gamma(x) = 0, x \geqq -1\}, \quad b = \min\{x; \Gamma(x) = \gamma, x \leqq 1\}.$$

We are going to show that

(26) $$a \underset{(=)}{>} -1 \text{ if } \alpha \underset{(=)}{>} 0 \quad \text{and} \quad b \underset{(=)}{<} 1 \text{ if } \beta \underset{(=)}{>} 0.$$

It is sufficient to treat the first inequality with a. Let first $\alpha > 0$, then we can choose a value $a_1 \in (-1, 0)$ such that $(1+x)^{\alpha/2}(1-x)^{1-\alpha/2} < M$ if $-1 \leqq x \leqq a_1$. For all sufficiently large ν we have $\alpha^{(\nu)} > \alpha/2$, therefore by (9)

$$|f_\nu(x)|^{\frac{1}{n^{(\nu)}}} < (1+x)^{\alpha^{(\nu)}}(1-x)^{1-\alpha^{(\nu)}} < (1+x)^{\frac{\alpha}{2}}(1-x)^{1-\frac{\alpha}{2}} < M$$

in $[-1, a_1]$, hence by (11*), (13) and (15) $|f_\nu(x)| < \|f_\nu\|$ here, by (12) $\Gamma_\nu(a_1) < 1/n_{(\nu)}$, consequently $\Gamma(a_1) = 0$, i.e. $a \geqq a_1 > -1$. On the other hand if $\alpha = 0$ then $\alpha^{(\nu)} = 0 = n_1^{(\nu)}$ because the sequence $\{\alpha^{(\nu)}\}$ is nondecreasing. Let us assume the contrary $a > -1$, then let $[-1, a]$ contain the zeros $x_1, \dots, x_{j_s - 1}$ (henceforth we drop the indices ν_s if the notations remain unambiguons). By (21) and (25) $\lim_{s\to\infty} j_s/n = \Gamma(a) = 0$. Let $\varepsilon \in (0, a+1)$ an arbitrary fixed number, then (18) guarantees a value $\bar\zeta_s \in [-1, -1+\varepsilon]$ with

$$\prod_{i=1}^{j_s} |\bar\zeta_s - x_i| > \left(\frac{\varepsilon}{4}\right)^{j_s},$$

then by (9)

(27) $$\|f_{\nu_s}\| \geqq |f_{\nu_s}(\bar\zeta_s)| > |\bar\zeta_s - 1|^{n_2} \left(\frac{\varepsilon}{4}\right)^{j_s} \prod_{i=j_s+1}^{n_3} |\bar\zeta_s - x_i|.$$

An other inequality can be derived from (12)

$$\|f_{\nu_s}\| = |f_{\nu_s}(\xi_{j_s})| < |\xi_{j_s} - 1|^{n_2} 2^{j_s} \prod_{i=j_s+1}^{n_3} |\xi_{j_s} - x_i|,$$

comparing this with (27) and using the inequalities $-1 \leqq \bar\zeta_s \leqq -1 + \varepsilon < a < x_{j_s} < \xi_{j_s} < x_i < 1$ we have

$$\left(\frac{1-a}{2-\varepsilon}\right)^{n_2} \prod_{i=j_s+1}^{n_3} \frac{x_i - a}{x_i + 1 - \varepsilon} > \left(\frac{\varepsilon}{8}\right)^{j_s},$$

but $(x_i - a)/(x_i + 1 - \varepsilon) < (1-a)/(2-\varepsilon)$, hence by extraction of nth root of both sides, letting s to ∞ we obtain $1 - a \geqq 2 - \varepsilon$ or $a \leqq -1 + \varepsilon$, which contradicts to the choice of ε and this completes the proof of (26).

Let $\psi_\nu(x)$ be defined by

$$\psi_\nu(x) = \int_{-1}^{1} \log|x-t| \cdot \varrho_\nu(t)\,dt. \tag{28}$$

The sequence $\{\psi_\nu(x)\}_{\nu=0}^{\infty}$ is equicontinuous, namely

$$|\psi_\nu(x+h)-\psi_\nu(x)| < \eta(h) \quad \text{for} \quad 0 \leqq h < \frac{3-\sqrt{5}}{2}, \tag{29}$$

where

$$\eta(h) = \sqrt{h} + 2\pi \int_{1-h-\sqrt{h}}^{1} \log\frac{1}{1-t}\frac{dt}{\sqrt{1-t^2}}, \tag{30}$$

and the bound $(3-\sqrt{5})/2$ is explained by the requirement $h+\sqrt{h} < 1$.

To show (29) we assume $-1 < x-\sqrt{h} < x+h+\sqrt{h} < 1$ (the proof for other values of x goes on similar way), then

$$|\psi_\nu(x+h)-\psi_\nu(x)| \leqq \int_{-1}^{x-\sqrt{h}} |\log(x+h-t)-\log(x-t)|\,\varrho_\nu(t)\,dt +$$
$$+ \int_{x-\sqrt{h}}^{x+h-\sqrt{h}} \big|\log|x-t|+\log|x+h-t|\big|\,\varrho_\nu(t)\,dt +$$
$$+ \int_{x+h+\sqrt{h}}^{1} |\log(t-x-h)-\log(t-x)|\,\varrho_\nu(t)\,dt.$$

Using the inequality $\log(1+x) < x$ we have

$$\int_{-1}^{x-\sqrt{h}} + \int_{x+h+\sqrt{h}}^{1} < \sqrt{h}\int_{-1}^{1} \varrho_\nu(t)\,dt < \sqrt{h},$$

and by (22)

$$\int_{x}^{x+h+\sqrt{h}} |\log(t-x)|\,\varrho_\nu(t)\,dt < \frac{\pi}{2}\int_{x}^{x+h+\sqrt{h}} |\log(t-x)|\frac{dt}{\sqrt{1-t^2}} \leqq \frac{\pi}{2}\int_{1-h-\sqrt{h}}^{1} \frac{|\log(1-t)|}{\sqrt{1-t^2}}\,dt,$$

$$\int_{x-h}^{x} |\log(x-t)|\,\varrho_\nu(t)\,dt < \frac{\pi}{2}\int_{1-h-\sqrt{h}}^{1} \frac{|\log(1-t)|}{\sqrt{1-t^2}}\,dt,$$

and estimating the integral of $\log|x+h-t|$ similarly we obtain (30).

To determine the connection between $\psi_\nu(x)$ and f_ν we recapitulate a result of F. Wenzl (see [5], especially Satz 2):

Lemma. If $P(x) = (x-x_1)(x-x_2)\dots(x-x_n)$, $-1 \leqq x_1 < x_2 < \dots < x_n \leqq 1$ and

$$\varrho^*(x) = \begin{cases} \dfrac{1}{n(x_{i+1}-x_i)} & \text{for} \quad x_i \leqq x \leqq x_{i+1} \quad (i=1, \dots, n-1) \\ 0 & \text{otherwise,} \end{cases}$$

then

$$\frac{1}{n}\log|P(x)| = \int_{-1}^{1}\log|x-t|\,\varrho^*(t)\,dt+\vartheta^* \quad \text{for} \quad x_i<x<x_{i+1},$$

where

$$\vartheta^* = \frac{\delta_1}{n}\log(1+x)+\frac{\delta_2}{n}\log(1-x)+\frac{\vartheta_1}{n}\log(x-x_i)+\frac{\vartheta_2}{n}\log(x_{i+1}-x)+\frac{1}{n}$$

and

$$0<\delta_1,\delta_2<\frac{1}{2},\quad 0<\vartheta_1<1,\quad -\frac{1}{2}<\vartheta_2<\frac{1}{2}.$$

Let the set $I_\nu(\varepsilon, c)$ be defined for $0<\varepsilon,\ c<1$ by

$$I_\nu(\varepsilon,c)=[-1,-1+\varepsilon]\cup[1-\varepsilon,1]\cup\bigcup_{i=1}^{n_3^{(\nu)}}\left[x_i^{(\nu)}-\frac{c}{n^{(\nu)}},\ x_i^{(\nu)}+\frac{c}{n^{(\nu)}}\right], \tag{31}$$

then by (20) and by the Lemma we have for $x\in[-1,1]\setminus I_\nu(\varepsilon,c)$

$$\frac{1}{n_3^{(\nu)}}\log\prod_{i=1}^{n_3^{(\nu)}}|x-x_i^{(\nu)}| = \int_{-1}^{1}\log|x-t|\frac{\varrho_\nu(t)}{\gamma^{(\nu)}}\,dt+$$

$$+\theta\left\{\frac{1}{n_3^{(\nu)}}\log\frac{1}{\sqrt{1-(1-\varepsilon)^2}}+\frac{3}{2n_3^{(\nu)}}\log\frac{n_3^{(\nu)}}{c}\right\}$$

with suitable $\theta\in(-1,1)$, hence by (9) and (29) the required relation is

$$\begin{aligned}\frac{1}{n^{(\nu)}}\log|f_\nu(x)| &= \alpha^{(\nu)}\log(1+x)+\beta^{(\nu)}\log(1-x)+\psi_\nu(x)+\\ &+\theta\left\{\frac{1}{n^{(\nu)}}\log\frac{1}{\sqrt{1-(1-\varepsilon)^2}}+\frac{3}{2n^{(\nu)}}\log\frac{\gamma^{(\nu)}n^{(\nu)}}{c}\right\}\end{aligned}\tag{32}$$

for all $x\in[-1,1]\setminus I_\nu(\varepsilon,c)$ with $|\theta|<1$. It is clear the the term in brackets on the right hand side tends to 0 if $n^\nu\to\infty$.

According to the Helly—Bray theorem we have by (20) and (23)

$$\lim_{s\to\infty}\psi_{\nu_s}(x)=\lim_{s\to\infty}\int_{-1}^{1}\log|x-t|\,d\Gamma_{\nu_s}(t)=\int_{0}^{b}\log|x-t|\,d\Gamma(t)\overset{\text{def}}{=}\psi(x). \tag{33}$$

It is obvious that the function $\psi(x)$ is continuous, therefore taking into account the relations (13), (14) and (32) it follows directly

$$\varphi(x)\overset{\text{def}}{=}\alpha\log(1+x)+\beta\log(1-x)+\psi(x)\leqq\log M \qquad (-1\leqq x\leqq 1). \tag{34}$$

It remains to show the main relation

$$\varphi(x)=\log M \quad \text{for} \quad a\leqq x\leqq b. \tag{35}$$

We prove this first for such an $x_0\in[a,b]\cap(-1,1)$ which fulfils the relation $\Gamma(x)>\Gamma(x_0)$ for all $x>x_0$. Such a value is for example by (25) $x_0=a$ if $\alpha\neq 0$.

Let $x \in (x_0, 1)$ be fixed then we have for all sufficiently large s $n^{(v_s)}[\Gamma_{v_s}(x) - \Gamma_{v_s}(x_0)] > > n^{(v_s)}\frac{1}{2}(\Gamma(x) - \Gamma(x_0)) > 2$, hence by (21) the interval $[x_0, x]$ contains at least two zeros of f_{v_s}, say, x_i and x_{i+1}. Let $\varepsilon = \min\{x_0 + 1, 1 - x\}$ and $c < \sqrt{\varepsilon(2-\varepsilon)}$, then $\xi_i \notin I_{v_s}(\varepsilon, c)$ because by (19)

$$\frac{1}{n^{(v_s)}} \leqq \int_{x_i}^{\xi_i} \frac{dt}{\sqrt{1-t^2}} < \frac{\xi_i - x_i}{\sqrt{1-(1-\varepsilon)^2}},$$

and similarly $x_{i+1} - \xi_i > \sqrt{\varepsilon(2-\varepsilon)}/n^{(v_s)}$ therefore by (12), (32), (34)

$$|\varphi(x_0) - \log M_{v_s}| = \left|\varphi(x_0) - \frac{1}{n^{(v_s)}} \log |f_{v_s}(\xi_i)|\right| <$$

$$< (\alpha - \alpha^{(v_s)}) |\log(1+x_0)| + (\beta - \beta^{(v_s)}) |\log(1-x_0)| + \alpha^{(v_s)} \left|\log \frac{1+x_0}{1+\xi_i}\right| +$$

$$+ \beta^{(v_s)} \left|\log \frac{1-x_0}{1-\xi_i}\right| + |\psi(x_0) - \psi_{v_s}(x_0)| + |\psi_{v_s}(x_0) - \psi_{v_s}(\xi_i)| +$$

$$+ \frac{1}{n^{(v_s)}} \log \frac{1}{\sqrt{\varepsilon(2-\varepsilon)}} + \frac{3}{2n^{(v_s)}} \log \frac{\gamma^{(v_s)} n^{(v_s)}}{c}.$$

Letting $s \to \infty$ this yields by (14), (33), (29)

$$|\varphi(x_0) - \log M| \leqq \alpha \left|\log \frac{1+x_0}{1+x}\right| + \beta \left|\log \frac{1-x_0}{1-x}\right| + \eta(x - x_0).$$

This estimation is valid for all $x > x_0$ in the vicinity of x_0. The function on the right hand side is continuous in x and vanishes at $x = x_0$, hence $\varphi(x_0) = \log M$. Similar consideration shows the validity of (35) if $\Gamma(x) < \Gamma(x_0)$ for all $x < x_0$. It remains now to prove that $\Gamma(x)$ is strictly increasing on $[a, b]$. Let us suppose the contrary, then there would be an interval $[u, v] \subset [a, b]$ where $\Gamma(x)$ would be constant. We may assume $u = \min\{x; \Gamma(x) = \Gamma(u), a \leqq x \leqq b\}$ and $v = \max\{x; \Gamma(x) = \Gamma(u), a \leqq x \leqq b\}$. For such u and v we proved already $\varphi(u) = \varphi(v) = \log M$, and we could write

$$\varphi(x) = \alpha \log(1+x) + \beta \log(1-x) + \int_a^u \log|x-t|\, d\Gamma(t) + \int_v^b \log|x-t|\, d\Gamma(t),$$

hence $\varphi(x)$ would be concave in (u, v), therefore

$$\varphi\left(\frac{u+v}{2}\right) > \frac{1}{2}[\varphi(u) + \varphi(v)] = \log M,$$

in contrary to (34). Since $\Gamma(x)$ has no interval included in (a, b), where it assumes a constant value the relation (35) holds.

The function $\Gamma(x)$ is absolutely continuous, it has a derivative $\gamma(x)$ almost everywhere with the properties

$$(36) \qquad \frac{d\Gamma(x)}{dx} = \gamma(x), \quad \int_a^b \gamma(x)\, dx = \gamma, \quad 0 \leqq \gamma(x) \leqq \frac{\pi}{2\sqrt{1-x^2}} \qquad \text{a.e. in } [-1, 1],$$

and by (34), (35)

$$(37)\qquad \alpha\log(1+x)+\beta\log(1-x)+\int_a^b \log|x-t|\,\gamma(t)\,dt = \log M \qquad (a\leqq x\leqq b).$$

This integral equation concerning $\gamma(x)$ has the solution (see [6])

$$\gamma(x) = \frac{\gamma}{\pi\sqrt{(b-x)(x-a)}}+\frac{1}{\pi^2\sqrt{(b-x)(x-a)}}\int_a^b\left(\frac{\alpha}{1+t}-\frac{\beta}{1-t}\right)\frac{\sqrt{(b-t)(t-a)}}{x-t}\,dt,$$

where the integral is defined as a Cauchy's principal value. By calculations we have

$$\gamma(x) = \frac{1-x^2-(1-x)\alpha\sqrt{(1+a)(1+b)}-(1+x)\beta\sqrt{(1+a)(1-b)}}{\pi(1-x^2)\sqrt{(b-x)(x-a)}}.$$

In the case $\alpha\beta\neq 0$ by (36) the numerator must vanish at $x=a$ and $x=b$, hence

$$(38)\qquad \gamma(x) = \frac{1}{\pi}\frac{\sqrt{(b-x)(x-a)}}{1-x^2} \qquad (a\leqq x\leqq b)$$

and

$$(39)\qquad \alpha = \frac{\sqrt{(1+a)(1+b)}}{2},\quad \beta = \frac{\sqrt{(1-a)(1-b)}}{2}$$

or

$$\left.\begin{matrix} b \\ a\end{matrix}\right\} = \alpha^2-\beta^2\pm\sqrt{(1+\alpha+\beta)(1+\alpha-\beta)(1-\alpha+\beta)(1-\alpha-\beta)}.$$

We can verify the validity of these formulae in the cases $\alpha\beta=0$, too.

Multiplying (37) by $1/\left(\pi\sqrt{(b-x)(x-a)}\right)$ and integrating over $[a, b]$ we have by (39)

$$M = \frac{1}{2}\sqrt{\mu(\alpha,\beta)},$$

as we stated in (16), hence the relation (16) holds in all cases.

Finally, we shall express the quantity α_0 by means of the newly introduced quantities. From (10) and (20) we have

$$\bar{x} = -\alpha^{(\nu)}+\beta^{(\nu)}+\frac{1}{n^{(\nu)}}\sum_{i=1}^{n_3^{(\nu)}} x_i^{(\nu)} = -\alpha^{(\nu)}+\beta^{(\nu)}+\int_{-1}^{1} x\,d\Gamma_\nu(x)+\frac{x_1^{(\nu)}+x_{n_3^{(\nu)}}^{(\nu)}}{2n^{(\nu)}},$$

hence by (23) and (38)

$$\bar{x} = -\alpha^2+\beta^2$$

therefore

$$(41)\qquad \alpha_0^2 = |\alpha^2-\beta^2|.$$

Having now all the required relations we can pass over to the proof of the theorems.

PROOF of Theorem 1. A polynomial (1) fulfils the relation

$$\|f\| \geqq M_{n,\alpha_0}(n_1, n_2) = \|f_0\| = M_0^n,$$

and by (15), (16) $\|f\| \geqq 2^{-n}\{\mu(\alpha, \beta)\}^{n/2}$. Using the fact that μ is increasing in both of its variables α, β, we have from (14), (7) $\|f\| \geqq 2^{-n}\left\{\mu\left(\frac{n_1}{n}, \frac{n_2}{n}\right)\right\}^{n/2}$, as it was stated. Concerning the footnote **) on page 251 we must consider first the extremal polynomial $\tilde{f}$ which minimizes $\|f\|$. The polynomial $\tilde{f}$ has again the form (1) and has zeros only in $[-1, 1]$, hence $\|f\| \geqq \|\tilde{f}\| \geqq 2^{-n}\{\mu(n_1/n, n_2/n)\}^{n/2}$.

PROOF of Theorem 2. We start with the study of the behaviour of $\mu(\alpha, \beta)$ along a hyperbola $\alpha^2 - \beta^2 = \bar{x}$ ($\bar{x}$ is fixed) in the domain $0 \leqq \alpha, \beta, \alpha+\beta \leqq 1$. Now β can be expressed by α, hence we may write $\beta = \beta(\alpha)$ and we have

(42)
$$\frac{d \log \mu(\alpha, \beta(\alpha))}{d\alpha} = \log \frac{(1+\alpha+\beta)(1+\alpha-\beta)}{(1-\alpha+\beta)(1-\alpha-\beta)} + \beta'(\alpha) \log \frac{(1+\alpha+\beta)(1-\alpha+\beta)}{(1+\alpha-\beta)(1-\alpha-\beta)} > 0,$$

because from $\alpha^2 - \beta^2 = \bar{x}$ it follows $\beta\beta' = \alpha$, therefore $\beta' \geqq 0$.

Since $|\bar{x}| \geqq \alpha_0^2$, by (42) we have $\mu(\alpha, \beta) \geqq \mu(\sqrt{|\bar{x}|}, 0) \geqq \mu(\alpha_0, 0)$, and by the same method as in the proof of Theorem 1 we have $\|f\| \geqq 2^{-n}\{\mu(\alpha_0, 0)\}^{n/2}$, as it was stated.

PROOF of Theorem 3. Now we have to find the minimum of $\mu(\alpha, \beta)$ under the restrictions $\alpha \geqq n_1/n$, $\beta \geqq n_2/n$ and $\alpha^2 - \beta^2 = \pm\alpha_0^2$. Let $n_1 \geqq n_2$. From (42) we conclude, this minimum can be attained only on the lines $\alpha = n_1/n$ and $\beta = n_2/n$. The corresponding points $P_i(\alpha_i, \beta_i)$ of the hyperbolas $\alpha^2 - \beta^2 = \pm\alpha_0^2$ are $P_1\left(n_1/n, \sqrt{(n_1/n)^2 - \alpha_0^2}\right)$ if $n_1^2 - n_2^2 \geqq \alpha_0^2 n^2$ or $P_1\left(\sqrt{\alpha_0^2 + (n_2/n)^2}, n_2/n\right)$ if $n_1^2 - n_2^2 < \alpha_0^2 n^2$ and $P_2\left(n_1/n), \sqrt{\alpha_0^2 + (n_1/n)^2}\right)$ if $1 - (2n_1)/n > \alpha_0^2$. Using the relation $\mu(\beta, \alpha) = \mu(\alpha, \beta)$ we have that the point $P_2(\beta_2, \alpha_2)$ (if P_2 exists at all) is lying on the same hyperbola as P_1, and by (42) $\mu(\alpha_2, \beta_2) = \mu(\beta_2, \alpha_2) \geqq \mu(\alpha_1, \beta_1)$. By the same method as in the proof of the Theorem 1 we obtain the desired inequalities.

REFERENCES

[1] ZOLOTAREFF, G.: Sur l'application des fonctions elliptiques aux questions de maxima et minima, *Bull. Acad. Sci. St. Petersbourg* **24** (1878) 305—310.

[2] ELBERT, Á.: Über eine Vermutung von Erdős betreffs Polynome II. *Studia Scient. Math. Hung.* **3** (1968) 299—324.

[3] PÓLYA, G.: Beitrag zur Verallgemeinerung des Verzerrungsssatzes auf mehrfach zusammenhängende Gebiete, *Sitzungsberichte d. Akad. Wiss. Berlin* (1928) 280—282.

[4] BERNSTEIN, S.: Sur quelques propriétés asymptotiques de polynomes, *Comptes Rendus* **157** (1913) 1055—1057.

[5] WENZL, F.: Nullstellendichte reeller Polynome und Tschebyscheffsche Approximation, *Math. Zeitschr.* **59** (1953) 17—39.

[6] CARLEMAN, T.: Über die Abelsche Integralgleichung mit konstanten Integrationsgrenzen, *Math. Z.* **15** (1122) 111—120.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received October 21, 1968; revised February 20, 1971)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 263—265.

О СУЩЕСТВОВАНИИ ТРЕХЧЛЕННЫХ АРИФМЕТИЧЕСКИХ ПРОГРЕССИЙ В ПЛОТНЫХ ПОДМНОЖЕСТВАХ НАТУРАЛЬНОГО РЯДА

Д. КАЖДАН

Рассматривается следующая задача: пусть M подмножество натуралного ряда, имеющее положительную верхную плотность, то есть $\overline{\lim\limits_{n\to\infty}} \frac{M_n}{n} > 0$ здесь через M_n обозначено число элементов $m \in M$ таких, что $m \leqq n$. Тогда M содержит трехчленные арифметические прогресии. Эти задачи другим методом была решена Ротом.

Обозначим через $M(n)$ подмножество в M состоящее из $m \in M$ таких что $m \leqq n$. Мы докажем, что $\exists$ функция $n(\varepsilon)$ такая, что если $\frac{M_n}{n} \geqq \varepsilon$ и $n > n(\varepsilon)$ то $M(n)$ содержил трехчленные прогрессии. Итак предположим что число α фиксировано. Мы хотим доказать, что если n досточно велико и $M \subset I_n$ где I_n подмножество натуралных чисел элементы которого не больше n, $|M| \geqq \alpha n$, то M содердит трехчленные прогрессии. Если $\alpha > \frac{1}{2}$, то теорема очевидна.

Хотелось бы доказывать эту теорему, показав $\exists \delta > 0$ такое что можно выбрать арифметическую прогрессию $D \subset I_n$ такую, что $|D \cap M| \geqq (\alpha + \delta)|D|$ и $|D| > f(n)$ где $f(n) \underset{n\to\infty}{\to} \infty$. (Очевидно, для этого достаточно построить множество $N \subset I_n$ такое, что $|N \cap M| \geqq (\alpha + \delta)|N|$ и распадается на длиные арифметические прогрессии.) Тогда бы мы теорема сразу доказывалась от противного. К сожалению, такие прогресии могут не существовать однако, можно доказать следующую альтернативу.

Или существует множество N разбывающееся на арифметические прогрессии

$$D, \quad |D| > c\sqrt{n}, \quad |D \cap M| > (\alpha + \delta(\alpha))|D|$$

или в M содержится арифметических прогрессий длины три сколько должно быть из вероятностных соображений (точную формулировку см. ниже.)

На самом деле мы будем считать, что M лежит не на отрезке I_n а на конечный „окружности" Z_p где p простое число и искать арифметические прогрессии в смысле операции на этой группе, то есть такие элементы m_1, m_2, m_3 в $M \subset Z_p$ что $m_2 - m_1 = m_3 - m_2$. Для того, чтобы убедится в эквивалентности этих задач достаточно заметить, что если мы выберем p такое, что $2n < p \leqq 4n$, что всегда возможно и вложить отрезок I_n в Z_p, то числа прогрессий в M рассматриваемого подмножества в I_n и как подмножества Z_p будут совпадать и оценки, которые мы получим будут нетривиальны. Итак пусть M подмножество в Z_p такое, что $|M| \geqq \alpha p$ и p большое простое число. Мы ищем $X(M)$

число решений уравнения $m_2-m_1=m_3-m_2$ или $m_3=2m_2-m_1$. Чтобы записать компактнее это число введем характеристическую функцию $\chi_M(x)$ множества M. Тогда, как легко видеть $X(M)=p^2\cdot\chi_M*\chi'_M*\chi_M$ (0) где $\chi'_M(x)=\chi_M(-\frac{1}{2}x)$ и $*$ обозначает операцию свертки:

$$f_1*f_2(x)=\frac{1}{p}\sum_{x'}f_1(x')f_2(x-x')$$

При изучении свертки функций удобно использовать преобразование Фурье. Напомним, что это такое. На группе Z_p существует характер $\chi_1(x)$ такое, что характеры $\chi_i(x)=\chi_1(ix)$ $(0\leqq i\leqq p-1)$ образуют ортонормированый базис в пространстве функций на Z_p, снабженым скалярным произведением $(f,g)=$ $=\frac{1}{p}\sum_x f(x)\cdot\bar{g}(x)$. Поэтому каждую функцию можно записать в виде $f(x)=$ $=\sum_{i=0}^{p-1}a_i\cdot\chi_i(x)$ где $a_i=(f,\chi_i)$ называются коэффициентами Фурье функции f и $(ff)=\sum_{i=0}^{p-1}|a_i|^2$. При этом, если a_i коэффициенты Фурье функции f, b_i функции g и c_i функции $f*g$ то $c_i=a_ib_i$. Поэтому, если λ_i это коэффициенты Фурье функции χ_M то,

$$\frac{X(M)}{p^2}=\sum_{i=0}^{p-1}\lambda_i^2\lambda_{-2i}$$

Легко видеть, что „нулевой" член этой суммы $\lambda_0^3=\left(\frac{|M|}{p}\right)^3=\alpha^3$ является тем ответом который следовало бы ожидать их вероятностных соображений. Оценим $f*g*h(0)$ в случае когда f, g и h три функции на Z_p коэффициенты которых равны a_i, b_i, c_i соответственно $(f,f)=(g,g)=(h,h)=\alpha$ и $|a_i|=o(1)$ Запишем

$$f*g*h(0)=\sum_{i=0}^{p-1}a_ib_ic_i=a_0b_0c_0+\sum_{i=1}^{p-1}a_ib_ic_i$$

Оценим последнюю сумму

$$\sum_{i=1}^{p-1}a_ib_ic_i$$

Легко доказать оценьку

$$\left|\sum_{i=1}^{p-1}a_ib_ic_i\right|<\alpha\max_{i\neq0}|a_ib_i|^{\frac{1}{2}}.$$

Итак $|f*g*h(0)-a_0c_0b_0|<\alpha\max\limits_{i\neq0}|a_i|^{\frac{1}{2}}\max\limits_{i\neq0}|b_i|^{\frac{1}{2}}$. Применом эту оценку к случаю, когда $f=\chi_M$, $g=\chi'_M$ и $h=\chi_M$. В этом случае $a_0=b_0=c_0=\alpha$. Мы видим что, если $|\lambda_i|<\frac{\alpha^2}{2}$ при $i\neq0$ то $X(M)\geqq p^2\frac{\alpha^3}{2}$. Осталось разобрать случай когда $\exists i\neq0$ такое, что $|\lambda_i|>\frac{\alpha^2}{2}$. Легко показать что $\exists c>0$ такое что множество тех чисел где $\operatorname{Re} \chi_i(x) \cdot \lambda_i^{-1} > c$ обладает свойствами требуеными в альтернативе. Альтернатива доказано.

Возможно, что из доказательства следует существование константа $\gamma > 0$ такое, что если $h > n_0$ и M подмножество I_n такое что $|M| > \gamma n / \log n$ то M содержит трехчленную прогрессию.

ЛИТЕРАТУРА

[1] Roth, K. F.: On certain sets of integers, *J. London Math. Soc.* **28** (1953), 104—109.
[2] Roth, K. F.: On certain sets of integers (II.), *J. London Math. Soc.* **29** (1954), 20—26.

Московский Государственный Университет

(Поступила 6-ого мая 1969 г.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 267—275.

DUAL AND PARAMETRIC METHODS IN DECOMPOSITION FOR LINEAR FRACTIONAL PROGRAM

by

P. ANAND

Abstract: This paper considers a large structure linear fractional program. The decomposition procedure based on dual simplex for optimizing the fractional programming problem is developed. This approach can also be used, for solving certain parametric linear fractional programs in which the parameter is either contained in the requirement vector, or in the costcoefficients of the numerator (or the denominator).

Introducton: Recently much attention has been paid to decomposition algorithms for solving large structure programming problems. DANTZIG and WOLFE [3, 4] developed a decomposition principle for solving large linear programs. The procedure developed has two main characteristics: — (1) The number of contraints in a linear program is reduced at the expense of introducing (in general) a large number of unknowns; then (ii) the simple algorithm is modified by the introduction of a generalized pricing operation" so as to render the new problem amenable to practical solution in spite of the large number of unknowns. Their approach is basically a primal method.

J. M. ABADIE and A. C. WILLIAMS in their paper [1] developed a decomposition algorithm which is basically a dual method. Their method also allows certain parametric linear programs to be solved by decomposition. The algorithm consists of constructing a sequence of admissible vectors such that the sequence of their values is monotonic strictly decreasing. Here the admissible vectors are drawn from a finite set, the convergence is thereby assured.

In this paper, we consider a large structure linear fractional functional program. A decomposition algorithm based on dual simplex method for optimizing a linear fractional program is developed. The dual decomposition algorithm can also be used for solving certain parametric linear fractional functional programs in which the parameter is either contained in the requirement vector or in the coefficients of the numerator (or the denominator). This paper is divided into three sections. Section I provides a dual decomposition procedure for a large structure linear fractional program. In section II, we modify the procedure for $b=0$ and in section III, we treat four parametric linear fractional programming problems to be solved by the decomposition method.

Preliminaries: A general linear fractional programming problem is as follows:

Maximize $$z = \frac{cx}{dx}$$

subject to

$$Ax = a$$
$$x \geqq 0.$$

Several primal methods have been given by various authors for its solution [2, 6], [8, 9]. In a recent paper, KANTI SWARUP [10] developed a dual simplex method for the solution of such problems under the assumption that the denominator of the objective function is possitive for all feasible basic solutions. The approach given resembles with the dual simplex method in a linear program. With well established notations if all $\Delta j \leqq 0$ and some of the basic variables are negative, the rules for making a change of basis are:

(i) The variable to leave the basis is the most negative x_{G_i} call it x_{G_r}

(ii) The variable x_k enters the basis, where k is determined by

$$\frac{\Delta_k}{d_G x_G \bar{u}_{rk}} = \operatorname*{Min}_j \frac{\Delta_j}{d_G x_G \bar{u}_{r_j}} \qquad (\bar{u}_{r_j} < 0)$$

where

$$\Delta_j = d_G x_G (c_j - c_G G^{-1} A_j) - c_G x_G (d_j - d_G G^{-1} A_j)$$

and

$$\bar{u}_{r_j} = \frac{x_{G_r}(d_j - d_G G^{-1} A_j)}{d_G x_G} + u_{r_j}$$

G being the basis matrix and u_{r_j} the rth component of $G^{-1} A_j$. The rth column of G is replaced by the kth column of A. The new basis inverse and a new basic solution with all $\Delta_j \leqq 0$ are computed. The next iteration then commences. The procedure is continued till all the basic variables are nonnegative.

Section I. We modify the dual simplex method described in the Preliminaries to treat a large linear fractional functional program of the form:

$$\text{Problem (1)} \begin{cases} \text{Maximize } z = \dfrac{c'x}{d'x} & (1.1) \\ \text{subject to} & \\ \qquad Ax = a & (1.2) \\ \qquad Bx = b & (1.3) \\ \qquad x \geqq 0 & (1.4) \end{cases}$$

where A is an $m_1 \times n$ matrix, B is an $m_2 \times n$ matrix, a and b are respectively m_1 and m_2 dimensional column vectors. c and d are n dimensional column vectors and x is the n dimensional column vector of unknowns. The denominator of the objective function is assumed to be positive for all feasible basic solutions. It is assumed that $a \neq 0$, $b \neq 0$. The set of points $x \geqq 0$ which satisfy $Bx = b$, is also assumed to be bounded convex set with only a finite number of extreme points. Under this assumption, any $x \geqq 0$ solving $Bx = b$ can be represented by a convex combination of the extreme points of the set of feasible solutions of $Bx = b$, $x \geqq 0$. Let $\varrho^1, \ldots, \varrho^K$ be all the extreme points of the convex set. We can represent any solution x by

$$(1.5) \qquad x = \sum_{k=1}^{K} \lambda_k \varrho^k, \quad \sum_{k=1}^{K} \lambda_k = 1, \qquad \lambda_k \geqq 0 \quad \text{for all} \quad k.$$

If $b=0$, the restriction requiring convex combination of solutions of $x\geqq 0$, $Bx=b$ should be dropped in the development that follows. In that case, any x can be represented as a convex combination of the extreme point solutions and nonnegative combinations of the homogeneous solutions. Thus, in place of (1. 5), we have

(1. 6)
$$x=\sum_{k=1}^{K}\lambda_k\varrho^k+\sum_{q=1}^{Q}\eta^q\mu_q$$

$$\sum_{k=1}^{K}\lambda_k=1,\qquad \lambda_k\geqq 0,\quad \mu_q\geqq 0 \text{ for all } k \text{ and } q.$$

$\eta^1,\dots,\eta^Q$ being a complete set of homogeneous solutions of $Bx=0$, $x\geqq 0$. We shall first consider the case of $b\neq 0$. In this case the linear fractional functional program (1) is equivalent to

$$\text{Problem (2)}\begin{cases}\text{Maximize } z=\dfrac{\sum\limits_{k=1}^{K}(c'\varrho^k)\lambda_k}{\sum\limits_{k=1}^{K}(d'\varrho^k)\lambda_k}\\ \sum\limits_{k=1}^{K}(A\varrho^k)\lambda_k=a\\ \sum\limits_{k=1}^{K}\lambda_k=1,\qquad \lambda_k\geqq 1 \text{ for all } k\end{cases}$$

in the sense that if λ_k^0 is optimal for (2), then $x^0=\sum\limits_{k=1}^{K}\varrho^k\lambda_k^0$ is optimal for (1). Thus we have replaced the linear fractional program (1) by an equivalent linear fractional program (2). Let us write $c'\varrho^k=f_k$; $d'\varrho^k=g_k$ and $A\varrho^k=q_k$ (1. 7). Then (2) becomes

(1. 8)
$$\text{Maximize } z=\frac{\sum\limits_{k=1}^{K}f_k\lambda_k}{\sum\limits_{k=1}^{K}g_k\lambda_k}$$

(1. 9)
$$\sum_{k=1}^{K}q_k\lambda_k=a$$

(1. 10)
$$\sum_{k=1}^{K}\lambda_k=1$$

(1. 11)
$$\lambda_k\geqq 0\quad \text{for all}\quad k.$$

We now solve the problem (2) by dual simplex method. In order to solve the linear fractional program (2) by the dual simplex method, we assume that we have, at each iteration (1), a basic solution (λ) i.e. a basic solution to the constraints (1. 9) (1. 10) [but which may not satisfy (1. 11)] with all $\Delta_j\leqq 0$ and (ii) an inverse matrix i.e. the inverse of the matrix whose columns are the columns corresponding to the various λ_k of the given basic solution.

Let λ_{B_1} be the basic solution with all $\Delta_j \leqq 0$. Let B_1 be the basis matrix of dimension $(m_1+1)\times(m_1+1)$. Let the ith row of the basis inverse be denoted by $(u_i, w_i)=(u_i^1, \dots, u_i^{m_1}, w_i)$. If all the components of the basic solution λ_{B_1} are nonnegative, then the current solution is optimal. If one or more $\lambda_i<0$, then the variable to be deleted from the basis is given by $\lambda_r=\underset{i}{\text{Min}}\ \lambda_i(\lambda_i<0)$ (1. 12) i.e. the Pivot row r has been selected. Now the vector to enter the basis is that vector which minimizes $v=\dfrac{\Delta_j}{\bar{u}_{r_j}}$ over the set of all vectors for which the denominator is negative.

$$\Delta_j = g_{B_1}\lambda_{B_1}[f_j-f_{B_1}(u, w)(q_j, 1)]-f_{B_1}\lambda_{B_1}[g_j-g_{B_1}(u, w)(q_j, 1)] =$$

$$= V_2[f_j-\sigma_1 q_j-\sigma_2]-V_1[g_j-\pi_1 q_j-\pi_2]$$

where

$$V_1=f_{B_1}\lambda_{B_1}, \quad V_2=g_{B_1}\lambda_{B_1}, \quad f_{B_1}u=\sigma_1$$

$$g_{B_1}u=\pi_1, \qquad f_{B_1}w=\sigma_2 \quad \text{and} \quad g_{B_1}w=\pi_2.$$

Here σ_1 and π_1 contain the first m_1 components of $f_{B_1}(u, w)$ and $g_{B_1}(u, w)$ respectively. σ_2 and π_2 are the (m_1+1)th component, of $f_{B_1}(u, w)$ and $g_{B_1}(u, w)$ respectively. Therefore,

$$\Delta_j = V_2(f_j-\sigma_1 g_j)-V_1(g_j-\pi_1 q_j)-V_2\sigma_2+V_1\pi_2$$

$$=[V_2(c-\sigma, A)-V_1(d-\pi_1 A)]\varrho^j-V_2\sigma_2+V_1\pi_2 \qquad \text{by (1. 7)}$$

and

$$\bar{u}_{r_j} = \lambda_r[g_j-g_{B_1}(u, w)(g_j, 1)]+V_2[(u_r, w_r)(q_j, 1)] =$$

$$= \lambda_r(g_j-\pi_1 q_j)+V_2 u_r q_j-\lambda_r\pi_2+V_2 w_r = [\lambda_r(d'-\pi_1 A)+V_2 u_r A]\varrho^j-\lambda_r\pi_2+V_2 w_r.$$

Thus, the vector to enter the basis is determined from

$$\text{(1.13)} \qquad \underset{j}{\text{Minimize}}\ v = \frac{[V_2(c'-\sigma_1 A)-V_1(d'-\pi_1 A)]\varrho^j+V_1\pi_2-V_2\sigma_2}{[\lambda_r(d'-\pi_1 A)+V_2 u_r A]\varrho^j+V_2 w_r-\lambda_r\pi_2}$$

over the set of all vectors for which the denominator is negative.

The problem which we are considering is thus reduced to the problem of finding that ϱ^j from among the ϱ for which the denominator of (1. 13) is negative and for which the ratio (1. 13) is a minimum over such ϱ. We note also that the numerator of this ratio is nonpositive. Therefore the subproblem takes the form: —

$$\text{(1.14)} \qquad \text{Minimize}\ v = \frac{[V_2(c'-\sigma_1 A)-V_1(d'-\pi_1 A)]x+V_1\pi_2-V_2\pi_2}{[\lambda_r(d'-\pi_1 A)+V_2 u_r A]x+V_2 w_r-\pi_2\lambda_r}$$

$$Bx=b$$

$$x\geqq 0$$

which is a linear fractional programming problem and can be solved by several well known methods. An optimal solution of the problem of minimizing (1. 14) subject to the conditions $Bx=b$, $x\geqq 0$ occurs at an extreme point of the convex set of feasible solutions to $Bx=b$ and $x\geqq 0$. Let us assume that the minimum occurs at ϱ^j and the denominator of (1. 13) is negative for ϱ^j. Then we shall introduce λ_j into the basis and delete λ_r from the basis. We compute a new basic solution and a new basis inverse. At each interation we maintain $\Delta_j \leqq 0$. The above procedure is continued till all the basic variables are nonnegative. We also note that if for some ϱ^j an optimal basic solution of (1. 14), we have

$$[\lambda_r(d'-\pi_1 A)+V_2 u_r A]\varrho^j + V_2 w_r - \pi_2 \lambda_r \geqq 0$$

then there is no admissible basic solution. Therefore the problem (1) has no optimal solution. Thus, we find that the above procedure is repeated to determine whether the new basic solution is optimal, or if not, what variable enters the basis of the next iteration.

Section II. *Modification* for $b=0$.

We shall now modify the above procedure for $b=0$.

In this case, we substitute (1. 6) into the linear fractional program (1) to replace it by an equivalent linear fractional program of the form:

$$\text{Problem (3)}\begin{cases} \text{Maximize } z = \dfrac{\sum_{k=1}^{K}(c'\varrho^k)\lambda_k + \sum_{q=1}^{Q}(c'\eta^q)\mu_q}{\sum_{k=1}^{K}(d'\varrho^k)\lambda_k + \sum_{q=1}^{Q}(d'\eta^q)\mu_q} & (2.\ 1)\\ \sum_{k=1}^{K}(A\varrho^k)\lambda_k + \sum_{q=1}^{Q}(A\eta^q)\mu_q = a & (2.\ 2)\\ \sum_{k=1}^{K}\lambda_k = 1 & (2.\ 3)\\ \lambda_k,\quad \mu^q \geqq 0 & (2.\ 4)\end{cases}$$

Let $(\lambda_{B_1}\ \mu_{B_1})$ be the basic solution with B_1 as a basic matrix. We assume that for each basic solution,

$$[V_2(c'-\sigma_1 A)-V_1(d'-\pi_1 A)]\varrho^k + V_1\pi_2 - V_2\sigma_2 \leqq 0 \qquad k=1,\dots,K$$

and for each homogeneous solution,

$$[V_2(c'-\sigma_1 A)-V_1(d'-\pi_1 A)]\eta^q \leqq 0 \qquad q=1,\dots,Q.$$

The pivot row will be determined from (1. 12). The vector to enter the basis for the columns of the type $A\varrho^k$ is determined by (1. 13), and the ratio (1. 13) for the columns of the type $A\eta^q$ is given by

$$\frac{[V_2(c'-\sigma_1 A)-V_1(d'-\pi_1 A)]\eta^q}{[V_2 u_r A+\lambda_r(d'-\pi_1 A)]\eta^q} \tag{2. 5}$$

i.e. we have to solve the following nonlinear program:

$$\text{Minimize } v = \frac{[V_2(c'-\sigma_1 A)-V_1(d'-\pi_1 A)]x}{[\lambda_r(d'-\pi_1 A)+V_2 u_r A]x}$$

$$Bx=0$$

$$x \geqq 0.$$

In this nonlinear program, the numerator is nonpositive and we have to minimize it over all vectors for which the denominator is negative.

Thus, in this case our problem is to find that ϱ^j or η^q from among the ϱ and η for which the denominators of (1. 13) and (2. 5) are negative and for which the ratios (1. 13) or (2. 5) is a minimum over such ϱ and η. The determination of the correct vector ϱ or η according to the above criterion is the "subproblem" which has to be solved at each iteration. This subproblem can be solved by several well know-methods.

The dual decomposition algorithm thus requires that on each step either we obtain an extreme point solution or we obtain a homogeneous solution. This algorithm has the same termination properties as in the case of $b\neq 0$, since the set of possible homogeneous solutions generated in this manner is finite.

Section III. *Parametric Linear Fractional Functional Programming*. Here, four types of parametric linear fractional functional programming problems have been considered which are to be solved by the decomposition method. We assume that the denominator of the objective function of the parametric linear fractional programming problems considered in this section to be positive for all feasible basic solutions. The parameter can take only nonnegative values.

Part I: — We find the optimal solution $x^0(\vartheta)$ a function of the parameter ϑ for the parametric linear fractional program

$$\text{Maximize } z = \frac{c'x}{d'x}$$

$$Ax = a + \vartheta\bar{a}$$

$$Bx = b$$

$$x \geqq 0.$$

Again, we consider the linear fractional program (3), where in (2. 2), we replace a by $a+\vartheta\bar{a}$. We assume that an optimal basic solutions (λ^0, μ^0) for $\vartheta=\underline{\vartheta}$ (initially $\underline{\vartheta}=0$) has been found, and we have also obtained a basis inverse. We now wish to compute an optimal solution for all $\vartheta>\underline{\vartheta}$ for which such solution exists.

Let us consider the problem

$$\text{Maximize } z = \frac{c'x}{d'x}$$

$$Ax = a + \vartheta\bar{a}$$

$$x \geqq 0$$

Suppose we have a basic optimal solution x^0 for $\vartheta=\underline{\vartheta}=0$. Let U be the inverse basis matrix. Then $x^0 = Ua+\vartheta U\bar{a} = \alpha+\vartheta\bar{\alpha}$ (say). If all $\bar{\alpha}_i\geqq 0$, ϑ can be increased without limit, and x^0 will be an optimal solution. If one or more $\bar{\alpha}_i<0$, we define $\bar{\vartheta} = \dfrac{\alpha_r}{-\bar{\alpha}_r} = \min\limits_{\bar{\alpha}_i<0} \dfrac{\alpha_i}{-\bar{\alpha}_i}$. Then $x^0(\vartheta) = \alpha_i+\vartheta\bar{\alpha}$ for $\underline{\vartheta}\leqq\vartheta\leqq\bar{\vartheta}$. If $\vartheta>\bar{\vartheta}$, the basis must be changed in order to maintain feasibility. When ϑ becomes slightly greater than $\bar{\vartheta}$, we have a basic nonfeasible solution with all $\Delta_j\leqq 0$. Then by the dual simplex method, we can determine the vector to be introduced into the basis by minimizing the ratio $\dfrac{\Delta_j}{x_r(d-\pi_1 a_j)+V_2 u_r a_j}$ over all vectors for which the denominator is negative.

Clearly, the parametric linear fractional programming problem by decomposition i.e. for a program of the type (3), is similarly reduced to the problem of selecting a vector for which the ratio (1. 13) or (2. 5) is a minimum, subject to the constraint that the denominator be negative.

Part II: — The second parametric linear fractional programming problem is

$$\text{Maximize } z = \frac{(c'+\vartheta\bar{c}')x}{d'x}$$

$$Ax=a$$

$$Bx=b$$

$$x\geqq 0$$

We reformulate it again in the form of the problem (3). We assume that an optimal solution (λ^0, ω^0) is achieved for $\vartheta=\underline{\vartheta}=0$ and also the inverse basis matrix.

Let
$$\bar{V}_1 = \sum_{k=1}^{K} (\bar{c}'_{B_1}\varrho^k)\lambda_k^0 + \sum_{q=1}^{Q} (\bar{c}'_{B_1}\eta^q)\mu_q^0$$

and
$$(\bar{\sigma}_1, \bar{\sigma}_2) = \bar{f}_{B_1}(u, w)$$

Now the vector to enter the basis (so as to compute $x^0(\vartheta)$ for $\vartheta>0$) is that vector for which

$$\frac{[V_2(c'-\sigma_1 A)-V_1(d'-\pi_1 A)]\varrho^j+V_1\pi_2-V_2\sigma_2}{[V_2(\bar{c}'-\bar{\sigma}_1 A)-\bar{V}_1(d'-\pi_1 A)]\varrho^j+\bar{V}_1\pi_2-V_2\bar{\sigma}_2}$$

or

$$\frac{[V_2(c'-\sigma_1 A)-V_1(d'-\pi_1 A)]\eta^q}{[V_2(\bar{c}'-\bar{\sigma}_1 A)-\bar{V}_1(d'-\pi_1 A)]\eta^q}$$

is a minimum over the set of all such vectors for which the denominator is negative (if the denominator is nonnegative for all ϱ^j and all η^q, then the current solution is optimal for all $\vartheta\geqq\underline{\vartheta}$). The problem is again reduced to the previous ones.

Part III:

The parametric linear fractional functional programming problem is as:

$$\text{Maximize } z = \frac{c'x}{(d'+\vartheta\bar{d}')x}$$

$$Ax=a$$

$$Bx=b$$

$$x\geqq 0.$$

Reformulate it again in the form of the problem (3). Let (λ^0, μ^0) be an optimal solution with (u, w) an inverse basis matrix for $\vartheta = \underline{\vartheta} = 0$

Let
$$\bar{V}_2 = \sum_{k=1}^{K} (\bar{d}'_{B_1} \varrho^k) \lambda_k^0 + \sum_{q=1}^{Q} (\bar{d}'_{B_1} \eta^q) \mu_q^0$$

and
$$(\bar{\pi}_1, \bar{\pi}_2) = \bar{g}_{B_1}(u, w)$$

The vector to be introduced into the basis (so as to compute $x^0(\vartheta)$ for $\vartheta > 0$) is that vector for which

$$\frac{[V_2(c' - \sigma_1 A) - V_1(d' - \pi_1 A)]\varrho^j + V_1 \pi_2 - V_2 \sigma_2}{[\bar{V}_2(c' - \sigma_1 A) - V_1(\bar{d}' - \bar{\pi}_1 A)]\varrho^j + V_1 \bar{\pi}_2 - \bar{V}_2 \sigma_2}$$

or

$$\frac{[V_2(c' - \sigma_1 A) - V_1(d' - \pi_1 A)]\eta^q}{[\bar{V}_2(c' - \sigma_1 A) - V_1(\bar{d}' - \bar{\pi}_1 A)]\eta^q}$$

is a minimum over the set of all such vectors for which the denominator is negative.

Part IV:

Lastly, we consider the following parametric fractional programming problem

$$\text{Maximize } z = \frac{(c' + \vartheta)x}{(d' + \vartheta)x}$$

$$Ax = a$$

$$Bx = b$$

$$x \geqq 0$$

We reformulate the problem in the form (3). We solve this problem for $\vartheta = \underline{\vartheta} = 0$. Let (λ^0, μ^0) be an optimal basic solution. Let (u, w) be the inverse basis matrix of dimension $(m_1 + 1) \times (m_1 + 1)$. We now wish to compute an optimal solution for all $\vartheta > \underline{\vartheta}$ for which such solutions exist. In this case a change of basis is required. The vector to be inserted into the basis is that vector for which

$$\frac{(V_2(c' - \sigma_1 A) - V_1(d' - \pi_1 A)]\varrho^j + V_1 \pi_2 - V_2 \sigma_2}{[\bar{V}(c' - \sigma_1 A) - \bar{V}(d' - \pi_1 A) + uA(V_1 - V_2)]\varrho^j + \bar{V}\pi_2 - \bar{V}\sigma_2 + V_2 - V_1 + w(V_1 - V_2)}$$

or

$$\frac{[V_2(c' - \sigma_1 A) - V_1(d' - \pi_1 A)]\eta^q}{[\bar{V}(c' - \sigma_1 A) - \bar{V}(d' - \bar{\pi}_1 A) + uA(V_1 - V_2)]\eta^q}$$

is the minimum over those vectors for which the denominator is negative.

Here
$$\bar{V} = \sum_{k=1}^{K} \varrho^k \lambda_k^0 + \sum_{q=1}^{Q} \eta^q \mu_q^0$$

(If the denominator is nonnegative for all ϱ^j and η^q then the current solution is optimal for all $\vartheta \geqq \underline{\vartheta}$). The problem is again reduced to the previous ones.

We remark that the decomposition procedure can be applied to parametric linear fractional functional programming problems in which the parameter can take any value not necessarly nonnegative, provided the denominator of the objective function remains positive for all feasible solutions and all basic solutions.

Acknowledgement: I would like to acknowledge gratefully the valuable help received from Dr. KANTI SWARUP and Dr. S. P. AGGARWAL Faculty of Mathematics, University of Delhi.

REFERENCES

[1] ABADIE, J. M. and WILLIAMS, A. C.: *Dual and parametric Methods in Decomposition,* Recent Advances in Mathematical Programming Edited by P. Wolfe and R. Grove, McGraw—Hill, London, 1966.
[2] CHARNES, A. and COOPER, W. W.: Programming with linear fractional functionals, *Nov. Res. Log. Quart.* **9**, 1962.
]3] DANTZIG, G. B. and P. WOLFE: A Decomposition Principle for Linear Programs, *Op. Res.,* **8**, 1960.
[4] The Decomposition Algorithm for Linear Programs *Econometrica,* **29**, No. 4, 1961.
[5] GASS, S. I.: *Linear Programming,* McGraw—Hill, 1958.
[6] ISBELL, J. R. and MARLOW, W. H.: Attrition Games, *Nav. Res. Log. Quart.,* **3**, 1956.
[7] LEMKE, C. E.: The dual method for solving the linear programming problem, *Nav. Res. Log. Quart.,* **1**, 1954.
[8] MARTOS, B.: Hyperbolic Programming, *Nav. Res. Log. Quart.,* **11**, 1964.
[9] SWARUP, K.: Linear fractional functionals Programming, *Op. Res.* **83**, No. 6, 1965.
[10] Some Acpests of Linear Fractional Functionals Programming, *The Austrelich Journal of Statistics,* **7**, No. 3, 1965.

Faculty of Mathematics, University of Delhi
(Residential adress: B — 4/28 Lodi Colony New Delhi — 3 INDIA)

(Received May 26, 1969.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 277—285.

SOME INEQUALITIES CONCERNING BESSEL FUNCTIONS OF FIRST KIND

by

Á. ELBERT

I.

If γ_ν is the first positive root of the Bessel function of order ν of the first kind then the well-known estimations (see WATSON [1])

$$\sqrt{\nu(\nu+2)} < \gamma_\nu < \sqrt{2(\nu+1)(\nu+3)} \qquad \nu \geqq 0 \tag{1}$$

hold. Now we want to obtain similar inequalities for γ_ν in the case $-1<\nu<0$, too. We shall prove the following inequalities:

$$\sqrt{(\nu+1)(\nu+5)} < \gamma_\nu < \sqrt{2(\nu+1)(\nu+3)} \qquad -1<\nu<0 \tag{2}$$

$$\{\gamma_\nu - \pi(1+\nu)\}\operatorname{sg}\left(\nu+\frac{1}{2}\right) < 0 \qquad -1<\nu<\infty,\ \nu \neq -\frac{1}{2} \tag{3}$$

$$\gamma_\nu \leqq \frac{\pi}{2} + A\left(\nu+\frac{1}{2}\right) \qquad -1<\nu<\infty, \tag{4}$$

where

$$A = \int_0^\pi \frac{\sin\varphi}{\varphi}\,d\varphi = 1{,}8519\ldots.$$

Proof of these inequalities. Let us consider the solution $Y_\lambda(x)$ of the differential equation

$$y'' + x^\lambda y = 0 \qquad (-2<\lambda<\infty) \tag{5}$$

with the initial conditions $Y_\lambda(0)=0$, $Y'_\lambda(0)=1$, then

$$Y_\lambda(x) = c_\lambda \sqrt{x}\, J_{1/(\lambda+2)}(x^{1+\lambda/2}/(1+\lambda/2))$$

and

$$Y'_\lambda(x) = c_\lambda x^{(\lambda+1)/2} J_{-1+1/(\lambda+2)}(x^{1+\lambda/2}/(1+\lambda/2)),$$

where c_λ is some constant. Let x_0 denote the first positive root of $Y'_\lambda(x)=0$, then

$$\gamma_{-1+\frac{1}{\lambda+2}} = \frac{x_0^{1+\frac{\lambda}{2}}}{1+\frac{\lambda}{2}}. \tag{6}$$

By (6) the inequalities (2) are equivalent to

$$\lambda+\frac{9}{4} < x_0^{\lambda+2} < \lambda+\frac{5}{2} \qquad -1<\lambda<\infty. \tag{7}$$

By transforming the differential equation (5) of $Y_\lambda(x)$ into the integral equation

$$Y_\lambda(x) = x - \int_0^x (x-t)t^\lambda Y_\lambda(t)\,dt = x - KY_\lambda(x), \tag{8}$$

we obtain

$$Y_\lambda'(x) = 1 - \int_0^x t^\lambda Y_\lambda(t)\,dt. \tag{9}$$

From (8) we have by induction on n

$$Y_\lambda(x) = \sum_{i=0}^{n} (-1)^i K^i x + (-1)^{n+1} K^{n+1} Y_\lambda(x) \qquad (K^0 x \equiv x).$$

Since $Y_\lambda'(x) > 0$ for $0 < x < x_0$, therefore $Y_\lambda(x) > 0$ for $0 < x \leqq x_0$ and $K^{n+1} Y_\lambda(x) > 0$ for $0 < x \leqq x_0$ and $n \geqq 0$ hence by (9)

$$1 - \int_0^x t^\lambda \sum_{i=0}^{2} (-1)^i K^i t\,dt \overset{\text{def}}{=} p(x^{\lambda+2}) < Y_\lambda'(x) < 1 - \int_0^x t^\lambda \sum_{i=0}^{3} (-1)^i K^i t\,dt \overset{\text{def}}{=} \tag{10}$$

$$\overset{\text{def}}{=} p^*(x^{\lambda+2}) \qquad (0 < x \leqq x_0).$$

By a simple computation we obtain

$$K^i t = \frac{t^{1+i(\lambda+2)}}{(\lambda+2)(\lambda+3)(2\lambda+4)(2\lambda+5)\dots(i\lambda+2i)(i\lambda+2i+1)} \qquad (i \geqq 1),$$

hence

$$p(z) = 1 - \frac{z}{\lambda+2} + \frac{z^2}{2(\lambda+2)^2(\lambda+3)} - \frac{z^3}{6(\lambda+2)^3(\lambda+3)(2\lambda+5)}.$$

If we take into account the inequality

$$\min_{0 \leqq z \leqq \lambda+9/4} \left(1 - \frac{z}{\lambda+2} + \frac{z^2}{2(\lambda+2)^2(\lambda+3)}\right) \geqq \frac{1}{4(\lambda+3)} + \frac{1}{32(\lambda+2)^2(\lambda+3)},$$

we get

$$\min_{0 \leqq z \leqq \lambda+9/4} p(z) > \frac{1}{4(\lambda+3)} - \frac{z^3}{6(\lambda+2)^3(\lambda+3)(2\lambda+5)} \geqq$$

$$\geqq \frac{1}{12(\lambda+3)} \min_{-1 \leqq \lambda < \infty} \left[3 - \frac{2}{2\lambda+5}\left(\frac{\lambda+\frac{9}{4}}{\lambda+2}\right)^3\right] > 0,$$

hence by (10) the left hand side inequality of (7) is true.

Similarly we have for $p^*(z)$

$$p^*(z) = 1 - \frac{z}{\lambda+2} + \frac{z^2}{2(\lambda+2)^2(\lambda+3)} - \frac{z^3}{6(\lambda+2)^3(\lambda+3)(2\lambda+5)} +$$

$$+ \frac{z^4}{24(\lambda+2)^4(\lambda+3)(2\lambda+5)(3\lambda+7)}.$$

The right inequality of (7) will be proved if we show $p^*(\lambda+5/2)<0$. Indeed,

$$96(\lambda+2)^4(\lambda+3)(2\lambda+5)(3\lambda+7)p^*(\lambda+5/2) = (\lambda+5/2)q(\lambda+5/2),$$

where

$$q(z) = -48z^4+108z^3-88z^2+30z-3,$$

but

$$q(z) = -1-14(z-1)-52(z-1)^2-48(z-1)^3(z+3/4),$$

hence $q(z)<0$ for $z\geqq 1$, i.e. $p(\lambda+5/2)<0$ if $\lambda\geqq -3/2$, which was to be proved.

We remark here, that we proved now somewhat more as we stated in (2) or in (7), respectively, namely that the right hand sides of these inequalities are true by (6) for $0\leqq\nu\leqq 1$, too, as in (1). But by a finer estimation for $p(z)$ the left inequality of (7) can be proved for $0\leqq\nu\leqq 1$, too. It is very possible, that this inequality holds for all $\nu>-1$.

To prove inequalities (3) and (4) we introduce the function $\varphi(x)$ by

$$\text{(11)}\qquad \operatorname{tg}\varphi(x) = \frac{x^{\frac{\lambda}{2}}Y_\lambda(x)}{Y_\lambda'(x)}, \qquad \varphi(0)=0,$$

which satisfies the differential equation

$$\text{(12)}\qquad \varphi' = x^{\frac{\lambda}{2}}+\frac{\lambda}{4x}\sin 2\varphi.$$

From (11) it is obvious that $\varphi(x_0)=\pi/2$. Taking into account the initial values of $Y_\lambda(x)$ we have by (11)

$$\text{(13)}\qquad \lim_{x\to+0}\varphi(x)x^{-\left(1+\frac{\lambda}{2}\right)} = \lim_{x\to+0}\frac{\varphi(x)}{\operatorname{tg}\varphi(x)}\frac{Y_\lambda(x)}{xY_\lambda'(x)} = 1,$$

and using the inequality $\sin x<x$ for $x>0$, we obtain from (12)

$$\lambda(\varphi'-x^{\frac{\lambda}{2}}) < \frac{\lambda^2}{2x}\varphi \qquad (\lambda\neq 0),$$

therefore

$$\lambda(\varphi x^{-\frac{\lambda}{2}})' < \lambda,$$

hence by (13)

$$\lambda\varphi(x) < \lambda x^{1+\frac{\lambda}{2}},$$

and especially at $x=x_0$

$$\lambda\left[\frac{\pi}{2}-\left(1+\frac{\lambda}{2}\right)\gamma_{-1+\frac{1}{\lambda+2}}\right]<0,$$

or

$$\left(\nu+\frac{1}{2}\right)[\pi(1+\nu)-\gamma_\nu]>0,$$

which proves the inequality (3). The case $\nu=-\frac{1}{2}$ is of no interest, while $\gamma_{-1/2}=\pi/2$.

For proving inequality (4) first we have to prove the following inequalities:

$$\varphi'(x) > 0 \qquad x > 0 \tag{14}$$

$$\lambda\left[\left(1+\frac{\lambda}{2}\right)\varphi - x\varphi'\right] > 0 \quad \text{if} \quad \lambda\neq 0,\ \lambda > -2, \text{ for } 0 < x < x_0. \tag{15}$$

Using the identity $\sin 2\varphi = 2\,\mathrm{tg}\,\varphi/(1+\mathrm{tg}^2\,\varphi)$ we have by (11) and (12)

$$x^{1-\frac{\lambda}{2}}(Y_\lambda'^2 + x^\lambda Y_\lambda^2)\varphi' = xY_\lambda'^2 + x^{\lambda+1}Y_\lambda^2 + \frac{\lambda}{2}Y_\lambda Y_\lambda' \stackrel{\text{def}}{=} g(x),$$

where $g(0)=0$, and by (5)

$$g'(x) = \left(1+\frac{\lambda}{2}\right)(Y_\lambda'^2 + x^\lambda Y_\lambda^2) > 0 \quad \text{for} \quad x > 0,$$

therefore $g(x) > 0$ for $x > 0$, hence (14) is true.

Concerning the inequality (15) we can point out that in the case $\lambda = 0$ we have $x\varphi' - \varphi \equiv 0$ for all values of x. By (13) we can write $\varphi(x) = O(x^{1+\lambda/2})$ for sufficiently small values of x, therefore from (12)

$$\varphi'(x) = x^{\frac{\lambda}{2}} + \frac{\lambda}{2x}\varphi(x) + O(x^{\frac{3}{2}\lambda+2}),$$

$$\varphi(x) = x^{\frac{\lambda}{2}}\int_0^x [\varphi(\xi)\xi^{-\frac{\lambda}{2}}]'\,d\xi = x^{1+\frac{\lambda}{2}} + O\left(x^{3\left(1+\frac{\lambda}{2}\right)}\right).$$

Repeating this procedure we have

$$\varphi'(x) = x^{\frac{\lambda}{2}} + \frac{\lambda}{2x}\varphi - \frac{\lambda}{3x}\varphi^3 + O(x^{\frac{5}{2}\lambda+4}),$$

$$\varphi(x) = x^{1+\frac{\lambda}{2}} - \frac{\lambda}{3(\lambda+3)}x^{3\left(1+\frac{\lambda}{2}\right)} + O\left(x^{5\left(1+\frac{\lambda}{2}\right)}\right),$$

therefore

$$x\varphi' = \left(1+\frac{\lambda}{2}\right)x^{1+\frac{\lambda}{2}} - \frac{\lambda(\lambda+2)}{2(\lambda+3)}x^{3\left(1+\frac{\lambda}{2}\right)} + O\left(x^{5\left(1+\frac{\lambda}{2}\right)}\right).$$

Using the notation

$$\psi(x) = \frac{1}{\lambda}\left[\left(1+\frac{\lambda}{2}\right)\varphi - x\varphi'\right] \qquad \lambda\neq 0,$$

$\psi(x)$ can be written in the form

$$\psi(x) = \frac{\lambda+2}{6(\lambda+3)}x^{3\left(1+\frac{\lambda}{2}\right)} + O\left(x^{5\left(1+\frac{\lambda}{2}\right)}\right).$$

Consequently the statement (15) is true if x is small enough. Let a be defined by $a = \inf\{x;\ \psi(x)\leqq 0,\ 0 < x \leqq x_0\}$. It is clear that $a > 0$. We show that the case $0 < a < x_0$

would be a contradiction. Indeed, in this case it would be $\psi(a)=0$, $0<\varphi(a)<\pi/2$, $0<\psi(x)$ for $0<x<a$, hence

$$\psi'(a) = \lim_{x\to a-0} \frac{\psi(a)-\psi(x)}{a-x} \leq 0. \tag{16}$$

On the other hand, the function $\psi(x)$ fulfils the differential equation

$$\psi' = \frac{\cos 2\varphi}{2x}\psi + \frac{2+\lambda}{8x}[\sin 2\varphi - 2\varphi\cdot\cos 2\varphi],$$

therefore $\psi(a)>0$, because the expression in the brackets is positive for $0<\varphi<\pi/2$, which contradicts (16), i.e. $a=x_0$, and the relation (15) is true.

Integrating the differential equation (12) over $[0, x_0]$ we get by (6)

$$\frac{\pi}{2} = \gamma_{\frac{1}{\lambda+2}-1} + \frac{\lambda}{4}\int_0^{x_0} \frac{\sin 2\varphi(x)}{x}\,dx, \tag{17}$$

hence for $\lambda=0$, i.e. $\nu=-1/2$ the relation (4) is true. For $\nu=-1+1/(\lambda+2)$, $\lambda>-2$, $\lambda\neq 0$ we have by (6), (14), (15), (17)

$$\frac{\pi}{2}-\gamma_{-1+\frac{1}{\lambda+2}} - \frac{\lambda}{2(\lambda+2)}\int_0^{\pi}\frac{\sin\varphi}{\varphi}\,d\varphi =$$

$$= \frac{\lambda}{4}\int_0^{x_0}\frac{\sin 2\varphi(x)}{x}\,dx - \frac{\lambda}{2(\lambda+2)}\int_0^{x_0}\frac{\sin 2\varphi(x)}{\varphi(x)}\varphi'(x)\,dx =$$

$$= \frac{\lambda}{2(\lambda+2)}\int_0^{x_0}\frac{\left(1+\frac{\lambda}{2}\right)\varphi - x\varphi'}{x\varphi'\varphi}\sin 2\varphi\cdot\varphi'\,dx > 0,$$

which proves the inequality (4).

II.

Let us consider now the solution $Z_\lambda(x)$ of the differential equation (5) with the initial conditions $Z_\lambda(0)=1$, $Z'_\lambda(0)=0$, where $0<\lambda<\infty$, and denote by $x'_0, x'_1, x'_2, \ldots$ the consecutive roots of the equation $Z'_\lambda(x)=0$. The values

$$\varrho_n^{(\lambda)} = Z_\lambda^2(x'_n) \qquad (n=1,2,\ldots)$$

have an interesting meaning [2], and it was given the estimation

$$\varrho_n^{(\lambda)} > \alpha_\lambda n^{-\frac{\lambda}{\lambda+2}} \qquad 0<\alpha_\lambda<1,\ \lambda>0.$$

Now we shall prove the following relations:

$$(18)\qquad 0{,}617 < \sqrt{\alpha_\lambda}\cdot\frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1+\lambda}{2+\lambda}\right)}\left(\frac{2}{\pi}\right)^{-\frac{\lambda}{2(\lambda+2)}} < 1,$$

$$(19)\qquad \lim_{n\to\infty}|Z_\lambda(x_n')|\,n^{\frac{\lambda}{2(\lambda+2)}} = \frac{\Gamma\left(\frac{1+\lambda}{2+\lambda}\right)}{\Gamma\left(\frac{1}{2}\right)}\left(\frac{2}{\pi}\right)^{\frac{\lambda}{2(\lambda+2)}},$$

where $\Gamma(x)$ is the gamma function.

Let $\mu = 1/(\lambda+2)$ and $z = x^{1+\lambda/2}/(1+\lambda/2)$, then the solution $Z_\lambda(x)$ can be expressed in the form

$$Z_\lambda(x) = \mu^\mu \Gamma(1-\mu)\sqrt{x}\,J_{-\mu}(z),$$

or using the formulae due to Schafheitlin (see e.g. [1])

$$(20)\qquad Z_\lambda(x) = \frac{\Gamma(1-\mu)2^{-2\mu+1}}{\Gamma\left(\frac{1}{2}-\mu\right)\Gamma\left(\frac{1}{2}\right)}\int_0^{\frac{\pi}{2}}\frac{\sin\left[z+\theta\left(\frac{1}{2}+\mu\right)\right]}{\cos^{\frac{1}{2}+\mu}\theta\cdot\sin^{1-2\mu}\theta}\,e^{-2z\operatorname{ctg}\theta}\,d\theta.$$

From this we have

$$|Z_\lambda(x)|\cdot\left\{\frac{\Gamma(1-\mu)2^{-2\mu+1}}{\Gamma\left(\frac{1}{2}-\mu\right)\Gamma\left(\frac{1}{2}\right)}\right\}^{-1} < \int_0^{\frac{\pi}{2}}\frac{e^{-2z\operatorname{ctg}\theta}}{\cos^{\frac{1}{2}+\mu}\theta\cdot\sin^{1-2\mu}\theta}\,d\theta =$$

$$= \int_0^{\frac{\pi}{2}}(\operatorname{tg}\theta)^{\frac{1}{2}+\mu}(\sin\theta)^{\frac{1}{2}+\mu}e^{-2z\operatorname{ctg}\theta}\frac{d\theta}{\sin^2\theta} < \int_0^\infty u^{-\frac{1}{2}-\mu}e^{-2zu}\,du = (2z)^{\mu-\frac{1}{2}}\cdot\Gamma\left(\frac{1}{2}-\mu\right)$$

therefore

$$(21)\qquad |Z_\lambda(x)| < \frac{\Gamma(1-\mu)2^{\frac{1}{2}-\mu}}{\Gamma\left(\frac{1}{2}\right)}z^{\mu-\frac{1}{2}}.$$

Now we need upper and lower estimations for the values of x_n'-s. Let z_n be the nth positive zero of the function $J_{-\mu}(z)$ and x_n the corresponding zero of $z_\lambda(x)$ $(n=1, 2, \ldots)$. Let $\chi=\chi_\lambda(x)$ be defined by

$$(22)\qquad \operatorname{tg}\chi = \frac{x^{\frac{\lambda}{2}}Z_\lambda(x)}{Z_\lambda'(x)},\qquad \chi(0) = \frac{\pi}{2},$$

then $\chi_\lambda(x)$ fulfils the differential equation

$$\chi' = x^{\frac{\lambda}{2}} + \frac{\lambda}{4x}\sin 2\chi \tag{23}$$

and we have $\chi(x_n') = \left(n+\frac{1}{2}\right)\pi$ $(n=1, 2, \ldots)$, $\chi(x_n) = n\pi$ $(n=1, 2, \ldots)$.

It was proved in [2] that the sequence $\left\{\int\limits_0^{x_n'} \frac{\lambda}{4x}\sin 2\chi_\lambda\, dx\right\}_{n=0}^{\infty}$ is decreasing, while the sequence $\left\{\int\limits_0^{x_n} \frac{\lambda}{4x}\sin 2\chi_\lambda\, dx\right\}_{n=1}^{\infty}$ is increasing. It can be easily seen that they have a common limit, therefore by (23) this limit is

$$\lim_{n\to\infty} \int\limits_0^{x_n'} \frac{\lambda}{4x}\sin 2\chi_\lambda\, dx = \lim_{n\to\infty} \int\limits_0^{x_n} \frac{\lambda}{4x}\sin 2\chi_\lambda\, dx = \lim_{n\to\infty}\left[\chi_\lambda(x_n) - \frac{\pi}{2} - \frac{x_n^{1+\frac{\lambda}{2}}}{1+\frac{\lambda}{2}}\right] =$$

$$= \lim_{n\to\infty}\left[\left(n-\frac{1}{2}\right)\pi - z_n\right] = \lim_{n\to\infty}\left[\left(n-\frac{1}{2}\right)\pi - n\pi + \frac{\pi}{2}\left(\frac{1}{2}+\mu\right)\right] = -\frac{\pi}{2}\left(\frac{1}{2}-\mu\right),$$

where $z_n = x_n^{1+\lambda/2}/(1+\lambda/2)$ is the nth zero of the Bessel function $J_{-1/(\lambda+2)}(z)$, having the asymptotic value (see WATSON [1]) $n\pi - \frac{\pi}{2}\left(\frac{1}{2}+\frac{1}{2+\lambda}\right)$. It follows from this immediately that

$$0 > \int\limits_0^{x_n'} \frac{\lambda}{4x}\sin 2\chi_\lambda\, dx > -\frac{\pi}{2}\left(\frac{1}{2}-\mu\right) > \int\limits_0^{x_n} \frac{\lambda}{4x}\sin 2\chi_\lambda\, dx \qquad (n=1, 2, \ldots),$$

therefore

$$n\pi \leqq z_n' = \int\limits_0^{x_n'} \left(\chi_\lambda' - \frac{\lambda}{4x}\sin 2\chi_\lambda\right) dx < n\pi + \frac{\pi}{2}\left(\frac{1}{2}-\mu\right) \qquad (n=0, 1, \ldots)$$

$$\left(n-\frac{1}{2}\right)\pi \leqq z_n < n\pi - \frac{\pi}{2}\left(\frac{1}{2}-\mu\right) \qquad (n=1, 2, \ldots).$$

Let $\zeta_n = n\pi + \frac{\pi}{2}\left(\frac{1}{2}-\mu\right)$, and ξ_n be defined by $\zeta_n = \xi_n^{1+\lambda/2}/(1+\lambda/2)$. It is clear that

$$z_n' < \zeta_n < z_{n+1}. \tag{24}$$

By (21) and the definition of ζ_n we have

$$|z_\lambda(x_n')| < \frac{\Gamma(1-\mu)}{\Gamma\left(\frac{1}{2}\right)}\left(\frac{2}{\pi}\right)^{\frac{1}{2}-\mu} n^{\mu-\frac{1}{2}}, \tag{25}$$

which is the right hand side of (18).

By (24) we have

$$(-1)^n Z_\lambda(x_n') > (-1)^n Z_\lambda(\xi_n),$$

hence by (20) and the substitution $\theta = \pi/2 - \vartheta$

$$(-1)^n Z_\lambda(x_n') \cdot \left\{ \frac{\Gamma(1-\mu)}{\Gamma\left(\frac{1}{2}-\mu\right)\Gamma\left(\frac{1}{2}\right)} 2^{1-2\mu} \right\}^{-1} >$$

$$> (-1)^n \int_0^{\frac{\pi}{2}} \frac{\sin\left[\zeta_n + \theta\left(\frac{1}{2}+\mu\right)\right]}{\cos^{\frac{1}{2}+\mu}\theta \cdot \sin^{1-2\mu}\theta} e^{-2\zeta_n \operatorname{ctg}\theta}\, d\theta = \int_0^{\frac{\pi}{2}} \frac{\cos\left(\frac{1}{2}+\mu\right)\vartheta}{\cos^{1-2\mu}\vartheta \cdot \sin^{\frac{1}{2}+\mu}\vartheta} e^{-2\zeta \operatorname{tg}\vartheta}\, d\vartheta.$$

Let ϑ_0 be a fixed value in $(0, \pi/2)$, then

$$(26) \qquad (-1)^n Z_\lambda(x_n') \cdot \left\{ \frac{\Gamma(1-\mu)}{\Gamma\left(\frac{1}{2}-\mu\right)\Gamma\left(\frac{1}{2}\right)} 2^{1-2\mu} \right\}^{-1} > \int_0^{\vartheta_0} \cos^{\frac{1}{2}+\mu}\vartheta \cdot \cos\left(\frac{1}{2}+\mu\right)\vartheta \cdot$$

$$\cdot (\operatorname{tg}\vartheta)^{-\frac{1}{2}-\mu} e^{-2\zeta_n \operatorname{tg}\vartheta}\, d \operatorname{tg}\vartheta > \cos^{\frac{1}{2}+\mu}\vartheta_0 \cdot \cos\left(\frac{1}{2}+\mu\right)\vartheta_0 \cdot (2\zeta_n)^{\mu-\frac{1}{2}} \int_0^{2\zeta_n \operatorname{tg}\vartheta_0} x^{-\frac{1}{2}-\mu} e^{-x}\, dx,$$

therefore

$$\varliminf_{n\to\infty} (-1)^n Z_\lambda(x_n') n^{\frac{1}{2}-\mu} \geqq \frac{\Gamma(1-\mu)}{\Gamma\left(\frac{1}{2}\right)} \left(\frac{2}{\pi}\right)^{\frac{1}{2}-\mu} \cos^{\frac{1}{2}+\mu}\vartheta_0 \cdot \cos\left(\frac{1}{2}+\mu\right)\vartheta_0$$

for all $\vartheta_0 \in (0, \pi/2)$, hence

$$\varliminf_{n\to\infty} (-1)^n Z_\lambda(x_n') n^{\frac{1}{2}-\mu} \geqq \frac{\Gamma(1-\mu)}{\Gamma\left(\frac{1}{2}\right)} \left(\frac{2}{\pi}\right)^{\frac{1}{2}-\mu},$$

and this implies with (25) the relation (19).

Using the inequalities

$$\int_0^{2\zeta_n \operatorname{tg}\vartheta_0} x^{-\frac{1}{2}-\mu} e^{-x}\, dx > e^{-2\zeta_n \operatorname{tg}\vartheta_0} \int_0^{2\zeta_n \operatorname{tg}\vartheta_0} x^{-\frac{1}{2}-\mu}\, dx = \frac{1}{\frac{1}{2}-\mu} (2\zeta_n \operatorname{tg}\vartheta_0)^{\frac{1}{2}-\mu} e^{-2\zeta_n \operatorname{tg}\vartheta_0}$$

and

$$\int_{2\zeta_n \operatorname{tg}\vartheta_0}^{\infty} x^{-\frac{1}{2}-\mu} e^{-x}\, dx < (2\zeta_n \operatorname{tg}\vartheta_0)^{-\frac{1}{2}-\mu} \int_{2\zeta_n \operatorname{tg}\vartheta_0}^{\infty} e^{-x}\, dx = (2\zeta_n \operatorname{tg}\vartheta_0)^{-\frac{1}{2}-\mu} e^{-2\zeta_n \operatorname{tg}\vartheta_0}$$

we have

$$\int_0^{2\zeta_n \operatorname{tg} \vartheta_0} x^{-\frac{1}{2}-\mu} e^{-x}\, dx > \frac{2\zeta_n \operatorname{tg} \vartheta_0}{2\zeta_n \operatorname{tg} \vartheta_0 + \frac{1}{2} - \mu} \int_0^\infty x^{-\frac{1}{2}-\mu} e^{-x}\, dx =$$

$$= \frac{2\zeta_n \operatorname{tg} \vartheta_0}{2\zeta_n \operatorname{tg} \vartheta_0 + \frac{1}{2} - \mu} \Gamma\left(\frac{1}{2} - \mu\right),$$

hence by (26)

$$(-1)^n Z_\lambda(x_n') \cdot \left\{ \frac{\Gamma(1-\mu)}{\Gamma\left(\frac{1}{2}-\mu\right)\Gamma\left(\frac{1}{2}\right)} 2^{1-2\mu} \right\}^{-1} > \cos^{\frac{1}{2}+\mu} \vartheta_0 \cos\left(\frac{1}{2}+\mu\right)\vartheta_0 \cdot (2\zeta_n)^{\mu-\frac{1}{2}} \times$$

$$\times \frac{2\zeta_n \operatorname{tg} \vartheta_0}{2\zeta_n \operatorname{tg} \vartheta_0 + \frac{1}{2} - \mu} \Gamma\left(\frac{1}{2}-\mu\right) \geqq \cos^2 \vartheta_0 \cdot \left(\frac{\zeta_1}{\pi}\right)^{\mu-\frac{1}{2}} (2n\pi)^{\mu-\frac{1}{2}} \frac{2\pi \operatorname{tg} \vartheta_0}{2\pi \operatorname{tg} \vartheta_0 + \frac{1}{2}} \Gamma\left(\frac{1}{2}-\mu\right).$$

Let $\operatorname{tg} \vartheta_0 = \frac{1}{2}$, then

$$(-1)^n Z_\lambda(x_n') \cdot \left\{ \frac{\Gamma(1-\mu)}{\Gamma\left(\frac{1}{2}\right)} 2^{1-2\mu} \right\}^{-1} > (2n\pi)^{\mu-\frac{1}{2}} \frac{2}{\sqrt{5}} \frac{4}{5} \frac{2\pi}{2\pi+1},$$

which proves the left inequality in (18).

REFERENCES

[1] WATSON, G. N.: *Theory of Bessel functions,* University Press, Cambridge, 1952.

[2] ELBERT, Á.: On the solutions of the differential equations $y'' + q(x)y = 0$, where $[q(x)]^\nu$ is concave I. *Acta Math. Acad. Sci. Hungar.* **20** (1969) 1—11.

Mathematical Institute of the Hungarian Academy of Sciences

(Received July 14, 1969)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 287—288.

ON A PROBLEM OF W. SCHMIDT

by
E. SZEMERÉDI

SCHMIDT asked the following question. Does there exist a set H of real numbers of infinite measure so that the ratio of two numbers in H is never an integer. In the present paper we are going to construct such a set H. After writing the paper I learned that HAIGHT about simultaneously also constructed such a set.

Let R denote the real line and μ the Lebesgue measure.

THEOREM. *There exists a subset $H \subset R$ with the following properties*

1. *H is Lebesgue measurable and $\mu(H)=\infty$.*
2. *For every $x, y \in H$ and for every integer $n \geqq 2$*

$$nx \neq y.$$

PROOF. Let $p_1, p_2, \ldots, p_m, \ldots$ denote the sequence of primes. We set

$$M_m = (p_m!)^{2^m} \quad \text{for} \quad m=1, 2, \ldots$$

$$M_{m,i} = \frac{M_m}{p_i} \quad \text{for} \quad m=1, 2, \ldots, 1 \leqq i \leqq m$$

$$y_m\left(\sum_{i=1}^{m} \frac{1}{p_i}\right) = \frac{1}{m \log m \log\log m}, \qquad m=2, 3, \ldots \tag{1}$$

$$N_{m,i} = \left[M_{m,i}, M_{m,i} + \frac{y_m}{p_i}\right], \qquad m=2, 3, \ldots \; 1 \leqq i \leqq m.$$

If $A = [a, b]$ we put $\dfrac{A}{t} = \left[\dfrac{a}{t}, \dfrac{b}{t}\right]$ for $t>0$.

We define H as follows:

$$H = \bigcup_{m=m_0}^{\infty} \bigcup_{i=1}^{m} N_{m,i} - \bigcup_{t=2}^{\infty} \bigcup_{m'=m_0}^{\infty} \bigcup_{j=1}^{m'} \frac{N_{m',j}}{t} \tag{2}$$

where m_0 is a sufficiently large integer.

H obviously satisfies 2. by construction. H is obviously Lebesgue measurable too.

To conclude our proof we prove

$$\mu(H) = +\infty. \tag{3}$$

First of all let us remark that by construction we have

(4) $$N_{m,i} \cap \frac{N_{m',j}}{t} = \emptyset \quad \text{for every} \quad m' \leqq m, \quad 1 \leqq i \leqq m, \quad 1 \leqq j \leqq m', \quad 2 \leqq t.$$

There is m_0 for which the following holds:

(5) $$m_0 \leqq m < m', \quad 1 \leqq i \leqq m, \quad 1 \leqq j \leqq m',$$

$$\mu\left(N_{m,i} \cap \left(\bigcup_{t=2}^{\infty} \frac{N_{m',j}}{t}\right)\right) \leqq \frac{\mu(N_{m,i})\mu(N_{m',j})}{M_{m,i}}.$$

On the other hand for suitable m_0 we also have

(6) $$\mu\left(N_{m,i} \cap \left({\bigcup_{t \geqq 2}}' \frac{N_{m',j}}{t}\right)\right) \leqq \frac{\mu(N_{m,i})\mu(N_{m',j})}{M_{m,i} \log\log m'}$$

where $\bigcup'$ is extended for those $t \geqq 2$ all whose prime factors are $\geqq \log m'$. Meditation shows that we have

(7) $$\bigcup_{t=2}^{\infty} \bigcup_{m'=m_0}^{\infty} \bigcup_{j=1}^{m'} \frac{N_{m',j}}{t} = \bigcup_{t=2}^{\infty} \bigcup_{m'=m_0}^{\infty} \bigcup_{j=1}^{\log m'} \frac{N_{m',j}}{t} \cup {\bigcup_{t \geqq 2}}' \bigcup_{m'=m_0}^{\infty} \bigcup_{j=1}^{m'} \frac{N_{m',j}}{t}$$

where $\bigcup'$ is extended all for those $t \geqq 2$ all whose prime factors are $\geqq \log m'$. Thus by (4) and (7) we have

$$\mu\left(N_{m,i} \cap \bigcup_{t=2}^{\infty} \bigcup_{m'=m_0}^{\infty} \bigcup_{j=1}^{m'} \frac{N_{m',j}}{t}\right) \leqq \mu\left(N_{m,i} \cap \left(\bigcup_{t=2}^{\infty} \bigcup_{m'>m} \bigcup_{j=1}^{m'} \frac{N_{m',j}}{t}\right)\right) \leqq$$

$$\leqq \mu\left(N_{m,i} \cap \left(\bigcup_{t=2}^{\infty} \bigcup_{m'>m} \bigcup_{j=1}^{\log m'} \frac{N_{m',j}}{t}\right)\right) + \mu\left(N_{m,i} \cap \left({\bigcup_{t \geqq 2}}' \bigcup_{m'>m} \bigcup_{j=1}^{m'} \frac{N_{m',j}}{t}\right)\right)$$

where $\bigcup'$ means the same as in (6) and (7). Thus by (5), (6) and (7) we have

$$\mu\left(N_{m,i} \cap \left(\bigcup_{t=2}^{\infty} \bigcup_{m'=m_0}^{\infty} \bigcup_{j=1}^{m'} \frac{N_{m',j}}{t}\right)\right) \leqq \frac{\mu(N_{m,i})}{M_{m,i}} \left(\sum_{m'>m} \sum_{j=1}^{\log m'} \mu(N_{m',j}) + \sum_{m'>m} \sum_{j=1}^{m'} \frac{\mu(N_{m',j})}{\log\log m'}\right).$$

Thus by (8) we have

$$\mu(H) \geqq \frac{1}{2} \sum_{m'=m_0}^{\infty} \sum_{j=1}^{m'} \mu(N_{m',j}) = \frac{1}{2} \sum_{m'=m_0}^{\infty} \sum y_{m'} \left(\sum_{i=1}^{m'} \frac{1}{p_i}\right) =$$

$$= \frac{1}{2} \sum_{m'=m_0}^{\infty} \frac{1}{m' \log m' \log\log m'} = \infty$$

Q.e.d.

Mathematical Intsitute of the Hungarian Academy of Sciences, Budapest

(Received October 13, 1969)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 289—300.

DEUX THÉORÈMES SUR LA PREMIÈRE FORME FONDAMENTALE DE LA GÉOMÉTRIE ÉQUIAFFINE

par

J. MERZA

Dans cet article nous allons déduire deux théorèmes à la caractérisation géométrique de la première forme fondamentale de la géométrie différentielle équiaffine. Par ces théorèmes la signification géométrique des quantités en question deviendra plus claire. Le premier donne une construction affine-métrique à la détermination de la première forme fondamentale des hypersurfaces de l'espace affine à $(n+1)$ dimensions, l'autre exprime quelle relation existe entre le changement des espaces tangents de l'hypersurface et une partie de la première forme fondamentale. Au cours des calculs l'accomplissement des conditions analytiques nécessaires sera toujours supposé. Commençons par l'énumeration des formules fondamentales de la théorie équiaffine des hypersurfaces.

On appelle hypersurface à n dimensions l'ensemble de points de l'espace affine A_{n+1} à $(n+1)$ dimensions donné par l'équation

$$\mathbf{r} = \mathbf{r}(u^1, \ldots, u^n) \tag{1}$$

si et seulement si

$$\operatorname{rang} \left\| \frac{\partial \mathbf{r}}{\partial u^\alpha} \right\| = n \qquad (\alpha = 1, \ldots, n).$$

On introduit la métrique équiaffine à l'aide des quantités

$$G_{\alpha\beta} = \frac{\Lambda_{\alpha\beta}}{|\Lambda|^{\frac{1}{n+2}}}, \qquad (\beta = 1, \ldots, n)$$

où

$$\Lambda = \det \|\Lambda_{\alpha\beta}\|$$

et les fonctions $\Lambda_{\alpha\beta}$ sont définies par la formule

$$\Lambda_{\alpha\beta} = \left| \frac{\partial^2 \mathbf{r}}{\partial u^\alpha \partial u^\beta}, \frac{\partial \mathbf{r}}{\partial u^1}, \ldots, \frac{\partial \mathbf{r}}{\partial u^n} \right|.$$

Les quantités $G_{\alpha\beta}$ sont les composantes d'un tenseur et la forme différentielle

$$\varphi = G_{\alpha\beta}\, du^\alpha du^\beta$$

est la première forme fondamentale de l'hypersurface. On écrit la seconde forme fondamentale (ou bien la forme cubique), pareillement au cas à deux dimensions, sous la forme

$$\psi = \frac{1}{|G|^{1/2}} \cdot \left| d^3\mathbf{r}, \frac{\partial \mathbf{r}}{\partial u^1}, \ldots, \frac{\partial \mathbf{r}}{\partial u^n} \right| - \frac{3}{2} d\varphi = A_{\alpha\beta\gamma}\, du^\alpha\, du^\beta\, du^\gamma,$$

puis on définit la normale affine à l'aide de la formule

$$\mathbf{n} = \frac{1}{n} \cdot \Delta \mathbf{r} = \frac{1}{n|G|^{1/2}} \cdot \frac{\partial}{\partial u^\beta} \left(|G|^{1/2} \cdot G^{\alpha\beta} \frac{\partial \mathbf{r}}{\partial u^\alpha} \right),$$

où $G = \det \|G_{\alpha\beta}\|$ et Δ dénote l'opérateur différentiel de second ordre de Beltrami. Les fonctions

$$\Gamma^{*\alpha}_{\beta\gamma} = \Gamma^\alpha_{\beta\gamma} + A^\alpha_{\beta\gamma}$$

où les $\Gamma^\alpha_{\beta\gamma}$ sont les symboles de Christoffel formés des coefficients de la première forme fondamentale, déterminent sur l'hypersurface une connexion symétrique affine. Ces paramètres de connexion nous permettent d'écrire les équations de Gauss sous la forme

$$\frac{\partial^2 \mathbf{r}}{\partial u^\alpha\, \partial u^\beta} = \Gamma^{*\varrho}_{\alpha\beta} \frac{\partial \mathbf{r}}{\partial u^\varrho} + G_{\alpha\beta}\, \mathbf{n}. \tag{2}$$

Rappelons encore une formule de la théorie affine des surfaces, notamment la relation

$$\left| \mathbf{n}, \frac{\partial \mathbf{r}}{\partial u^1}, \ldots, \frac{\partial \mathbf{r}}{\partial u^n} \right| = |G|^{1/2}. \tag{3}$$

Considérons une courbe $u^\alpha = u^\alpha(t)$ sur l'hypersurface (1). On peut déterminer la longueur d'arc équiaffine de cette courbe en partant de l'équation

$$\mathbf{r}^*(t) = \mathbf{r}[u^1(t), \ldots, u^n(t)]$$

par l'emploi de la formule

$$s = \int_{t_1}^{t_2} |\dot{\mathbf{r}}^*, \ddot{\mathbf{r}}^*, \ldots, \overset{(n)}{\mathbf{r}}{}^*|^{\frac{2}{n(n+1)}}\, dt.$$

Naturellement, nous ne nous occupons que des courbes dont la longueur d'arc existe. Nous écrivons l'équation de la courbe rapportée à son arc affine sous la forme $u^\alpha = u^\alpha(s)$ et, pour simplifier, nous appliquerons la notation $\mathbf{r}(s)$ à la fonction $\mathbf{r}[u^1(s), \ldots, u^n(s)]$. Choisissons le point P_0 de la courbe et comptons de ce point la longueur d'arc affine. Soit P un autre point de la courbe qui appartient à la valeur s du paramètre. Mesurons la distance équiaffine du point P_0 au point P dans l'espace selon la formule

$$d = \frac{1}{|G|^{1/2}} \cdot \left| \mathbf{r}(s) - \mathbf{r}(0), \left(\frac{\partial \mathbf{r}}{\partial u^1}\right)_0, \ldots, \left(\frac{\partial \mathbf{r}}{\partial u^n}\right)_0 \right|. \tag{4}$$

Cette expression est la généralisation naturelle de la formule appliquée par W. Blaschke à la détermination de la distance d'un point de l'espace à un point d'une surface.

Développons en série la différence figurante dans la première colonne du déterminant de l'expression (4) et substituons le résultat reçu

$$\mathbf{r}(s)-\mathbf{r}(0)=\mathbf{r}'(0)\cdot s+\mathbf{r}''(0)\frac{s^2}{2}+0(3)$$

dans la formule. A cause de la décomposition

$$\mathbf{r}'=\frac{\partial \mathbf{r}}{\partial u^\sigma}u'^\sigma$$

on peut supprimer le vecteur $\mathbf{r}'(0)\cdot s$ de la première colonne du déterminant

$$d=\frac{1}{|G|^{1/2}}\cdot\left|\mathbf{r}'(0)\cdot s+\mathbf{r}''(0)\frac{s^2}{2}+O(3),\ \left(\frac{\partial \mathbf{r}}{\partial u^1}\right)_0,\ \ldots,\ \left(\frac{\partial \mathbf{r}}{\partial u^n}\right)_0\right|.$$

Par l'emploi des équations (2) de Gauss on peut substituer le vecteur $\mathbf{r}''$ pa

$$\mathbf{r}''=\frac{\partial^2\mathbf{r}}{\partial u^\alpha\,\partial u^\beta}u'^\alpha u'^\beta+\frac{\partial \mathbf{r}}{\partial u^\alpha}u''^\alpha=$$

$$=\left(\Gamma^{*\varrho}_{\alpha\beta}\frac{\partial \mathbf{r}}{\partial u^\varrho}+G_{\alpha\beta}\mathbf{n}\right)u'^\alpha u'^\beta+\frac{\partial \mathbf{r}}{\partial u^\alpha}u''^\alpha=$$

$$=(u''^\varrho+\Gamma^{*\varrho}_{\alpha\beta}u'^\alpha u'^\beta)\frac{\partial \mathbf{r}}{\partial u^\varrho}+G_{\alpha\beta}u'^\alpha u'^\beta\mathbf{n}.$$

Remarquons que ce n'est que la composante dans la direction du vecteur **n** qu'on doit substituer dans le déterminant parce que les autres membres seraient éliminés par la soustraction des autres colonnes multipliées par un facteur convenable. De cette manière on obtient le résultat

$$d=\frac{1}{|G|^{1/2}}\cdot\left|G_{\alpha\beta}u'^\alpha u'^\beta\cdot\frac{s^2}{2}\mathbf{n}+O(3),\ \frac{\partial \mathbf{r}}{\partial u^1},\ \ldots,\ \frac{\partial \mathbf{r}}{\partial u^n}\right|$$

(tous les coefficients sont mis au point P_0) qu'on peut transformer en

$$d=\frac{G_{\alpha\beta}u'^\alpha u'^\beta}{|G|^{1/2}}\cdot\frac{s^2}{2}\cdot\left|\mathbf{n},\ \frac{\partial \mathbf{r}}{\partial u^1},\ \ldots,\ \frac{\partial \mathbf{r}}{\partial u^n}\right|+O(3).$$

Prenons en considération l'équation (3), divisons par s^2 et passons à la limite pour énoncer le théorème suivant.

THÉORÈME. *Soit d la distance de deux points d'une courbe sur la surface, où cette distance est calculée à l'aide de la formule* (4), *soit s la longueur d'arc affine de la courbe et* $\varphi(s) = G_{\alpha\beta} u'^{\alpha} u'^{\beta}$ *la valeur de la première forme fondamentale de la surface associée à la direction du vecteur tangent de la courbe. En ce cas on a la relation*

$$\lim_{s \to 0} \frac{2d}{s^2} = \varphi(0).$$

Remarque. Si l'on définit la distance centro-affine du point $\mathbf{z}$ au point $\mathbf{r}$ d'une surface quelconque par

$$d_c = \frac{\left| \mathbf{z} - \mathbf{r}, \dfrac{\partial \mathbf{r}}{\partial u^1}, \ldots, \dfrac{\partial \mathbf{r}}{\partial u^n} \right|}{\left| \mathbf{r}, \dfrac{\partial \mathbf{r}}{\partial u^1}, \ldots, \dfrac{\partial \mathbf{r}}{\partial u^n} \right|}$$

le théorème précédent sera valable dans la géométrie centro-affine aussi en supposant qu'on remplace les quantités équiaffines par celles de la géométrie centro-affine.

Pour prouver le second théorème, considérons l'hypersurface (1) et supposons que le point $\mathbf{r}_0 = \mathbf{r}(u^1, \ldots, u^n)$ est fixé. Choisissons pour repère les $n+1$ vecteurs qui sont composés des vecteurs tangents des lignes de paramètre passant par ce point et de la normale affine. Les calculs seront faits par rapport à ce système. Fixons sur les lignes de paramètre passantes par le point r_0 les points

$$\mathbf{r}_i = \mathbf{r}(u^1, \ldots, u^{i-1}, u^i + \Delta u^i, u^{i+1}, \ldots, u^n) \qquad (i = 1, \ldots, n)$$

et ajoutons à ceux le point

$$\mathbf{r}_{n+1} = \mathbf{r}(u^1 + \Delta u^1, u^2 + \Delta u^2, \ldots, u^n + \Delta u^n).$$

La construction se commence par la détermination des équations des plans tangents de la surface aux $n+2$ points reçus. Rappelons encore une fois que les calculs seront faits par rapport au repère au point $\mathbf{r}(u^1, \ldots, u^n)$ et désignons les coordonnées relatives à ce système par $y^1, \ldots, y^{n+1}$. Faisons encore une convention: nous ne ferons pas sommation dans ce qui suit aux indices latins, les indices de sommation seront désignés par des caractères grecques. Nous appliquons la convention d'Einstein mais si ce nous paraît utile nous écriverons des signes de sommation aussi.

Il est trivial que l'équation du plan tangent au point $\mathbf{r}_0$ est

$$y^{n+1} = 0.$$

Pour déterminer les équations des plans tangents aux points $\mathbf{r}_i$ ($i = 1, \ldots, n+1$) employons la formule

$$\left| \mathbf{R} - \mathbf{r}_i, \left(\frac{\partial \mathbf{r}}{\partial u^1}\right)_i, \ldots, \left(\frac{\partial \mathbf{r}}{\partial u^n}\right)_i \right| = 0. \tag{5}$$

Nous voulons avoir les équations dans le système de coordonnées $y^1, \ldots, y^{n+1}$ c'est pourquoi il faut que nous décomposons les vecteurs de la formule (5) dans le repère mobile. En partant du développement

$$\mathbf{r}_i = \mathbf{r}(u^1, \ldots, u^{i-1}, u^i + \Delta u^i, u^{i+1}, \ldots, u^n) =$$

$$= \mathbf{r}_0 + \left(\frac{\partial \mathbf{r}}{\partial u^i}\right)_0 \Delta u^i + \left(\frac{\partial^2 \mathbf{r}}{\partial (u^i)^2}\right)_0 \frac{(\Delta u^i)^2}{2} + O(3) =$$

$$= \mathbf{r}_0 + \left(\frac{\partial \mathbf{r}}{\partial u^i}\right)_0 \Delta u^i + \underset{(0)}{\Gamma^{*\varrho}_{ii}} \left(\frac{\partial \mathbf{r}}{\partial u^\varrho}\right)_0 \frac{(\Delta u^i)^2}{2} + \underset{(0)}{G_{ii}} \mathbf{n}_0 \frac{(\Delta u^i)^2}{2} + O(3)$$

$$(\varrho = 1, \ldots, n)$$

et en tenant compte de la relation

$$\frac{\partial \mathbf{r}}{\partial u^i} = \delta^\varrho_i \frac{\partial \mathbf{r}}{\partial u^\varrho}$$

on obtiendra au vecteur $\underset{(i)}{\mathbf{y}}$ du point $\mathbf{r}_i$ l'expression (nous omettons les indices 0, auf le cas $\mathbf{r}_0$)

$$\underset{(i)}{\mathbf{y}} = \mathbf{r}_i - \mathbf{r}_0 = \left(\delta^\varrho_i \Delta u^i + \Gamma^{*\varrho}_{ii} \frac{(\Delta u^i)^2}{2} + O(3)\right) \frac{\partial \mathbf{r}}{\partial u^\varrho} + \left(G_{ii} \frac{(\Delta u^i)^2}{2} + O(3)\right) \mathbf{n}.$$

Le développement

$$\left(\frac{\partial \mathbf{r}}{\partial u^k}\right)_i = \left(\frac{\partial \mathbf{r}}{\partial u^k}\right)_0 + \left(\frac{\partial^2 \mathbf{r}}{\partial u^k \partial u^i}\right)_0 \Delta u^i + O(2) =$$

$$= \left(\frac{\partial \mathbf{r}}{\partial u^k}\right)_0 + \underset{(0)}{\Gamma^{*\varrho}_{ki}} \left(\frac{\partial \mathbf{r}}{\partial u^\varrho}\right)_0 \Delta u^i + G_{ki} \Delta u^i \mathbf{n} + O(2),$$

ou, omettant les indices 0,

$$\left(\frac{\partial \mathbf{r}}{\partial u^k}\right)_i = (\delta^\varrho_k + \Gamma^{*\varrho}_{ki} \Delta u^i + O(2)) \frac{\partial \mathbf{r}}{\partial u^\varrho} + (G_{ki} \Delta u^i + O(2)) \mathbf{n}$$

nous donne les expressions pour les dérivées partielles nécessaires. En utilisant les coordonnées $\underset{(i)}{y^j}$ relatives au repère choisi, l'équation du plan tangent S_i au point $\mathbf{r}_i$ $(i = 1, \ldots, n)$ sera

$$\left| \mathbf{Y} - \underset{(i)}{\mathbf{y}}, \left(\frac{\partial \mathbf{r}}{\partial u^1}\right)_i, \ldots, \left(\frac{\partial \mathbf{r}}{\partial u^n}\right)_i \right| = 0,$$

ou en détail

$$\begin{vmatrix} y^1 - \underset{(i)}{y^1} & \delta^1_1 + O(1) \ldots & \delta^1_n + O(1) \\ y^2 - \underset{(i)}{y^2} & \delta^2_1 + O(1) \ldots & \delta^2_n + O(1) \\ \cdot & \cdot \quad \cdots & \cdot \\ y^n - \underset{(i)}{y^n} & \delta^n_1 + O(1) \ldots & \delta^n_n + O(1) \\ y^{n+1} - \underset{(i)}{y^{n+1}} & G_{1i} \Delta u^i + O(2) \ldots & G_{ni} \Delta u^i + O(2) \end{vmatrix} = 0.$$

En développant le déterminant selon la première colonne nous avons l'équation

$$\sum_{\varrho=1}^{n}(y^{\varrho}-\underset{(i)}{y^{\varrho}})(-1)^{\varrho+1}\begin{vmatrix} \delta_1^1+O(1) & \dots & \delta_n^1+O(1) \\ \cdot & \dots & \cdot \\ \delta_1^{\varrho-1}+O(1) & \dots & \delta_n^{\varrho-1}+O(1) \\ \delta_1^{\varrho+1}+O(1) & \dots & \delta_n^{\varrho+1}+O(1) \\ \cdot & \dots & \cdot \\ \delta_1^n+O(1) & \dots & \delta_n^n+O(1) \\ G_{1i}\Delta u^i+O(2) & \dots & G_{ni}\Delta u^i+O(2) \end{vmatrix}+$$

$$+(-1)^{n+2}(y^{n+1}-\underset{(i)}{y^{n+1}})\begin{vmatrix} \delta_1^1+O(1) & \dots & \delta_n^1+O(1) \\ \cdot & \dots & \cdot \\ \delta_1^n+O(1) & \dots & \delta_n^n+O(1) \end{vmatrix}=$$

$$=\sum_{\varrho=1}^{n}(y^{\varrho}-\underset{(i)}{y^{\varrho}})(-1)^{\varrho+1}(-1)^{n-\varrho}\begin{vmatrix} \delta_1^1+O(1) & \dots & \delta_n^1+O(1) \\ \cdot & \dots & \cdot \\ \delta_1^{\varrho-1}+O(1) & \dots & \delta_n^{\varrho-1}+O(1) \\ G_{1i}\Delta u^i+O(2) & \dots & G_{ni}\Delta u^i+O(2) \\ \delta_1^{\varrho+1}+O(1) & \dots & \delta_n^{\varrho+1}+O(1) \\ \cdot & \dots & \cdot \\ \delta_1^n+O(1) & \dots & \delta_n^n+O(1) \end{vmatrix}+$$

$$+(-1)^{n+2}(y^{n+1}-\underset{(i)}{y^{n+1}})\begin{vmatrix} \delta_1^1+O(1) & \dots & \delta_n^1+O(1) \\ \cdot & \dots & \cdot \\ \delta_1^n+O(1) & \dots & \delta_n^n+O(1) \end{vmatrix}=0$$

et en divisant par $(-1)^{n+1}$ nous obtenons l'équation du plan S_i:

$$(y^{\varrho}-\underset{(i)}{y^{\varrho}})\big(G_{\varrho i}\Delta u^i+O(2)\big)-(y^{n+1}-\underset{(i)}{y^{n+1}})\big(1+O(1)\big)=0.$$

On sait déjà que

$$\underset{(i)}{y^{j}}=\delta_i^j\Delta u^i+O(2)$$

$$\underset{(i)}{y^{n+1}}=G_{ii}\frac{(\Delta u^i)^2}{2}+O(3),$$

c'est pourquoi

$$\big(y^{\varrho}-\delta_i^{\varrho}\Delta u^i+O(2)\big)\big(G_{\varrho i}\Delta u^i+O(2)\big)-\left(y^{n+1}-G_{ii}\frac{(\Delta u^i)^2}{2}+O(3)\right)\big(1+O(1)\big)=0,$$

d'où, après avoir supprimé les termes d'ordre supérieur, on reçoit l'équation du plan S_i sous la forme

$$G_{\varrho i}\Delta u^i y^{\varrho}-y^{n+1}=\frac{1}{2}G_{ii}(\Delta u^i)^2.$$

Si l'on veut trouver l'équation du plan S_{n+1} il ne faut qu'apercevoir qu'en augmentant toutes les variables on doit faire sommation à l'indice i aussi. Cela constaté, en partant des formules

$$\underset{(n+1)}{y^j} = \delta^j_\varrho \Delta u^\varrho + \frac{1}{2} \Gamma^{*j}_{\varrho\sigma} \Delta u^\varrho \Delta u^\sigma + O(3)$$

$$\underset{(n+1)}{y^{n+1}} = \frac{1}{2} G_{\varrho\sigma} \Delta u^\varrho \Delta u^\sigma + O(3)$$

on recevra l'équation

$$G_{\varrho\sigma} \Delta u^\sigma y^\varrho - y^{n+1} = \frac{1}{2} G_{\varrho\sigma} \Delta u^\varrho \Delta u^\sigma$$

du plan S_{n+1}.

Ayant les équations des plans S_k $(k = 0, 1, \dots, n+1)$ nous commençons à déterminer les sommets du simplex à $n+1$ dimensions formé par ces plans. On obtiendra les $n+2$ points nécessaires comme les solutions des $n+1$ équations qui se résultent du système

$$-y^{n+1} = 0$$

$$G_{\varrho i} \Delta u^i y^\varrho - y^{n+1} = \frac{1}{2} G_{ii} (\Delta u^i)^2$$

$$G_{\varrho\sigma} \Delta u^\sigma y^\varrho - y^{n+1} = \frac{1}{2} G_{\varrho\sigma} \Delta u^\varrho \Delta u^\sigma$$

par la suppression de l'une des équations. Si c'est la $l^{\text{ième}}$ équation qui est supprimée, le point déterminé par le système reçu sera désigné par P_{l-1}.

Le déterminant formé des coefficients du système qui sert à la détermination du point P_o est

$$\begin{vmatrix} G_{11} \Delta u^1 & G_{21} \Delta u^1 & \dots & G_{n1} \Delta u^1 & -1 \\ G_{12} \Delta u^2 & G_{22} \Delta u^2 & \dots & G_{n2} \Delta u^2 & -1 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ G_{1n} \Delta u^n & G_{2n} \Delta u^n & \dots & G_{nn} \Delta u^n & -1 \\ G_{1\sigma} \Delta u^\sigma & G_{2\sigma} \Delta u^\sigma & \dots & G_{n\sigma} \Delta u^\sigma & -1 \end{vmatrix}.$$

Si l'on soustrait la première, deuxième, ..., $(n-1)^{\text{ième}}$ lignes l'une après les autres de la ligne dernière on aura la valeur $(n-1) G \Delta u^1 \dots \Delta u^n$ du déterminant.

Nous verrons plus loin que la détermination des quantités $\underset{(0)}{y^1}, \dots, \underset{(0)}{y^n}$ est superflue et c'est seulement la coordonnée $\underset{(0)}{y^{n+1}}$ qui est nécessaire. On recevra sa valeur en développant le déterminant

$$\begin{vmatrix} G_{11} \Delta u^1 & G_{21} \Delta u^1 & \dots & G_{n1} \Delta u^1 & \frac{1}{2} G_{11} (\Delta u^1)^2 \\ G_{12} \Delta u^2 & G_{22} \Delta u^2 & \dots & G_{n2} \Delta u^2 & \frac{1}{2} G_{22} (\Delta u^2)^2 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ G_{1n} \Delta u^n & G_{2n} \Delta u^n & \dots & G_{nn} \Delta u^n & \frac{1}{2} G_{nn} (\Delta u^n)^2 \\ G_{1\sigma} \Delta u^\sigma & G_{2\sigma} \Delta u^\sigma & \dots & G_{n\sigma} \Delta u^\sigma & \frac{1}{2} G_{\varrho\sigma} \Delta u^\varrho \Delta u^\sigma \end{vmatrix}.$$

L'application de la soustraction réitérée des lignes de la $(n+1)^{\text{ième}}$ ligne et l'introduction de la notation

$$\Phi = \varphi - \sum_{\varrho=1}^{n} G_{\varrho\varrho}(\Delta u^{\varrho})^2$$

nous donne le résultat $\frac{1}{2}\Phi G\Delta u^1 \dots \Delta u^n$ c'est pourquoi

$$\underset{(0)}{y}^{n+1} = \frac{\Phi}{2(n-1)}.$$

Passons ensuite au calcul des coordonnées des points P_i $(i=1, \dots, n)$. Ces points sont situés dans le plan $y^{n+1}=0$, à cause de cela nous ne devons calculer que les coordonnées $\underset{(i)}{y}^1, \dots, \underset{(i)}{y}^n$. Le déterminant du système

$$-y^{n+1} = 0$$

$$G_{\varrho k}\,\Delta u^k\, y^{\varrho} - y^{n+1} = \frac{1}{2} G_{kk}(\Delta u^k)^2 \qquad (k=1, \dots, i-1, i+1, \dots, n)$$

$$G_{\varrho\sigma}\,\Delta u^{\sigma}\, y^{\varrho} - y^{n+1} = \frac{1}{2} G_{\varrho\sigma}\,\Delta u^{\varrho}\,\Delta u^{\sigma}$$

qui sert à la détermination des coordonnées du point P_i est

$$\begin{vmatrix} 0 & 0 & \dots 0 & -1 \\ G_{11}\,\Delta u^1 & G_{21}\,\Delta u^1 & \dots G_{n1}\,\Delta u^1 & -1 \\ \cdot & \cdot & \cdots\ \cdot & \cdot \\ G_{1,i-1}\,\Delta u^{i-1} & G_{2,i-1}\,\Delta u^{i-1} & \dots G_{n,i-1}\,\Delta u^{i-1} & -1 \\ G_{1,i+1}\,\Delta u^{i+1} & G_{2,i+1}\,\Delta u^{i+1} & \dots G_{n,i+1}\,\Delta u^{i+1} & -1 \\ \cdot & \cdot & \cdots\ \cdot & \cdot \\ G_{1n}\,\Delta u^n & G_{2n}\,\Delta u^n & \dots G_{nn}\,\Delta u^n & -1 \\ G_{1\sigma}\,\Delta u^{\sigma} & G_{2\sigma}\,\Delta u^{\sigma} & \dots G_{n\sigma}\,\Delta u^{\sigma} & -1 \end{vmatrix}.$$

On peut l'écrire immédiatement sous la forme

$$(-1)^{n+3}\begin{vmatrix} G_{11}\,\Delta u^1 & G_{21}\,\Delta u^1 & \dots G_{n1}\,\Delta u^1 \\ \cdot & \cdot & \cdots\ \cdot \\ G_{1,i-1}\,\Delta u^{i-1} & G_{2,i-1}\,\Delta u^{i-1} & \dots G_{n,i-1}\,\Delta u^{i-1} \\ G_{1,i+1}\,\Delta u^{i+1} & G_{2,i+1}\,\Delta u^{i+1} & \dots G_{n,i+1}\,\Delta u^{i+1} \\ \cdot & \cdot & \cdots\ \cdot \\ G_{1n}\,\Delta u^n & G_{2n}\,\Delta u^n & \dots G_{nn}\,\Delta u^n \\ G_{1\sigma}\,\Delta u^{\sigma} & G_{2\sigma}\,\Delta u^{\sigma} & \dots G_{n\sigma}\,\Delta u^{\sigma} \end{vmatrix}.$$

En effectuant la soustraction réitérée des lignes de la ligne dernière puis en transposant la $n^{\text{ième}}$ ligne dans la $i^{\text{ième}}$ ligne on aura au déterminant l'expression $(-1)^{3-i}G\Delta u^1 \dots \Delta u^n$.

On peut commencer le calcul du déterminant

$$\begin{vmatrix} 0 & \dots 0 & 0 & 0 & \dots 0 & -1 \\ G_{11}\Delta u^1 & \dots G_{k-1,1}\Delta u^1 & \frac{1}{2}G_{11}(\Delta u^1)^2 & G_{k+1,1}\Delta u^1 & \dots G_{n1}\Delta u^1 & -1 \\ . & \dots & . & . & \dots & . \\ G_{1,i-1}\Delta u^{i-1} & \dots G_{k-1,i-1}\Delta u^{i-1} & \frac{1}{2}G_{i-1,i-1}(\Delta u^{i-1})^2 & G_{k+1,i-1}\Delta u^{i-1} & \dots G_{n,i-1}\Delta u^{i-1} & -1 \\ G_{1,i+1}\Delta u^{i+1} & \dots G_{k-1,i+1}\Delta u^{i+1} & \frac{1}{2}G_{i+1,i+1}(\Delta u^{i+1})^2 & G_{k+1,i+1}\Delta u^{i+1} & \dots G_{n,i+1}\Delta u^{i+1} & -1 \\ . & \dots & . & . & \dots & . \\ G_{1n}\Delta u^n & \dots G_{k-1,n}\Delta u^n & \frac{1}{2}G_{nn}(\Delta u^n)^2 & G_{k+1,n}\Delta u^n & \dots G_{nn}\Delta u^n & -1 \\ G_{1\sigma}\Delta u^\sigma & \dots G_{k-1,\sigma}\Delta u^\sigma & \frac{1}{2}G_{\varrho\sigma}\Delta u^\varrho \Delta u^\sigma & G_{k+1,\sigma}\Delta u^\sigma & \dots G_{n\sigma}\Delta u^\sigma & -1 \end{vmatrix}$$

qui sert à la détermination des coordonnées $\underset{(i)}{y^k}$, à la manière précédente puis on développe le déterminant reçu

$$\frac{(-1)^{n+3+n-i}}{2}\times$$

$$\times\begin{vmatrix} G_{11}\Delta u^1 & \dots G_{k-1,1}\Delta u^1 & G_{11}(\Delta u^1)^2 & G_{k+1,1}\Delta u^1 & \dots G_{n1}\Delta u^1 \\ . & \dots & . & . & \dots \quad . \\ G_{1,i-1}\Delta u^{i-1} & \dots G_{k-1,i-1}\Delta u^{i-1} & G_{i-1,i-1}(\Delta u^{i-1})^2 & G_{k+1,i-1}\Delta u^{i-1} & \dots G_{n,i-1}\Delta u^{i-1} \\ G_{1,i}\Delta u^i & \dots G_{k-1,i}\Delta u^i & \varphi-\sum_{\varrho=1}^{n} G_{\varrho\varrho}(\Delta u^\varrho)^2+G_{ii}(\Delta u^i)^2 & G_{k+1,i}\Delta u^i & \dots G_{ni}\Delta u^i \\ G_{1,i+1}\Delta u^{i+1} & \dots G_{k-1,i+1}\Delta u^{i+1} & G_{i+1,i+1}(\Delta u^{i+1})^2 & G_{k+1,i+1}\Delta u^{i+1} & \dots G_{n,i+1}\Delta u^{i+1} \\ . & \dots & . & . & \dots \quad . \\ G_{1n}\Delta u^n & \dots G_{k-1,n}\Delta u^n & G_{nn}(\Delta u^n)^2 & G_{k+1,n}\Delta u^n & \dots G_{nn}\Delta u^n \end{vmatrix}$$

selon la $k^{\text{ième}}$ colonne. On obtient l'expression

$$\frac{(-1)^{3-i}}{2}\left\{\sum_{\substack{\varrho=1\\ \varrho\neq i}}^{n} G_{\varrho\varrho}(\Delta u^\varrho)^2 GG^{\varrho k}\Delta u^1\dots\Delta u^{\varrho-1}\Delta u^{\varrho+1}\dots\Delta u^n+\right.$$

$$\left.+[\Phi+G_{ii}(\Delta u^i)^2]GG^{ik}\Delta u^1\dots\Delta u^{i-1}\Delta u^{i+1}\dots\Delta u^n\right\}=$$

$$=\frac{(-1)^{3-i}}{2}G\left[\sum_{\varrho=1}^{n} G_{\varrho\varrho}(\Delta u^\varrho)^2 G^{\varrho k}\Delta u^1\dots\Delta u^{\varrho-1}\Delta u^{\varrho+1}\dots\Delta u^n+\right.$$

$$\left.+\Phi G^{ik}\Delta u^1\dots\Delta u^{i-1}\Delta u^{i+1}\dots\Delta u^n\right].$$

Si l'on introduit la notation

$$B^i=\sum_{\varrho=1}^{n} G_{\varrho\varrho}(\Delta u^\varrho)^2 G^{i\varrho}\Delta u^1\dots\Delta u^{\varrho-1}\Delta u^{\varrho+1}\dots\Delta u^n$$

on trouve à la coordonnée $\underset{(i)}{y^k}$ la relation

$$\underset{(i)}{y^k} = \frac{(-1)^{3-i}G}{2(-1)^{3-i}G\,\Delta u^1\ldots\Delta u^n}\,(B^k+\Phi G^{ik}\,\Delta u^1\ldots\Delta u^{i-1}\,\Delta u^{i+1}\ldots\Delta u^n) =$$

$$= \frac{1}{2\prod\limits_{\varrho=1}^{n}\Delta u^{\varrho}}\,(B^k+\Phi G^{ik}\,\Delta u^1\ldots\Delta u^{i-1}\,\Delta u^{i+1}\ldots\Delta u^n).$$

En employant les moyens déjà vus on a pour les coordonnées du point P_{n+1} les expressions

$$\underset{(n+1)}{y^k} = \frac{(-1)^{n+3}GB^k}{2(-1)^{n+3}G\prod\limits_{\varrho=1}^{n}\Delta u^{\varrho}} = \frac{B^k}{2\prod\limits_{\varrho=1}^{n}\Delta u^{\varrho}} \qquad (k=1,\ldots,n)$$

et c'est trivial que $\underset{(n+1)}{y^{n+1}}=0$.

Le volume du simplex à $n+1$ dimensions déterminé par les points $P_0, P_1, \ldots$ $\ldots, P_{n+1}$ est fourni par la formule

$$V = \frac{1}{(n+1)!}\,|\underset{(1)}{\mathbf{y}}-\underset{(0)}{\mathbf{y}},\ \underset{(2)}{\mathbf{y}}-\underset{(0)}{\mathbf{y}},\ \ldots,\ \underset{(n)}{\mathbf{y}}-\underset{(0)}{\mathbf{y}},\ \underset{(n+1)}{\mathbf{y}}-\underset{(0)}{\mathbf{y}}|$$

qu'on peut écrire sous la forme

$$V = \frac{(-1)^{n+1}}{(n+1)!}\left|\begin{array}{cccc} \underset{(0)}{y^1} & \underset{(1)}{y^1} & \ldots & \underset{(n+1)}{y^1} \\ \vdots & \vdots & & \vdots \\ \underset{(0)}{y^{n+1}} & \underset{(1)}{y^{n+1}} & \ldots & \underset{(n+1)}{y^{n+1}} \\ 1 & 1 & \ldots & 1 \end{array}\right|.$$

Substituons les coordonnées dans ce déterminant. Nous écrivons le résultat schématiquement de la manière suivante

$$V = \frac{(-1)^{n+1}}{(n+1)!}\left|\begin{array}{ccc} \begin{array}{c}\underset{(0)}{y^1}\\ \vdots \\ \underset{(0)}{y^n}\end{array} & \dfrac{B^k+\Phi G^{ik}\Delta u^1\ldots\Delta u^{i-1}\,\Delta u^{i+1}\ldots\Delta u^n}{2\prod\limits_{\varrho=1}^{n}\Delta u^{\varrho}} & \dfrac{B^k}{2\prod\limits_{\varrho=1}^{n}\Delta u^{\varrho}} \\ \dfrac{\Phi}{2(n-1)} & 0\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 0 & 0 \\ 1 & 1\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 1 & 1 \end{array}\right| =$$

$$= \frac{(-1)\Phi}{(n+1)!\,2(n-1)}\left|\begin{array}{cc} \dfrac{\Phi G^{ki}\Delta u^1\ldots\Delta u^{i-1}\,\Delta u^{i+1}\ldots\Delta u^n}{2\prod\limits_{\varrho=1}^{n}\Delta u^{\varrho}} & \dfrac{B^k}{2\prod\limits_{\varrho=1}^{n}\Delta u^{\varrho}} \\ 0\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 0 & 1 \end{array}\right|.$$

Le développement de ce déterminant, par l'emploi de l'identité

$$\det \|G^{ik}\| = \frac{1}{G}$$

donne au volume V_S^* du simplex

$$V_S^* = \frac{\Phi^{n+1}}{(n+1)!(1-n)2^{n+1} G \prod_{\varrho=1}^{n} \Delta u^\varrho}.$$

Après avoir eu le volume du simplex déterminé par les plans S_k ($k = 0, 1, \dots, n+1$) nous le comparons au volume du simplex déterminé par les vecteurs $\mathbf{r}_i - \mathbf{r}_0$ ($i = 1, \dots, n$) et par le vecteur normal. En partant du développement

$$\mathbf{r}_i - \mathbf{r}_0 = \frac{\partial \mathbf{r}}{\partial u^\varrho} \Delta u^\varrho + O(2)$$

relatif au repère mobile, on voit que le volume cherché est donné par

$$V_R^* = \frac{1}{(n+1)!} \begin{vmatrix} \Delta u^1 & 0 & \dots 0 & 0 \\ 0 & \Delta u^2 & \dots 0 & 0 \\ . & . & \dots . & . \\ 0 & 0 & \dots \Delta u^n & 0 \\ O(2) & O(2) & \dots O(2) & 1 \end{vmatrix} = \frac{\prod_{\varrho=1}^{n} \Delta u^\varrho}{(n+1)!}$$

et par conséquence

$$V_S^* \cdot V_R^* = \frac{\Phi^{n+1}}{[(n+1)!](1-n)2^{n+1} G}.$$

On sait que la valeur V^* du volume calculée dans le repère mobile et la valeur V du volume mesurée à l'aide des vecteurs de base de l'espace sont jointes par la relation

$$V = (-1)^n |G|^{1/2} V^*.$$

En appliquant ce fait dans notre calcul nous voyons que

$$\Phi^{n+1} = 2^{n+1}(1-n)[(n+1)!]^2 V_S \cdot V_R$$

d'où, par l'emploi de la relation

$$\Phi = \varphi - \sum_{\varrho=1}^{n} G_{\varrho\varrho}(\Delta u^\varrho)^2 = 2 \sum_{\binom{n}{2}} G_{\varrho\sigma} \Delta u^\varrho \Delta u^\sigma,$$

on déduit

$$\sum_{\binom{n}{2}} G_{\varrho\sigma} \Delta u^\varrho \Delta u^\sigma \approx \sqrt[n+1]{|[(n+1)!]^2 (n-1) V_S \cdot V_R|}$$

ce que nous voulions démontrer.

Si l'on introduit la notation

$$c = \sqrt[n+1]{[(n+1)!]^2(n-1)}$$

on peut énoncer le théorème suivant.

THÉORÈME. *Considérons sur les lignes de paramètre passant par le point* $\mathbf{r}_0$ *de la surface les points*

$$\mathbf{r}_i = \mathbf{r}(u^1, \ldots, u^{i-1}, u^i + \Delta u^i, u^{i+1}, \ldots, u^n) \qquad (i = 1, \ldots, n)$$

et le point $\mathbf{r}_{n+1} = \mathbf{r}(u^1 + \Delta u^1, \ldots, u^n + \Delta u^n)$. *Désignons par* V_S *le volume du simplex déterminé par les plans tangents aux points* $\mathbf{r}_0, \mathbf{r}_1, \ldots, \mathbf{r}_{n+1}$ *et par* V_R *celui du simplex formé des vecteurs* $\mathbf{r}_i - \mathbf{r}_0$, $\mathbf{n}$ *(*$\mathbf{n}$ *est la normale affine). Ces volumes sont liés à la première forme fondamentale de la surface par la relation*

$$\sum_{\alpha \neq \beta} G_{\alpha\beta}\, \Delta u^\alpha \Delta u^\beta \approx c \cdot \sqrt[n+1]{|V_S \cdot V_R|}\,.$$

BIBLIOGRAPHIE

[1] BLASCHKE, W.: *Vorlesungen über Differentialgeometrie* II., Berlin, 1923.
[2] MERZA, J.: L'introduction de la différentiation absolue dans l'espace affine, *Publ. Math. Debrecen* **5** (1958), 330—337.
[3] SCHIROKOW, P. A.—SCHIROKOW, A. P.: *Affine Differentialgeometrie,* Leipzig, 1962.

Institut de Mathématiques de l'Académie des Sciences de Hongrie

(Reçu le 17 Octobre 1969.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 301—306.

A NOTE ON OPERATOR TRANSFORMATIONS

by

T. FÉNYES and P. KOSIK

Introduction. In his paper [1] GESZTELYI has introduced the concept F of operator transformation as a linear map of M into M where M is the field of Mikusiński operators. GESZTELYI [1] defined also the continuity and multiplicity of operator transformation and proved interesting theorems related to such operator transformations.

The purpose of this paper is to show, in addition to GESZTELYI's interesting paper, some elementary properties of operator transformations especially of T^α and U_k.

1. On operator transformations of logarithms

Notation: The set of linear, continuous and multiplicative operator transformations is denoted by $\mathcal{T}_1$. The set of logarithms is denoted by Λ.

THEOREM 1. *An $F \in \mathcal{T}_1$ maps Λ into itself.*

The PROOF is analogous to that of the special case given by GESZTELYI [1] and the formula

$$F(e^{-\lambda w}) = e^{-\lambda F(w)}, \qquad w \in \Lambda \tag{1}$$

holds.

Remark. If F has a continuous inverse F^{-1}, we conclude that F is a bijective map of Λ onto itself.

Examples.

$$F = T^\alpha \quad \text{Here is} \quad F^{-1} = T^{-\alpha}$$

$$F = U_k \quad \text{Here is} \quad F^{-1} = U_{\frac{1}{k}}$$

2. On the operator transformations T^α and U_k

In the sequel $T^\alpha(x)$, $U_k(x)$ will be considered as operator functions of the (real) variable α $(-\infty < \alpha < \infty)$, respectively k $(0 < k < \infty)$.

Notation: Let $-\infty < \lambda_1 \leqq \lambda \leqq \lambda_2 < \infty$. The symbol $\mu_\nu(\lambda_1, \lambda_2)$ denotes the set of operator functions being ν times continuously differentiable on the finite interval $\lambda_1 \leqq \lambda \leqq \lambda_2$ in the sense of MIKUSIŃSKI [2]. The symbols $\mu_\nu(-\infty, \infty)$, $\mu_\nu(\varrho, \infty)$ denote the set of operator functions being ν times continuously differentiable on every finite interval contained in $-\infty<\lambda<\infty$, respectively $\varrho\leqq\lambda<\infty$.

First we prove the following

LEMMA A. *If* $f(\lambda)\in\mu_1(\lambda_1,\lambda_2)$ *and* $f(\lambda)\in\mu_1(\lambda_3,\lambda_4)$, *where* $\lambda_1<\lambda_3<\lambda_2<\lambda_4$ *then*

$$f(\lambda)\in\mu_1(\lambda_1,\lambda_4)$$

PROOF. Since

$$f(\lambda)=q_1\{f_1(\lambda,t)\},\quad q_1=\frac{\{a_1\}}{\{b_1\}}\in M,\qquad \lambda_1\leqq\lambda\leqq\lambda_2$$

and

$$f(\lambda)=q_2\{f_2(\lambda,t)\},\quad q_2=\frac{\{a_2\}}{\{b_2\}}\in M,\qquad \lambda_3\leqq\lambda\leqq\lambda_4$$

where $\frac{\partial f_1}{\partial\lambda},\ \frac{\partial f_2}{\partial\lambda}$ exist and are continuous in $\lambda_1\leqq\lambda\leqq\lambda_2$, $t\geqq0$, respectively $\lambda_3\leqq\lambda\leqq\lambda_4$, $t\geqq0$, we get

$$(2)\qquad \begin{aligned} f(\lambda)&=\frac{\{F_1(\lambda,t)\}}{\{b_1\}\{b_2\}},\qquad \lambda_1\leqq\lambda\leqq\lambda_2\\ f(\lambda)&=\frac{\{F_2(\lambda,t)\}}{\{b_1\}\{b_2\}},\qquad \lambda_3\leqq\lambda\leqq\lambda_4\end{aligned}$$

where

$$(3)\qquad \begin{aligned}\{F_1(\lambda,t)\}&=\{a_1\}\{b_2\}\{f_1(\lambda,t)\}\\ \{F_2(\lambda,t)\}&=\{a_2\}\{b_1\}\{f_2(\lambda,t)\}\end{aligned}$$

It can be seen from (2) that

$$F_1(\lambda)=F_2(\lambda)\quad\text{if}\quad \lambda_3\leqq\lambda\leqq\lambda_2$$

Consequently (2) can be written as

$$(4)\qquad f(\lambda)=\frac{\{F(\lambda,t)\}}{\{b_1\}\{b_2\}}\qquad \lambda_1\leqq\lambda\leqq\lambda_4$$

where $\frac{\partial F}{\partial\lambda}$ exists and is continuous on $\lambda_1\leqq\lambda\leqq\lambda_4$, $t\geqq0$

THEOREM 2. *Let* $x\in M$ *be arbitrary and fixed. If*

$$T^\alpha(x)\in\mu_1(\alpha_1,\alpha_2)$$

then

$$T^\alpha(x)\in\mu_\infty(-\infty,\infty)$$

and the following formula holds:

$$(5)\qquad \frac{d^iT^\alpha(x)}{d\alpha^i}=(-1)^iD^{(i)}[T^\alpha(x)],\qquad i=1,2,\dots,$$

If

$$U_k(x)\in\mu_1(k_1,k_2)\qquad k_1,k_2>0$$

then

$$U_k(x) \in \mu_\infty(\varrho, \infty) \quad \text{for} \quad \varrho > 0$$

and

$$\frac{dU_k(x)}{dk} = -\frac{s}{k} D[U_k(x)] \tag{6}$$

holds.

PROOF of the case $T^\alpha(x)$. If $x = \{f\} \in C$, then $T^\alpha(f) = \{e^{\alpha t} f\}$ and the theorem is trivially true.

If $x = \frac{\{f\}}{\{g\}} \in M$, from $T^\alpha(x) \in \mu_1(\alpha_1, \alpha_2)$ the formula

$$\frac{d}{d\alpha} T^\alpha(x) = \frac{d}{d\alpha} \frac{T^\alpha(f)}{T^\alpha(g)} = \frac{-D[T^\alpha(f)]\, T^\alpha(g) + D[T^\alpha(g)]\, T^\alpha(f)}{[T^\alpha(g)]^2} \qquad (\alpha_1 \leqq \alpha \leqq \alpha_2) \tag{7}$$

follows. Since the operator transformations T^α, D are commutable, we obtain

$$\frac{d}{d\alpha} T^\alpha(x) = \frac{T^\alpha[fD(g) - gD(f)]}{T^\alpha(\{g\}^2)} = T^\alpha \left[\frac{fDg - gDf}{g^2} \right] = \tag{8}$$

$$= -T^\alpha[D(x)] = -D[T^\alpha(x)]$$

We refer to the following result of GESZTELYI [1] (page 184, Theorem 5. 4).

LEMMA B. *If F is a linear, continuous operator transformation and $f(\lambda) \in \mu_1(\lambda_1, \lambda_2)$, then $F[f(\lambda)] \in \mu_1(\lambda_1, \lambda_2)$.*

So we obtain that $T^\alpha(x) \in \mu_\infty(\alpha_1, \alpha_2)$ and the formula

$$\frac{d^i T^\alpha(x)}{d\alpha^i} = (-1)^i D^{(i)}[T^\alpha(x)], \qquad i = 1, 2, \ldots, \ (\alpha_1 \leqq \alpha \leqq \alpha_2) \tag{9}$$

holds.

Now we show that $T^\alpha(x) \in \mu_\infty(-\infty, \infty)$. In fact, let β be fixed such that $0 < \beta < \alpha_2 - \alpha_1$. Then

$$T^{-\beta}[T^\alpha(x)] = T^{\alpha-\beta}[x] \in \mu_1(\alpha_1 + \beta, \alpha_2 + \beta) \tag{10}$$

Applying T^β we see by Lemma B that $T^\alpha(x) \in \mu_1(\alpha_1 + \beta, \alpha_2 + \beta)$ and by Lemma A that

$$T^\alpha(x) \in \mu_1(\alpha_1, \alpha_2 + \beta) \tag{11}$$

We can easily conclude that $T^\alpha(x) \in \mu_1(\alpha_1, \alpha_2 + \gamma)$ where $\gamma > 0$ is arbitrary, i.e.

$$T^\alpha(x) \in \mu_1(\alpha_1, \infty)$$

and similarly $T^\alpha(x) \in \mu_1(-\infty, \alpha_2)$ so that

$$T^\alpha(x) \in \mu_1(-\infty, \infty)$$

But then

$$T^\alpha(x) \in \mu_\infty(-\infty, \infty) \tag{12}$$

PROOF of the case $U_k(x)$. If $x=\{f\}\in C$, then

$$U_k(f)=\{kf(kt)\}=s\left\{k\int_0^t f(k\tau)\,d\tau\right\}=s\left\{\int_0^{kt} f(u)\,du\right\}$$

and

$$\frac{dU_k(f)}{dk}=s\{tf(kt)\}=\frac{s}{k}\{ktf(kt)\}=-\frac{s}{k}DU_k(f) \tag{13}$$

By Lemma B, $DU_k(f)\in\mu_1(\varrho,\infty)$ and also $\dfrac{DU_k(f)}{k}\in\mu_1(\varrho,\infty)$. Since

$$DU_k(f)=\frac{1}{k}U_kD(f),$$

$$\frac{dU_k(f)}{dk}=-\frac{s}{k^2}U_k[D(f)]$$

we easily conclude that $U_k(f)\in\mu_\infty(\varrho,\infty)$. Consequently, if $x=\{f\}$, the Theorem is true.

If $x=\dfrac{\{f\}}{\{g\}}\in M$, from $U_k(x)\in\mu_1(k_1,k_2)$ the formula

$$\frac{d}{dk}U_k(x)=\frac{d}{dk}\frac{U_k(f)}{U_k(g)}=\frac{-\frac{s}{k}DU_k(f)U_k(g)+\frac{s}{k}DU_k(g)U_k(f)}{[U_k(g)]^2}= \tag{14}$$

$$=-\frac{s}{k}\frac{DU_k(f)U_k(g)-DU_k(g)U_k(f)}{[U_k(g)]^2}=-\frac{s}{k}D\left[\frac{U_k(f)}{U_k(g)}\right]=$$

$$=-\frac{s}{k}DU_k\left(\frac{f}{g}\right)=-\frac{s}{k}DU_k(x)\qquad(k_1\leqq k\leqq k_2)$$

follows and it can easily be seen that $U_k(x)\in\mu_\infty(k_1,k_2)$.

Now we show that $U_k(x)\in\mu_\infty(\varrho,\infty)$. In fact, let k_0 be fixed such that $1<k_0<\dfrac{k_2}{k_1}$.

Then

$$U_{\frac{1}{k_0}}U_k(x)=U_{\frac{k}{k_0}}(x)\in\mu_1(k_0k_1,k_0k_2) \tag{15}$$

Applying U_{k_0} we see by Lemma B that $U_k(x)\in\mu_1(k_0k_1,k_0k_2)$ and by Lemma A that

$$U_k(x)\in\mu_1(k_1,k_0k_2) \tag{16}$$

We can easily conclude that

$$U_k(x)\in\mu_1(k_1,\gamma k_2)$$

where $\gamma > 1$ is arbitrary, i.e. $U_k(x) \in \mu_1(k_1, \infty)$ and similarly $U_k(x) \in \mu_1(\varrho, k_2)$, so that

$$U_k(x) \in \mu_1(\varrho, \infty) \quad \text{for} \quad \varrho > 0$$

But then

$$U_k(x) \in \mu_\infty(\varrho, \infty) \quad \text{for} \quad \varrho > 0 \tag{17}$$

The derivative of $T^\alpha(x)$ and of $U_k(x)$ has an interesting property which will be proved in

THEOREM 3. *If $T^\alpha(x) \in \mu_1(-\infty, \infty)$ and its derivative vanishes at any $\alpha = \alpha_0$, then x is a complex number and the derivative of $T^\alpha(x)$ vanishes identically on $-\infty < \alpha < \infty$. Consequently if $x \notin K$, the derivative of $T^\alpha(x)$ can not vanish anywhere.*

The same is true in the case of $U_k(x)$ $(k > 0)$

PROOF. By the above assumption

$$\frac{d}{d\alpha} T^\alpha(x)\big|_{\alpha=\alpha_0} = -D[T^{\alpha_0}(x)] = 0 \tag{18}$$

holds. It follows from a result of MIKUSIŃSKI [3] that $T^{\alpha_0}(x) = c$ where c is a complex number, so that $x = T^{-\alpha_0}(c) = c$, $T^\alpha(x) = T^\alpha(c) = c$ for every α. Consequently

$$\frac{d}{d\alpha} T^\alpha(x) \equiv 0$$

On the other hand we have

$$\left.\frac{dU_k(x)}{dk}\right|_{k=k_0} = -\frac{s}{k_0} D[U_{k_0}(x)] = 0 \tag{19}$$

So we obtain $D[U_{k_0}(x)] = 0$, $U_{k_0}(x) = c$ so that $x = U_{\frac{1}{k_0}}(c) = c$, $U_k(x) = U_k(c) = c$ for every $k > 0$. Consequently

$$\frac{d}{dk} U_k(x) \equiv 0$$

will be obtained.

Finally we expose the following questions:

1. Are the operator functions $T^\alpha(x)$, $U_k(x)$ continuously differentiable for every $x \in M$, or not?

2. Do a non-logarithmic operator $x \in M$ and a non-trivial $F \in \mathcal{T}_1$ exist for which the operator

$$y = F(x)$$

is a logarithm?

The authors are indebted to Prof. G. FREUD for his valuable remarks.

REFERENCES

[1] GESZTELYI, E.: Über lineare Operatortransformationen. *Publicationes Mathematicae* **19** (1967) 169—206.
[2] MIKUSIŃSKI, J.: *Operational Calculus.* Pergamon Press, New York 1959.
[3] MIKUSIŃSKI, J.: Remarks on the algebraic derivative in the Operational Calculus. *Studia Math.* **19** (1960) 187—192.

Mathematical Institute of the Hungarian Academy of Sciences

(Received December 10. 1969)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 307—316.

О РЕГУЛЯРНЫХ ПРОГРАММАХ

Л. ДЬЮРИШ

В настоящей работе подробно излагается „язык регулярных программ", очерченный в [6][1]. Определяется синтаксис языка с помощью операций введенных в [2] (и их обобщений). Появление операционных параметров в этих операциях может получать существенную роль в исследованиях теории программирования. Способ построения системы $P(\mathfrak{A}, \mathfrak{B})$, определен нами и тоже иллюстрирует и подчеркнует это. Теория регулярных программ даст возможность совместно исследовать синтаксические и семантические (теоретические и практические) проблемы программирования [4].

В §1. определяем регулярные программы, в §2. рассматриваем соотношения между элементами системы $P(\mathfrak{A}, \mathfrak{B})$.

§ 1.

Перед определением системы регулярных программ кратко обоснуем метод построения.

Из определения [2] микропрограммно-алгебраическо системы $(\mathfrak{A}, \mathfrak{B})$ вытекает, что язык описывающий микропрограммы является универсальным, т. е. в случае удачного выбора исходных операторов и логических условий любую трансформацию состояний мохно выразить микропрограммой (или последовательностью микропрограмм). — Система $(\mathfrak{A}, \mathfrak{B})$ оказалась безусловно эффективным средством с точки зрения конструирования вычислительных машин и др. Некоторые проблемы программирования [6] требуют её дальнейшего развития.

При описании сложных программ или систем программ, и вообще в связи с синтаксическими проблемами программирования оказывается необходимым принимать во внимание тот практический факт, что по существу одна микропрограмма состоит из последовательности операций, которые соответствуют одному „циклу", т. е. служит для конкретного описания одной команды [2].

Поэтому имеем в виду то подмножество множества всевозможных выражений представимых в системе Глушкова, элементы которого соответствуют командам. Выражения, созданные над этим множеством с помощью умножения, условной дизъюнкции и условной итерации опишут „грубую струк-

[1] Понятие регулярных программ было введено в [1] В. Глушковым. Здесь оперируем иным определением (эквивалентным [1]).

руту" (и только эту) программ. Легко понять, что „грубую структуру" любой программы можно выразить в таком смысле.

Очевидно, что можно формулизировать значение программ (как трансформацию состояний [4]) с помощью микропрограмм, выражающих „тонкую структуру". Значит синтаксические и семантические проблемы программирования являются формулизируемыми с помощью системы $P(\mathfrak{A}, \mathfrak{B})$ построенной таким образом.

(I) часть определения системы $P(\mathfrak{A}, \mathfrak{B})$ по существу состоит из системы $(\mathfrak{A}, \mathfrak{B})$ определенной в [2], хотя требуются некоторых модификаций или специализаций. Пусть дано базисное множество M (множество внутренних состояний вычислительной машины). — Если не имеем в виду интерпретации (на M), тогда речь идёт о „схемах" микропрограмм и программ [4].

(I) Регулярные микропрограммы

I. Обозначим конечное множество операторов (исходных микроопераций) через $\mathfrak{U}=\{x_1, x_2, \ldots, x_n, e\}$. (Отображения базисного множества в себе, е: символ идентичного отображения.) $\mathfrak{B}=\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$: конечное множество исходных логических условий. (Логические условия вполне (не частично) определены над базисным множеством.)

II. Основные операции:

В $\mathfrak{A}$: ассоциативное умнежоние (означает последовательное выполнение операторов); $\mathfrak{A}$ является полугруппой.

В $\mathfrak{B}$: конъюнкция, дизъюнкция, отрицание. (Тождества булевой алгебры выполняются в $\mathfrak{B}$.)

Внешние операции:

1. Умножение оператора и логического условия: значение условия $\beta = X\cdot\alpha$ равно значению α после выполнения оператора X.
2. α-дизъюнкция: оператор $Z=(X \overset{\alpha}{\vee} У)$ равно X в случае $\alpha=1$, иначе $У$.
3. α-итерация: оператор $Z=\underset{\alpha}{\{X\}}$ равно первому такому элементу последовательности $e, X, X^2, \ldots, X^n, \ldots$, для которого уже $\alpha=1$.

III. Выражения, полученные из элементов $\mathfrak{A}$ и $\mathfrak{B}$ конечное число раз применяя операции II. называются *регулярными микропрограммами*.

Очевидно можем говорить о безусловных и условных регулярных микропрограммах.

Отметим, что принятие $(\mathfrak{A}, \mathfrak{B})$ в качестве „системы", а не „алгебры" является несущественным отклонением от [2]. Те же самые операции здесь служат конструированию новых выражений.

Система, определенная на основах I—III. называется системой регулярных микропрограмм, и обозначается через $(\mathfrak{A}, \mathfrak{B})$.

Для дальнейшего необходимо точно сформулировать, какое подмножество множества всевозможных вырежений описываемих в системе $(\mathfrak{A}, \mathfrak{B})$ может соответствовать командам. Те регулярные микропрограммы являются командами или элементарными программами, которые опишут функционирование композиции автоматов введенной в [2] с некоторого начального состаяния до терминального состояния управляющего автомата (т. е. которые относятся к одному циклу).

Условимся в том, что существует специальная регулярная микропрограмма E среди приказов, соответствующая идентичной трансформации базисного множества.

(2) Регулярные программы

I. Каждая команда является (элементарной) регулярной программой.

II. Если P_1, P_2 и P_3 регулярные программы, тогда результат следующих операций регулярная программа тоже:

1. Умножение: $P = P_1 \cdot P_2$. (Означает последовательное выполнение трансформации описанных P_1 и P_2.)
2. Условная дизъюнкция: $P = P_1 \vee_{P_3} P_2$ равно P_1, если условие у P_3 выполняется, иначе P_2.
3. Условная итерация: $P = *_{P_2}(P_1)$ равно первому элементу из $E, P_1, P_1^2, \ldots, P_1^n, \ldots$, для которого условие у P_2 уже выполняется.

III. *Регулярные программы:* регулярные программы, предписанные I. и II. (и только эти). Система регулярных программ: $P(\mathfrak{A}, \mathfrak{B})$.

Примечания.

1. В (2) II. 2,3 условие считается выполненным в случае безусловных регулярных программ, в противном слущае речь идёт о выполнении условия описанного условными регулярными программами.
2. Каждую программу (микропрограмму) можно сформулировать регулярной программой (микропрограммой) [1]. Наше определение и определение [1] являются эквивале нтнымив содержательном смысле, поэтому каждую программу можно представить в системе $P(\mathfrak{A}, \mathfrak{B})$.

Пример.

‚Схема программы' [3] $\mathfrak{A} \equiv P_1 \underset{1}{\uparrow} A_1 P_2 \underset{2}{\uparrow} \underset{1}{\downarrow} A_2 \underset{2}{\downarrow}$ интерпретируя соответствует ‚программе' [7].

$\mathfrak{A}$ как регулярная программа имеет следующий вид:

$$\mathfrak{A} \equiv \big(A_1 \cdot (A_2 \vee_{P_2} E)\big) \vee_{P_1} A_2 .$$

Мы можем дать значение (семантическое описание) для $\mathfrak{A}$ таким образом, что конкретно сформулируем её элементы на языке [2]:

$$A_1 = \underset{\alpha_1}{(X_1 \vee X_2)} \cdot \underset{\alpha_2}{\{X_3 \cdot X_2\}} \cdot X_1 ,$$

$$A_2 = X_1 \cdot X_2 \cdot \underset{\alpha_2}{(X_2 \vee X_3)} \cdot \underset{\alpha_2}{\{X_4\}},$$

$$P_1 = \underset{\alpha_2}{\{X_2 \cdot X_3\}} \cdot X_4 \cdot X_1 \cdot \alpha_1 ,$$

$$P_2 = \underset{\alpha_1}{(X_1 \vee X_2)} \cdot \underset{\alpha_2}{\{X_2\}} \cdot X_3 \cdot \alpha_2 ;$$

E: команда, соответствующая идентичной трансформации;

x_1, x_2, x_3, x_4: данные отображения множества состояний,

α_1, α_2: логические условия, определенные над множеством состояний.

Этот пример тоже иллюстрирует, что построение системы $P(\mathfrak{A}, \mathfrak{B})$ может оказаться полезным с точки зрения программирования.

§ 2.

Определение (по [2]): Регулярные микропрограммы (и программы) являются *эквивалентными* тогда и только тогда, если опишут ту же самую трансформацию состояний.

Не занимаемся тождествами системы $(\mathfrak{A}, \mathfrak{B})$ (см. [2]).

Некоторые из следующих соотнощений непосредственно вытекают из определения системы.

Мы делим соотношения эквивалентности на две части:

I. Предположим, что операционные параметры служат только выражению условий, не опишут трансформации состояний.

II. Любая регулярная программа может быть операционным параметром.

I.

Пусть P_1, P_2, P_3 и P_4 будут произвольными, Q, Q_1 и Q_2 соответствующимц предположению I. регулярными программами.

I. (1) $$P_1 \vee_Q P_1 = P_1.$$

I. (2) $$(P_1 \vee_Q P_2) \cdot P_3 = P_1 \vee_Q (P_2 \vee_Q P_3).$$

I. (3) $$(P_1 \cdot P_2) \cdot P_3 = P_1 \cdot (P_2 \cdot P_3).$$

I. (4) $$(P_1 \vee_Q P_2) \cdot P_3 = (P_1 \cdot P_3) \vee_Q (P_2 \cdot P_3).$$

Очевидно, что если условие у Q выполняется, тогда обе стороны опишут трансформацию состояний соответствующую $(P_1 \cdot P_3)$, в противном случае $(P_2 \cdot P_3)$.

I. (5) $$P_1 \cdot (P_2 \vee_Q P_3) = (P_1 \cdot P_2) \vee_Q (P_1 \cdot P_3).$$

I. (6) $P_1 \vee_{Q_1} P_2 = P_2 \vee_{Q_2} P_1$ справедливо только тогда, если отрицание условия у Q_1 равно условию у Q_2 или наоборот.

I. (7) $$(P_1 \vee_Q P_2) \cdot (P_3 \vee_Q P_4) = (P_1 \cdot P_3) \vee_Q (P_1 \cdot P_4) \vee_Q (P_2 \cdot P_3) \vee_Q (P_2 \cdot P_4).$$

Из-за ассоциативности условной дизъюнкции (I. (2)) опишем правую сторону в следующем виде:

$$(((P_1 \cdot P_3) \vee_Q (P_1 \cdot P_4)) \vee_Q (P_2 \cdot P_3)) \vee_Q (P_2 \cdot P_4).$$

Так легко установить, что обе стороны означают трансформацию состояний, соответствующую $(P_1 \cdot P_3)$ в случае, если условие у Q выполняется, иначе $(P_2 \cdot P_4)$.

I. (8) $$*_Q(*_Q(P_1)) = *_Q(P_1).$$

По определению условной итерации левая сторона опишет трансформацию соответствующую первой регулярной программе из

$$E, *_Q(P_1), (*_Q(P_1))^2, \dots, (*_Q(P_1))^k, \dots$$

для которой уже условие у Q выполняется. Если условие уже вначале выполняется, тогда речь идёт об E, иначе нужно рассматривать $*_Q(P_1)$. А это означает трансформацию, описанную первым элементом

$$E, P_1, P_1^2, \dots, P_1^n, \dots$$

для которого уже условие у Q выполняется. Если такой элемент существует, тогда трансформация, соответствующая этому, равна трансформации, описанной левой стороной; если нет, тогда не существует тоже

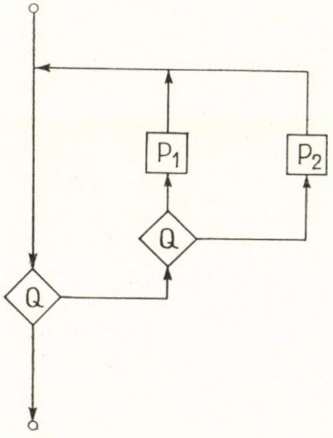
$$\text{для } (*_Q(P_1)^2), \text{ для } (*_Q(P_1))^3, \dots$$

Имея в виду и значение выражения правой стороны очевидно справедливость тождества.

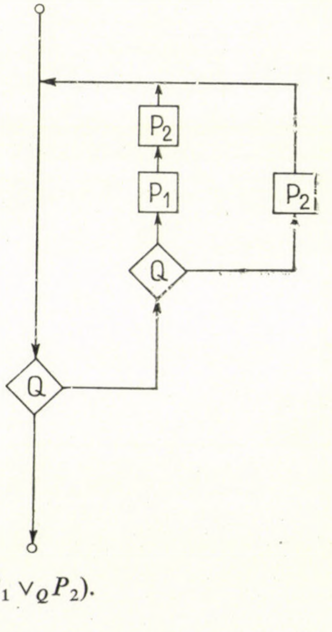
I. (9) $$*_Q(P_1 \vee_Q P_2) = *_Q((P_1 \cdot P_2) \vee_Q P_2).$$

Доказательство подобно I. (*8*), поэтому только иллюстрируем с помощью граф-схемной интерпретации.

Левая сторона:



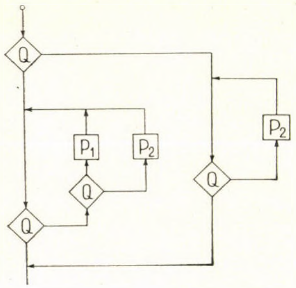
Правая сторона:



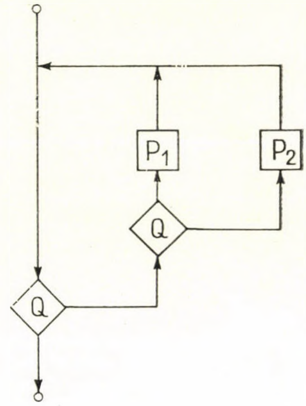
I. (10) $$(*_Q(P_1 \vee_Q P_2)) \vee_Q (*_Q(P_2)) = *_Q(P_1 \vee_Q P_2).$$

Только иллюстрируем и это

Левая сторона:



Правая сторона:

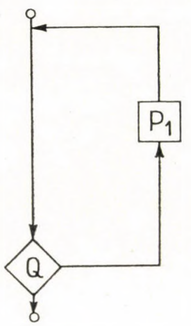


I. (11) $$P_1 \cdot E = E \cdot P_1 = P_1.$$

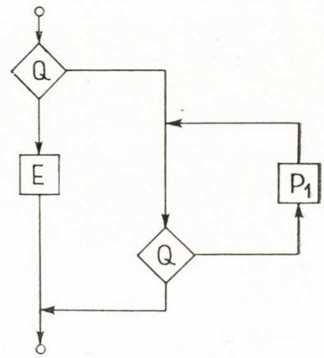
I. (12) $$*_Q(P_1) = E \vee_Q \left(*(P_1)\right).$$

Интерпретируя с граф-схемами:

Левая сторона:



Правая сторона:



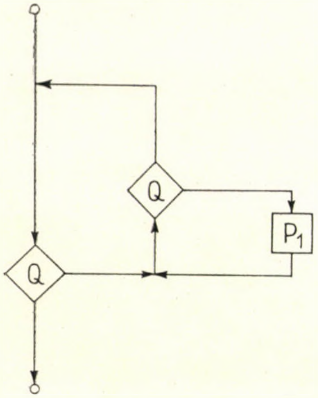
I. (13) $$\left(*_Q(P_1)\right) \vee_Q \left(E \vee_Q \left(*_Q(P_1)\right)\right) = E \vee_Q \left(*_Q(P_1)\right).$$

Эти соотнешения тождественно выполняются для всех программ, поэтому это тождество можно получать з помощью I. (1) и I. (12) таким образом, что в I. (1) заменяем P_1 с $*_Q(P_1)$ и $E \vee_Q\left(N_Q(P_1)\right)$ из-за I. (12). Значит обе стороны опишут ту же самую трансформацию состояний.

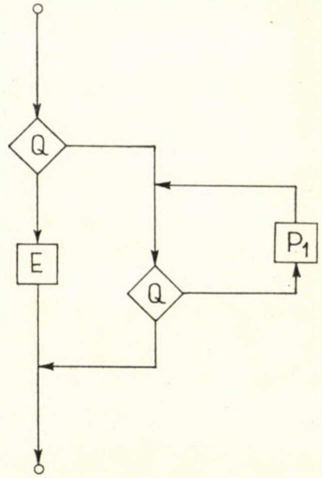
I. (14) $$*_Q\left(*_Q(P_1)\right) = E \vee_Q \left(*_Q(P_1)\right).$$

Это вытекает из I. (8) и I. (12) подобно предыдущему, но можем представить и с помощью граф-схем.

Левая сторона:



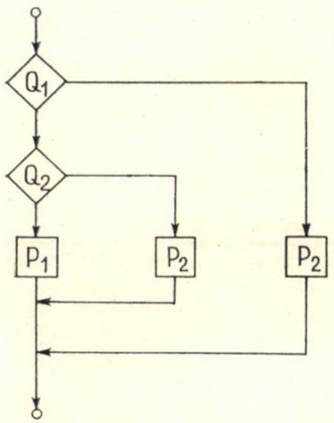
Правая сторона:



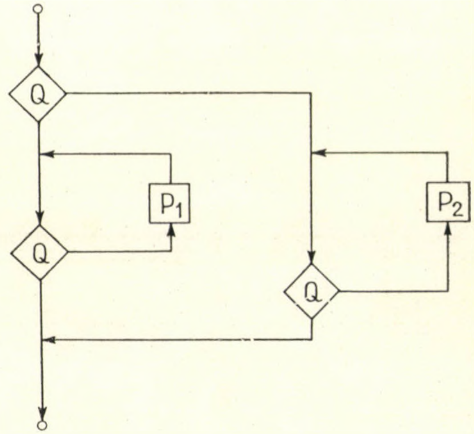
I. (15) $$P_1 \vee_{Q_1 \cdot Q_2} P_2 = (P_1 \vee_{Q_2} P_2) \vee_{Q_1} P_2 .$$

В граф-схемной интерпретации:

Левая сторона:



Правая сторона:



I. (16) $$\big(*_Q\big((P_1 \cdot (P_2 \vee_Q P_3)) \vee_Q ((P_1 \vee_Q P_2) \cdot P_3)\big)\big) \vee_Q \big(*_Q ((P_1 \vee_Q P_2) \cdot P_3)\big) = \\ = *_Q\big(((P_1 \cdot P_2) \vee_Q (P_1 \cdot P_3)) \vee_Q ((P_1 \cdot P_3) \vee_Q (P_2 \cdot P_3))\big).$$

Можно доказать подстановками в I. (10) на основе I. (5), I. (4).

I. (17) $*_Q(P_1) = E$, если условие у Q выполняется.

I. (18) $$*_Q\big((P_1 \cdot P_2) \vee_Q P_2\big) = \big(*_Q(P_1 \vee_Q P_2)\big) \vee_Q \big(*_Q(P_2)\big).$$

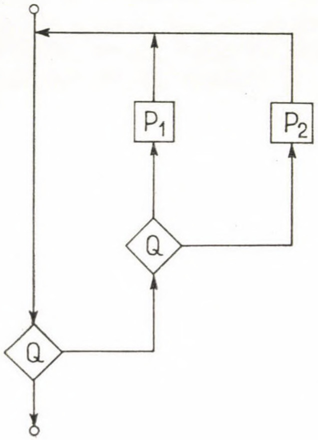
I. (19) $P_1 \vee_{Q_1 \cdot Q_2} P_2 = P_2$, если условие у Q_1 выполняется в противном случае, как условие у Q_2.

I. (20) $$*_Q(P_1 \vee_Q (P_1 \vee_{Q_1 \cdot Q_2} P_2)) = *_Q((P_1 \cdot ((P_1 \vee_{Q_2} P_2) \vee_{Q_1} P_2)) \vee_Q P_2).$$

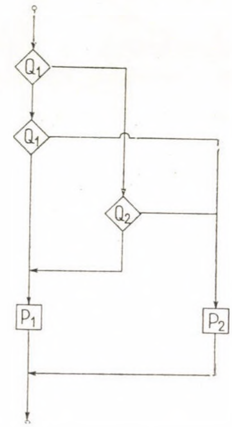
I. (21) $$(*_Q(P_1)) \vee_Q (*_Q(P_2)) = *_Q(P_1 \vee_Q P_2).$$

Интерпретируя с граф-схемами:

Левая сторона:



Правая сторона:

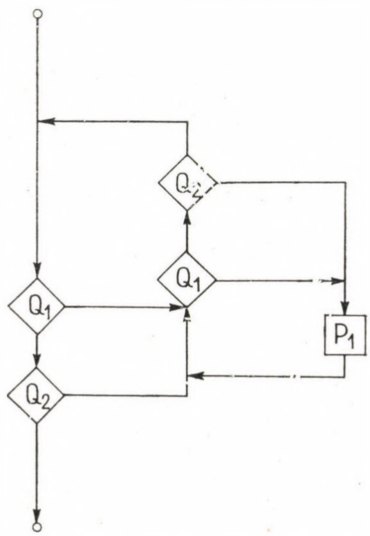


I. (22) $$P_1 \vee_{Q_1 \vee_{Q_1} Q_2} P_2 = P_1,$$

если условие I. (6) выполняется для Q_1 и Q_2.

В граф-схемной интерпретации:

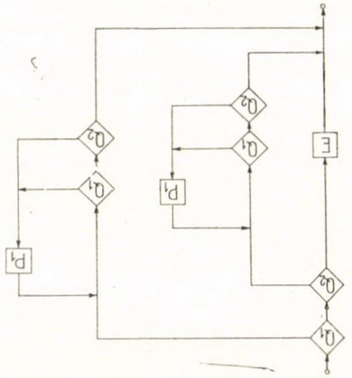
Левая сторона:



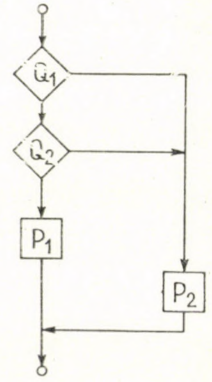
I. (23) $$*_{Q_1 \cdot Q_2}\left(*_{Q_1 \cdot Q_2}(P_1)\right) = \left(E \vee_{Q_2}\left(*_{Q_1 \cdot Q_2}(P_1)\right)\right) \vee_{Q_1}\left(*_{Q_1 \cdot Q_2}(P_1)\right).$$

Интерпретируя:

Левая сторона:



Правая сторона:



II.

Здесь исследуем соотношения системы $P(\mathfrak{A}, \mathfrak{B})$ без ограничительного условия I., значит *любая* регулярная программа может быть операционным параметром.

Естественно I. (3) и I. (11), не инеющие условных операций выполняются без изменений. В дальнейшем мы занимаемся только соотношениями содержащими и условные операции. Некоторые из таких взятых I. тоже справедливы без изменений и здесь.

Легко показать, что I. (4),
I. (5),
I. (9) и
I. (15) выполняются. (Это следует из описанных в 1.)

В случае I. (6) имеем в виду и условие, что два операционных параметра описывают ту же самую трансформацию состояний.

Остальные тождества не выполняются без изменений, но легко выполнить необходимые модификации, поэтому подробно с этим не занимаемся. Покажем с помошью примера, какие модификации могут быть целесообразными.

Соотношение соответствующее I. (1) здесь имеет вид:

$$P_1 \vee_P P_1 = S \cdot P_1,$$

где $P = S \cdot Q$ так, что S является некоторой ‚безусловной' регулярной программой, описывающей некоторую трансформацию состояний; а Q: ‚условная' регулярная программа соответствующая идентичной трансформации состояний.

Примечания:

1. При построенни $P(\mathfrak{A}, \mathfrak{B})$ I. мы уделяли внимание и таким простым соотношениям, как тождества де Моргана в математической логике (I. (21)), которые тоже могут оказаться полезными.
2. Условия выполнимости эквивалентных программ являются тождественными [5], поэтому можно исползовать (т. е. применять) соотношения эквивалентности упомянутых типов и в случае практических проблем оптимизации программирования.
3. Из синтаксической эквивалентности вытекает и семантическая эквивалетность. Таким образом эти результаты можно связать например исрезультатами Ю. Янова. Однако формальная система Ю. Янова не донускате произвольных повторений операторов. Кроме этого понадобятся хорошо внбраннюе соотношения и в том случае, если разработана некоторая формальная система для данной области.

ЛИТЕРАТУРА

[1] В. М. Глушков: *Автоматно-алгебраические аспекты оптимизации управляющих устройств*, Труды Международного Конгресса Математиков (Москва, 1966), Изд-во „Мир”, Москва, 1968, 595—602.

[2] В. М. Глушков: Теория автоматов и формальные преобразования микропрограмм, журн. *Кибернетика*, **5**, К., 1965, I—9.

[3] А. А. Ляпунов: О логических схемах программ, сб. *Проблемы кибернетики*, **I**, М., 1958, 46—74.

[4] А. П. Ершов, А. А. Ляпунов: О формализации понятия программы, журн. *Кибернетика*, **5**, К., 1967, 40—57.

[5] Л. Гюриш: Об исследовании выполнимости программ, журн. *Кибернетика*, **4.**, К, 1970, 55—56.

[6] Gyuris L.: A gluskovi mikroprogram-rendszer módosításáról, Magyar Tudományos Akadémia Számítástechnikai Központja Közlemények, **3** (1967), 61—67.

[7]. Gyuris L.: Interpretált algoritmus-sémák ekvivalenciaproblémájáról, A Magyar Tudományos Akadémia III. Osztálya Közleményei, **18** (1968) 269—272.

(Поступило 8-ого апреля 1970 г.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 317—333.

LIMIT THEOREMS FOR TOTAL VARIATION OF CARTESIAN PRODUCT MEASURES *

by

I. VAJDA

1. Introduction. In the present paper we shall deal with asymptotic properties of the total variation $V(P^n, Q^n)$ of two Cartesian product measures P^n, Q^n. It is supposed that P^n, Q^n are defined on a product measurable space $(X^n, \mathscr{X}^n)$ by

$$P^n = \prod_{i=1}^{n} P_i, \quad Q^n = \prod_{i=1}^{n} Q_i, \qquad n=1, 2, \ldots, \infty, \tag{1}$$

where P_i, Q_i are totally finite measures defined on a measurable space $(X, \mathscr{X})$.

In basic theorems of the present paper, new necessary and sufficient conditions for validity of the relation

$$\lim_n [P^n(X^n)+Q^n(X^n)-V(P^n, Q^n)] = 0 \tag{2}$$

are found and the rate of convergence in (2) is investigated.

One of the fields where such considerations and results are relevant is statistics. Suppose that $P^n(X^n)=Q^n(X^n)=1$, i.e. that P^n, Q^n are probability measures, $n=1, 2, \ldots$ $\ldots, \infty$. If we consider the problem of testing the simple hypothesis $H:P^n$ against the simple alternative $K:Q^n$ on the basis of one observation of a random vector $(\xi_1, \xi_2, \ldots, \xi_n)$ from X^n distributed by P^n or Q^n, we meet with an intimate connection between (2) and asymptotic properties of both Bayes and Neyman—Pearson tests. For example, it will be shown that (2) holds iff (if and only if) P^∞ and Q^∞ are singular on $\mathscr{X}^\infty$ (in symbols $P^\infty \perp Q^\infty$), i.e. iff the hypotheses under consideration are asymptotically discernable with zero error (in the sense of the Bayes, Neyman—Pearson, or any other reasonable test). As to the rate of convergence in (2), it is connected with an asymptotic efficiency of the tests we have considered.

To link up the present paper with other works orientated in a similar direction let us mention that in the stationary case

$$P_1=P_2=\cdots=P, \quad Q_1=Q_2=\cdots=Q \tag{3}$$

the rate of convergence in (2) has been investigated by H. CHERNOFF [1]. He considered rather more general problem but, from our viewpoint, the following statement of his cited work is intrinsic: Let $\pi \in (0, 1)$ be an a priori probability of the hypothesis $H:P^n$ and put

$$e_\pi(P^n, Q^n) = \min_{E \in \mathscr{X}^n} [\pi P^n(E)+(1-\pi)Q^n(X^n-E)]. \tag{4}$$

---

* A former version of this paper was scheduled to be presented at the abortive Fifth Prague Conference on Information Theory, September 9—13, 1968.

Then

$$e_\pi(P^n, Q^n) = \lambda(P, Q)^{n+o(n)} \quad \text{(see (3))}, \tag{5}$$

where $\lambda(P, Q) \in [0, 1]$. The quantity $D(P, Q) = -\log \lambda(P, Q)$ has been called asymptotic efficiency of the Bayes test:

$$\text{„reject } H: P^n \text{ iff } (\xi_1, \xi_2, \dots, \xi_n) \in F_n\text{" *}, \tag{6}$$

where $F_n \in \mathscr{X}^n$ is the set that minimizes (4). Here $F_n = \{\pi p^n < (1-\pi) q^n\} \in \mathscr{X}^n$ for

$$p^n = \frac{dP^n}{d\mu^n}, \quad q^n = \frac{dQ^n}{d\mu^n}, \tag{7}$$

where μ^n is the product measure on $(X^n, \mathscr{X}^n)$ generated by the totally finite measure

$$\mu = \sum_{i=1}^{n} 2^{-i} \frac{P_i + Q_i}{P_i(X) + Q_i(X)} \quad \text{on } \mathscr{X}. \tag{8}$$

The measure μ or μ^n is uniformly dominating the family $\{P_i, Q_i\}$ or $\{P^n, Q^n\}$ respectively.

Remark that a decreasing influence of the a priori distribution upon test properties when the sample size tends to infinity has been observed already by R. von Mises [2]; this fact is conclusively illustrated by (5), where the right-hand asymptotic parameter $\lambda(P, Q)$ does not depend on $\pi \in (0, 1)$ **. Taking into account this together with the equality

$$e_{1/2}(P^n, Q^n) = \frac{1}{4}[2 - V(P^n, Q^n)] = \frac{1}{2} \int_{X^n} \min(p^n, q^n)\, d\mu^n \tag{9}$$

one can argue that the concept of asymptotic efficiency of the Bayes test (6) may be based merely on the rate of convergence in (2).

As to the statement (2) itself, the first who found a necessary and sufficient condition for its validity was S. Kakutani [3]. He proved that (2) holds iff

$$\prod_{i=1}^{\infty} H_{1/2}(P_i, Q_i) = 0, \tag{10}$$

where $H_{1/2}(P_i, Q_i)$ is the Hellinger's integral [4] of P_i, Q_i defined by

$$H_{1/2}(P_i, Q_i) = \int_X \sqrt{p_i q_i}\, d\mu, \tag{11}$$

where

$$p_i = \frac{dP_i}{d\mu} \quad q_i = \frac{dQ_i}{d\mu} \quad \text{(see (8))}. \tag{12}$$

* $e_\pi(P^n, Q^n)$ is the minimum probability of error (Bayes risk) related to the above described problem of testing H against K. It is attained by the Bayes test (6).

** The decreasing influence can be easily understood on the basis of inequality (22) proved below, in the special case of two simple hypotheses we have investigated.

The asymptotically errorless Bayes and Neyman—Pearson tests investigated in Sec. 3 below were characterized by T. NEMETZ [5] in terms of the Shannon's information contained in the sample $(\xi_1, \xi_2, \dots, \xi_n)$ concerning the hypotheses H, K.

The rate of convergence in (2), in the general case where (3) need not be necessarily true, has been studied previously in [6, 7].

From a formal point of view, the present paper is divided into two parts. In the first part (Sec. 2—4) probability measures P_i, Q_i, P^n, and Q^n are considered. Results of this part are extended in the second part (Sec. 5) to the more general case where the measures are not necessarily probabilistic.

2. Analitical properties of the α-entropy and its relation to the Bayes risk. Let P, Q be probability measures on a measurable space $(X, \mathscr{X})$ given by their Radon—Nikodym densities p, q with respect to a dominating measure μ. Denote by ξ the likehood ratio of P, Q, i.e. $\xi = p/q$. The quantity

$$(13) \qquad H_\alpha(P, Q) = E_Q \xi^\alpha = \int_X p^\alpha q^{1-\alpha} \, d\mu \qquad (\text{cf. } (11))$$

defined for all real α, will be called, in accordance with another authors, α-entropy of P with respect to Q. However, this terminology is rather inconsistent because it is usually required that every numerical measure of entropy (uncertainty) of P with respect to Q should be increasing when the "divergence" between P and Q is ncreasing. From this point of view

$$(14) \qquad \frac{1}{\alpha-1} \log H_\alpha(P, Q) \qquad \alpha \neq 1$$

or

$$(15) \qquad \operatorname{sign}(1-\alpha) H_\alpha(P, Q), \quad 1 - H_\alpha(P, Q)$$

are entropies but not $H_\alpha(P, Q)$, because, for $\alpha \in (0, 1)$, $H_\alpha(P, Q) = 0$ or 1 iff $P \perp Q$ or $P = Q$ respectively. The adoption of our terminology is motivated by the fact that, at one hand, it is desirable to emphasize the „entropic" aspects of $H_\alpha(P, Q)$ and, on the other hand, $H_\alpha(P, Q)$ itself is much more simple for analysis than any of the functionals in (14), (15).

A special version of the α-entropy for $\alpha = 1/2$ has been considered by E. HELLINGER [4] in his investigation of unitary invariants of self-adjoint operators in Hilbert space. Following H. HAHN [8], $H_{1/2}(P, Q)$ is called Hellinger's integral. From this point of view, the expression Hellinger's integral of order α should be suggested for (13). The quantity $-\log H_{1/2}(P, Q)$ was used by A. BHATTACHARYYA [9] in a statistical context; it is known in the literature also as Bhattacharyya-distance. H. CHERNOFF has considered $H_\alpha(P, Q)$ as the moment generating function of the likelihood-ratio logarithm $\log \xi$ (cf. (13)). A. RÉNYI introduced in [10] the quantity (14) called by him information of order α.

Let us denote by $I = I(P, Q)$ the set of all α such that the corresponding α-entropy is finite.

LEMMA 1. *$[0, 1] \subset I$ and for every $\alpha \in [0, 1]$*

$$0 \leqq H_\alpha(P, Q) \leqq 1 \tag{16}$$

where, if $\alpha \neq 0$, $\alpha \neq 1$, $H_\alpha(P, Q) = 0$ or 1 iff $P \perp Q$ or $P = Q$ respectively.

PROOF. The inclusion $[0, 1] \subset I$ follows from the Cauchy—Schwarz inequality. The remainder is clear.

LEMMA 2. *The set I is always an interval. If $I' \subset I$ is an arbitrary open subset then $H_\alpha(P, Q)$, as a function of α, is analytic in I' and*

$$\frac{d^k}{d\alpha^k} H_\alpha(P, Q) = \int_X p^\alpha q^{1-\alpha} \left(\log \frac{p}{q}\right)^k d\mu \quad \textit{everywhere on } I'. \tag{17}$$

PROOF. As noticed by L. H. KOOPMANS [11], $H_\alpha(P, Q)$ is the real restriction of

$$H(z) = \int_{-\infty}^{+\infty} \exp(zu)\, dF(u),$$

where $z = \alpha + \sqrt{-i}\beta$ is a complex number and $F(u) = Q(\{\log \xi \leqq u\})$. It follows from the theory of bilateral Laplace transform (see [12]) that all the properties stated in Lemma 2 are common fol all moment generating functions including $H_\alpha(P, Q)$.

LEMMA 3. *$H_\alpha(P, Q)$ is strictly convex in I unless $P \perp Q$ or $P = Q$.*

PROOF. This Lemma follows directly from Lemma 1 and (17) for $k = 2$.

LEMMA 4. *$[H_\alpha(P, Q)]^{1/\alpha}$ is a non-decreasing function of α for $\alpha \in [0, +\infty)$.*

PROOF. See M. LOÈVE [13], § 9. 3.

In this section, our attention will be paid mainly to clarify the relation between $e_\pi(P, Q)$ and $H_\alpha(P, Q)$ or $e_{1/2}(P, Q)$ respectively. It seems intuitively that, for any fixed $\pi, \alpha \in (0, 1)$, both these relations should be „monotone" in the sense that $e_\pi(P, Q)$ is increasing when $H_\alpha(P, Q)$ or $e_{1/2}(P, Q)$ is increasing (for a parametric system of distributions P, Q).

As to the relation between $e_\pi(P, Q)$ and $H_\alpha(P, Q)$, T. KAILATH [14] as the first found the inequality

$$\frac{1}{4}\left(1 - \sqrt{1 - H_{1/2}^2(P, Q)}\right) \leqq e_{1/2}(P, Q) \leqq \frac{1}{2} H_{1/2}(P, Q)$$

which yields

$$e_{1/2}(P, Q) \leqq \frac{1}{2} H_{1/2}(P, Q) \leqq 2\sqrt{e_{1/2}(P, Q)}.$$

The letter relation has been extended by T. NEMETZ [5] who proved

$$e_\pi(P, Q) \leqq \sqrt{\pi(1-\pi)}\, H_{1/2}(P, Q) \leqq \sqrt{\frac{e_\pi(P, Q)}{\pi(1-\pi)}}.$$

The left-hand inequality in this relation can be very easily extended into the form

$$e_\pi(P, Q) \leqq \pi^\alpha(1-\pi)^{1-\alpha} H_\alpha(P, Q), \qquad \alpha \in (0, 1),$$

noticing that $\varphi(u) \leqq u^\alpha(1-u)^{1-\alpha}$, $\alpha \in (0, 1)$, and then applying (20) and (21) below (see also (23) below). The right-hand inequality has been sharpened recently by A. PEREZ [15], who proved the following:

$$\pi^\alpha(1-\pi)^{1-\alpha} H_\alpha + \pi^{1-\alpha}(1-\pi)^\alpha H_{1-\alpha} \leqq e_\pi^\alpha(1-e_\pi)^{1-\alpha} + e_\pi^{1-\alpha}(1-e_\pi)^\alpha,$$

where

$$H_\alpha = H_\alpha(P, Q) \quad e_\pi = e_\pi(P, Q).$$

Our estimates of $e_\pi(P, Q)$ in terms of $H_\alpha(P, Q)$ or $e_{1/2}(P, Q)$ respectively will be based on the following preliminaries.

Let us mention that by $e_\pi(P, Q)$, in acordance with (4), we denote the quantity

$$e_\pi(P, Q) = \min_{E \in \mathscr{X}} [\pi P(E) + (1-\pi) Q(X-E)] = \int_X \min(\pi p, (1-\pi) q)\, d\mu$$

interpreted as the minimum probability of error related to the Bayes test of the simple hypothesis $H: P$ against the simple alternative $K: Q$ provided that a priori probabilities of H or K are π or $1-\pi$ respectively.

LEMMA 5. *If we put, for some* $\pi \in (0, 1)$, $\mu^* = \pi P + (1-\pi) Q$ *on* $\mathscr{X}$, $u = \pi dP/d\mu^*$ *on* X, *and* $\varphi(u) = \min(u, 1-u)$ *on* $[0, 1]$, *then*

$$0 \leqq u \leqq 1 \quad \textit{and} \quad \varphi(u) \in [0, 1/2] \textit{ on } X, \tag{18}$$

$$e_{1/2}(P, Q) = \int_X \min\left(\frac{u}{2\pi}, \frac{1-u}{2(1-\pi)}\right) d\mu^* \quad \textit{for every} \quad \pi \in (0, 1), \tag{19}$$

$$H_\alpha(P, Q) = \int_X \left(\frac{u}{\pi}\right)^\alpha \left(\frac{1-u}{1-\pi}\right)^{1-\alpha} d\mu^* \quad \textit{for every} \quad \alpha, \pi \in (0, 1), \tag{20}$$

and

$$e_\pi(P, Q) = \int_X \varphi(u)\, d\mu^*. \tag{21}$$

PROOF. Clear.

THEOREM 1. *For every* $\alpha, \pi \in (0, 1)$,

$$2\min(\pi, 1-\pi) e_{1/2}(P, Q) \leqq e_\pi(P, Q) \leqq 2\max(\pi, 1-\pi) e_{1/2}(P, Q) \tag{22}$$

and

$$e_\pi(P, Q) \leqq \pi^\alpha(1-\pi)^{1-\alpha} H_\alpha(P, Q) \leqq e_\pi(P, Q)^{\varphi(\alpha)}(1-e_\pi(P, Q))^{1-\varphi(\alpha)}. \tag{23}$$

PROOF. Denote

$$\psi(u) = \min\left(\frac{u}{2\pi}, \frac{1-u}{2(1-\pi)}\right),^* \qquad u \in [0, 1].$$

* Throughout the proof we suppose that $\alpha, \pi \in (0,1)$ are arbitrary fixed. Relations (22) and (23) obviously hold also for $\pi = 0$ or 1 and the left-hand inequality in (23) for $\alpha = 0$ or 1.

If $\psi(u) = u/2\pi$, then $\psi(u) \in [0, 1/2]$ for $u \in [0, 1]$ and $\varphi(u) = \varphi(2\pi\psi(u))$, where

$$(24) \qquad 2\psi\varphi(\pi) \leqq \varphi(2\pi\psi) \leqq 2\pi\psi$$

for every $\psi \in [0, 1/2]$. If $\psi(u) = (1-u)/2(1-\pi)$, then $\psi(u) \in [0, 1/2]$ for $u \in [0, 1]$ and $\varphi(u) = \varphi(2(1-\pi)\psi(u))$, where

$$(25) \qquad 2\psi\varphi(\pi) = 2\psi\varphi(1-\pi) \leqq \varphi(2(1-\pi)\psi) \leqq 2(1-\pi)\psi$$

for every $\psi \in [0, 1/2]$. Integrating (24) over the set $A = \{\psi(u) = u/2\pi\} \in \mathcal{X}$ and (25) over $B = \{\psi(u) = (1-u)/2(1-\pi)\} = X - A$, we obtain

$$2\varphi(\pi)\int\limits_X \psi(u)\,d\mu^* \leqq e_\pi(P, Q) \leqq 2\pi\int\limits_A \psi(u)\,d\mu^* + 2(1-\pi)\int\limits_B \psi(u)\,d\mu^*$$

These inequalities yield (22).

To prove (23) let us notice first that

$$(26) \qquad \pi^\alpha(1-\pi)^{1-\alpha}H_\alpha(P, Q) = \int\limits_A \varphi(u)^\alpha(1-\varphi(u))^{1-\alpha}\,d\mu^* + \int\limits_B (1-\varphi(u))^\alpha\varphi(u)^{1-\alpha}\,d\mu^*,$$

where $A = \{u = \varphi(u)\} \in \mathcal{X}$, $B = \{u = 1-\varphi(u)\} = X - A$. Since

$$\max\,[u^{1-\alpha}(1-u)^\alpha, u^\alpha(1-u)^{1-\alpha}] = u^{\varphi(\alpha)}(1-u)^{1-\varphi(\alpha)} \quad \text{for every} \quad u \in [0, 1/2],$$

(18) and (26) imply

$$(27) \qquad \pi^\alpha(1-\pi)^{1-\alpha}H_\alpha(P, Q) \leqq \int\limits_X \varphi(u)^{\varphi(\alpha)}(1-\varphi(u))^{1-\varphi(\alpha)}\,d\mu^*.$$

On the other hand, both $u^{1-\alpha}(1-u)^\alpha$ and $u^\alpha(1-u)^{1-\alpha}$ are concave functions of u in the interval $u \in [0, 1/2]$ which take on values 0 and 1 at the end-points of this interval so that

$$u \leqq \min\,[u^{1-\alpha}(1-u)^\alpha, u^\alpha(1-u)^{1-\alpha}] \quad \text{for every} \quad u \in [0, 1/2].$$

Hence, by (21) and (26), the left-hand inequality in (23) holds. Applying the Jensen's inequality to the concave function (of u) $u^{\varphi(\alpha)}(1-u)^{1-\varphi(\alpha)}$ and using (21) and (27) we obtain the right-hand inequality in (23).

It is to be emphasized that the technique of a convex envelope used in the proof above can be applied to obtain upper and lower estimates of integrals with integrands being concave or convex functions of a random variable ξ in terms of another integrals of the same type. It seems that analogical techniques might be applicable not only in the statistics or information theory, where various convex or concave functionals play an important role, but whereever such functionals occur.

Before formulating an important corollary of Th. 1 we shall adopt a notational convention. Throughout all this paper P_i, Q_i, $i = 1, 2, \ldots$ is supposed to be an arbitrary fixed sequence of measures. Therefore, in the sequel we shall be allowed to use a more simple notation H_α^i instead of $H_\alpha(P_i, Q_i)$.

Corollary. For every $0<\alpha\leqq\alpha(i)\leqq 1$ and $0\leqq\beta(i)\leqq\beta<1$,

$$\frac{1}{4}\prod_{i=1}^{n}(H^i_{\beta(i)})^{2\beta/\beta(i)}\leqq e_{1/2}(P^n,Q^n)\leqq\frac{1}{2}\prod_{i=1}^{n}(H^i_{\alpha(i)})^{\alpha/\alpha(i)}. \tag{28}$$

If

$$\alpha^*=\sup_{i=1,2,\dots}\alpha(i),\quad \beta_*=\inf_{i=1,2,\dots}\beta(i),$$

then

$$\frac{1}{4}\left(\prod_{i=1}^{n}H^i_{\beta(i)}\right)^{2\beta/\beta_*}\leqq e_{1/2}(P^n,Q^n)\leqq\frac{1}{2}\left(\prod_{i=1}^{n}H^i_{\alpha(i)}\right)^{\alpha/\alpha^*}. \tag{29}$$

PROOF. Since

$$p^n=\prod_{i=1}^{n}p_i,\quad q^n=\prod_{i=1}^{n}q_i\qquad \text{(cf. (7) and (12))}, \tag{30}$$

we can write for any real α

$$H_\alpha(P^n,Q^n)=\prod_{i=1}^{n}H^i_\alpha. \tag{31}$$

Relation (31) together with (23) imply

$$\frac{1}{4}\left(\prod_{i=1}^{n}H^i_\beta\right)^2\leqq e_{1/2}(P^n,Q^n)\leqq\frac{1}{2}\prod_{i=1}^{n}H^i_\alpha.$$

This together with Lemma 4 yields (28). Inequalities (29) follow from the fact that H^u is a non-increasing function of u if $H\in[0,1]$.

3. Limit theorems for total variation. In Sec. 1 the Bayes test was defined by (6), where $\pi\in(0,1)$ is implicitely figuring as a parameter. It was quantitatively characterized by the risk $e_\pi(P^n,Q^n)$ which can be interpreted as an average probability of error.

Define now what we mean by the Neyman—Pearson test. By the Neyman—Pearson lemma, for every $\beta\in(0,1)$ there exists a set $F_n(\beta)\in\mathscr{X}^n$ such that

$$Q^n(X^n-F_n(\beta))=\inf_{\substack{E\in\mathscr{X}^n\\ P^n(E)\leqq\beta}}Q^n(X^n-E). \tag{32}$$

The Neyman—Pearson test we define as follows:

$$\text{“reject}\quad H:P^n\quad\text{iff}\quad(\xi_1,\xi_2,\dots,\xi_n)\in F_n(\beta).\text{”} \tag{33}$$

Quantities β and $\alpha_n(\beta)=Q^n(X^n-F_n(\beta))$, called probabilities of the second and first kind, serve as quantitative characteristics of the test (33). Obviously,

$$e_\pi(P^n,Q^n)\leqq\pi\alpha_n(\beta)+(1-\pi)\beta. \tag{34}$$

We shall say that the Bayes test is asymptotically errorless (with respect to the hypotheses $H:P^n$, $K:Q^n$) if, for every $\pi\in(0,1)$,

$$\lim_n e_\pi(P^n,Q^n)=0\qquad\text{(cf. (22))}. \tag{35}$$

Analogically, we shall say that the Neyman—Pearson test is asymptotically errorless (with respect to $H: P^n$, $K: Q^n$) if for every $\beta \in (0, 1)$

(36) $$\lim_n \alpha_n(\beta) = 0.$$

It is to see that the asymptotical errorlessness, as we have introduced it, is a property of the measures P^∞, Q^∞. Th. 2 below implies that it takes place iff the infinite vector of observations $(\xi_1, \xi_2, \ldots)$ contains „full information" concerning the hypotheses under consideration, in the sense of T. NEMETZ [5].

THEOREM 2. *Relations* (35) *or* (36) *hold for every* $\pi \in (0, 1)$ *or* $\beta \in (0, 1)$ *respectively iff* (2) *holds, i.e. iff*

(37) $$\lim_n e_{1/2}(P^n, Q^n) = 0.$$

The latter condition is equivalent to $P^\infty \perp Q^\infty$ *.

PROOF. It follows from Th. 1 and (31) that (37) holds iff

$$\lim_n \prod_{i=1}^n H_\alpha^i = 0,$$

i.e. iff

(38) $$H_\alpha(P^\infty, Q^\infty) = \prod_{i=1}^\infty H_\alpha^i = 0 \;{**}$$

for some $\alpha \in (0, 1)$. Hence, by Lemma 1, we can argue that (38) holds iff $P^\infty \perp Q^\infty$. Thus, the statements (35), (37), and $P^\infty \perp Q^\infty$ are mutually equivalent. Inequality (30) implies that (36) holds for every $\beta \in (0, 1)$ only if $P^\infty \perp Q^\infty$. The "if" part is also simple: if $\lim_n e_{1/2}(P^n, Q^n) = 0$ then, for sufficiently large n, $P^n(F_n) < \beta$ for any fixed $\beta \in (0, 1)$ so that $\alpha_n(\beta) \leqq Q^n(X^n - F_n) \leqq 2e_{1/2}(P^n, Q^n)$ (see (4) and (6)).

Relation (38) enables us to express the conditions under which $P^\infty \perp Q^\infty$ holds in terms of more easily evaluable functionals H_α^i. Exactly this is a sense of the KAKUTANI's result mentioned in Sec. 1. The next theorem goes deeper in this direction.

THEOREM 3. *If* $P^\infty \perp Q^\infty$, *then, for every* $\alpha_i \in (0, 1)$ *such that*

(39) $$\alpha_* = \liminf_i \alpha_i > \alpha, \quad \alpha^* = \limsup_i \alpha_i < \beta$$

for some $\alpha, \beta \in (0, 1)$, *the following relation holds*

(40) $$\prod_{i=1}^\infty H_{\alpha_i}^i = 0.$$

---

* cf., for example, Th. 4 in [5] or assertion (ii) and (2.12) in [6].

** This relation need not be clear at the first sight. However, it can be easily established using the concavity of the intergrand in (13), for $\alpha \in (0,1)$, and then applying the well known semimartingale argument.

Conversely, if there exist $\alpha_i \in (0, 1)$ *satisfying* (39) *and* (40), *then* $P^\infty \perp Q^\infty$. *Relation* (40) *is equivalent to*

$$\sum_{i=1}^{\infty} (1 - H^i_{\alpha_i}) = +\infty \tag{41}$$

unless $H^i_\alpha = 0$ *for some* i *and* $\alpha \in (0, 1)$ *(i.e. unless* P_i, Q_i *are singular for some* i*).*

PROOF. If we define

$$\alpha(i) = \begin{cases} \alpha_i & \text{for } \alpha_i \geqq \alpha \\ \alpha & \text{for } \alpha_i < \alpha \end{cases} \quad \text{and} \quad \beta(i) = \begin{cases} \alpha_i & \text{for } \alpha_i \leqq \beta \\ \beta & \text{for } \alpha_i > \beta \end{cases}$$

then the conditions of the Corollary above are satisfied so that $P^\infty \perp Q^\infty$ implies

$$\lim_n \prod_{i=1}^{n} H^i_{\beta(i)} = 0 \qquad (\text{cf. (37) and (29) with } \beta_* = \alpha_*).$$

It follows from our construction of the numbers $\beta(i)$ that k exist such that $\beta(i) = \alpha_i$ for every $i \geqq k$, i.e.

$$\prod_{i=1}^{k-1} H^i_{\beta(i)} \lim_n \prod_{i=k+1}^{n} H^i_{\alpha_i} = 0. \tag{42}$$

If $H^i_{\beta(i)} = 0$ for some $1 \leqq i \leqq k-1$ then, by Lemma 1, $H^i_\alpha = 0$ for every $\alpha \in (0, 1)$ and, consequently, (42) implies (40).

Analogous considerations as those used above yield that (40) implies

$$\prod_{i=1}^{\infty} H^i_{\alpha(i)} = 0$$

and it remains to apply the right-hand inequality in (29) to obtain (37). The equivalence of (40) and (41) easily follows from the following inequalities: $1 - u \leqq \exp(-u)$, $1 - u \geqq \exp\left(-u/(1-u)\right)$.

The numbers $H^1_\alpha, H^2_\alpha, \ldots$ are not the simplest of all the "distance" measures which can be defined on the pairs $(P_1, Q_1), (P_2, Q_2), \ldots$. More simple seem to be the following ones

$$\Delta_i = \frac{1}{2} V(P_i, Q_i) = 1 - 2e_{1/2}(P_i, Q_i) \in [0, 1] \qquad (\text{cf. [6]}), \tag{43}$$

which may be called variation numbers corresponding to P_i, Q_i. A simplicity viewpoint is the goal point in favour of the following considerations, where necessary and sufficient conditions for the singularity $P^\infty \perp Q^\infty$ (as well as for a rate of convergence in (2)) in terms of $\Delta_1, \Delta_2, \ldots$ are given.

THEOREM 4. *If* $P^\infty \perp Q^\infty$, *then*

$$\prod_{i=1}^{\infty} (1 - \Delta_i) = 0 \tag{44}$$

which is equivalent to

$$\sum_{i=1}^{\infty} \Delta_i = +\infty \tag{45}$$

unless $\Delta_i=1$ *for some (i.e. unless* $P_i \perp Q_i$ *for some* i*). Conversely, if*

$$\sum_{i=1}^{\infty} \Delta_i^2 = +\infty, \tag{46}$$

then $P^\infty \perp Q^\infty$*. The condition*

$$\sum_{i=1}^{n} \Delta_i = O(n) \tag{47}$$

implies (46).

PROOF. It follows from Th. 1 that

$$2e_{1/2}(P_i, Q_i) \leqq H_{1/2}^i \leqq \sqrt{4e_{1/2}(P_i, Q_i)(1-e_{1/2}(P_i, Q_i))} \tag{48}$$

so that, by (43),

$$1-\Delta_i \leqq H_{1/2}^i \leqq \sqrt{1-\Delta_i^2} \leqq \exp\left(-\frac{1}{2}\Delta_i^2\right)$$

or

$$\prod_{i=1}^{n} (1-\Delta_i) \leqq \prod_{i=1}^{n} H_{1/2}^i \leqq \exp\left(-\frac{1}{2}\sum_{i=1}^{n} \Delta_i^2\right). \tag{49}$$

This together with Th. 3 yields the desired assertions concerning (44) and (46). As to the equivalence of (44) and (45), we may refer to the inequalities at the end of the proof of Th. 3 above. To prove that (47) implies (46) let us remark that if (46) is not satisfied, then $\lim_i \Delta_i = 0$. According to a well known theorem of analysis, this implies

$$\lim_n \frac{1}{n}\sum_{i=1}^{n} \Delta_i = 0 \qquad \text{Q.E.D.}$$

Let us remark that (44) under $P^\infty \perp Q^\infty$ can be drawn without any reference to Th. 1 or Th. 3. If we denote, for $m<n$,

$$p^{(n,m)} = \prod_{i=m+1}^{n} p_i, \quad q^{(n,m)} = \prod_{i=m+1}^{n} q_i$$

then, obviously,

$$\min(p^n, q^n) \geqq \min(p^{(n,m)}, q^{(n,m)}) \min(p^m, q^m)$$

and, consequently (cf. (9)),

$$2e_{1/2}(P^n, Q^n) \geqq 2e_{1/2}\left(\prod_{i=m+1}^{n} P_i, \prod_{i=m+1}^{n} Q_i\right) 2e_{1/2}(P^m, Q^m) \tag{50}$$

for every $1 \leqq m < n$ (this inequality has been first established in [16]). From this inequality and (43) we obtain

$$2e_{1/2}(P^n, Q^n) \geqq \prod_{i=1}^{n} (1-\Delta_i)$$

and it remains to apply Th. 2.

Let us remark that neither (45) nor (46) is equivalent to $P^\infty \perp Q^\infty$. For this purpose let us consider a sequence $\delta_i \in [0, 1)$ and define discrete distributions P_i, Q_i by

$$P_i = \left(\frac{1+\delta_i}{2}, \frac{1-\delta_i}{2}\right), \quad Q_i = \left(\frac{1-\delta_i}{2}, \frac{1-\delta_i}{2}\right).$$

It is easy to see that in this case $\Delta_i = \delta_i$, $H^i_{1/2} = \sqrt{1-\delta_i^2}$, so that $P^\infty \perp Q^\infty$ iff (46) holds. If we consider another example, namely,

$$P_i = (0, 1), \quad Q_i = (\delta_i, 1-\delta_i)$$

hen we obtain $\Delta_i = \delta_i$, $H^i_{1/2} = \sqrt{1-\delta_i}$ so that $P^\infty \perp Q^\infty$ iff (47) holds.

4. Rate of convergence. The aim of this section will be to generalize the statement (5) of H. CHERNOFF to the non-stationary case where (3) need not be necessarily true. As we said in Sec. 1, it will suffice to restrict ourselves to $\pi = 1/2$, i.e. to investigate the rate of convergence in (2) or (37) only. We shall say that a sequence of numbers e_n converges exponentially to $e \in [0, +\infty]$ if there exists $\lambda \in [0, 1)$ or $\lambda \in (1, +\infty)$ such that, for all sufficiently large n, $|e_n - e| \leqq \lambda^n$ or $e_n \geqq \lambda^n$ depending on whether $e \in [0, +\infty)$ or $e = +\infty$ respectively (in symbols, $\lim_n e_n = e(\text{Exp})$). According to (5) and (22), in what follows we shall restrict ourselves to a study of exponential rate of convergence of $e_{1/2}(P^n, Q^n)$ to zero. All results concerning $e_{1/2}(P^n, Q^n)$ can be directly extended to $e_\pi(P^n, Q^n)$ by (22).

THEOREM 5. *The statement*

$$\lim_n e_{1/2}(P^n, Q^n) = 0 \text{ (Exp.)} \tag{51}$$

holds iff, for $\alpha_i \in (0, 1)$ *defined in* Th. 3,

$$\sum_{i=1}^{n} -\log H^i_{\alpha_i} = O(n).$$

PROOF. As it was said in the proof of Th. 3, the following inequalities hold

$$\frac{1}{4}\left(\prod_{i=1}^{n} H^i_{\beta(i)}\right)^{2\beta/\alpha_*} \leqq e_{1/2}(P^n, Q^n) \leqq \frac{1}{2}\left(\prod_{i=1}^{n} H^i_{\alpha(i)}\right)^{\alpha/\alpha^*}$$

for α, β, α_*, α^* defined by (39). The remainder is now clear.

THEOREM 6. *If* (51) *holds, then*

$$\sum_{i=1}^{n} -\log(1-\Delta_i) = O(n) \tag{52}$$

Conversely, if

$$\sum_{i=1}^{n} \Delta_i = O(n) \tag{53}$$

then (51) *holds. For* $0 \leqq \Delta_i \leqq \Delta < 1$, $i = 1, 2, \ldots$ (51) *holds iff* (53) *holds.*

PROOF. The first assertion of this theorem follows from Th. 5 and (49). If $\Delta_i \leqq \leqq \Delta < 1$, then (52) implies (53), so that it remains to prove the second assertion of the theorem. This assertion has been proved by the author in [7]. However, for the sake of completeness, we shall present the proof here. If we shall prove that there exist $E_n, \tilde{E}_n \in \mathscr{X}^n$ and $\lambda \in (0, 1)$ such that $X^n - \tilde{E}_n \subset E_n$ and

$$P^n(\tilde{E}_n) \leqq \lambda^n \tag{54}$$

$$Q^n(E_n) \leqq \lambda^n \tag{55}$$

for every $n \geqq n_0$, the desired assertion will be proved as well, because, in view of (4), we can succesively write

$$e_{1/2}(P^n, Q^n) \leqq \frac{1}{2}[P^n(\tilde{E}_n) + Q^n(X^n - \tilde{E}_n)] \leqq \frac{1}{2}[P^n(\tilde{E}_n) + Q^n(E_n)].$$

It follows from the definition of Δ_i in (43) that (cf. (9)) $\Delta_i = P_i(G_i) - Q_i(G_i)$ for some $G_i \in \mathscr{X}$. Therefore (53) implies that there exists $\delta \in (0, 1)$ and n_0 such that for every $n > n_0$

$$\sum_{i=1}^{n} P_i(G_i) - n\delta > \sum_{i=1}^{n} Q_i(G_i) + n\delta.$$

Now, if we define $\zeta_i = \eta_i - Q_i(G_i)$ where $\eta_i = 1$ or 0 depending on whether $x \in G_i$ or $x \notin G_i$ respectively, and if we put

$$E_n = \left\{\sum_{i=1}^{n} \zeta_i > n\delta\right\} \in \mathscr{X}^n$$

$$\tilde{E}_n = \left\{\sum_{i=1}^{n} [\zeta_i + Q_i(G_i) - P_i(G_i)] \leqq -n\delta\right\} \in \mathscr{X}^n,$$

then it is to see that the inclusion $X^n - \tilde{E}_n \subset E_n$ holds for every $n \geqq n_0$. We shall prove that the set E_n satisfies (55) for $\lambda = \exp(-\delta^2/4)$ and $n \geqq n_0$. Inequality (54) can be proved by the same manner.

Let us notice first that

$$\prod_{i=1}^{n} \int_X \exp(\delta\zeta_i)\, dQ_i = \int_{X^n} \exp\left(\delta \sum_{i=1}^{n} \zeta_i\right) dQ^n \geqq \exp(\delta_n^2) Q^n(E_n)$$

so that

$$-\frac{1}{n}\log Q^n(E_n) \geqq \delta^2 - \frac{1}{n}\sum_{i=1}^{n} \log \int_X \exp(\delta\zeta_i)\, dQ_i. \tag{56}$$

Since $|\zeta_i| \leqq 1$ and $\alpha \in (0, 1)$ we can write

$$\int_X \exp(\delta\zeta_i)\, dQ_i = \sum_{k=0}^{\infty} \frac{\delta^k}{k!} \int_X \zeta_i^k\, dQ_i,$$

where

$$\delta^k \int_X \zeta_i^k\, dQ_i \begin{cases} = 0 & \text{if } k = 1 \\ \leqq \delta^2 & \text{if } k > 1. \end{cases}$$

Hence

$$\int_X \exp(\delta\zeta_i)\,dQ_i \leqq 1+\delta^2 \sum_{k=2}^{\infty} \frac{1}{k!} = 1+\delta^2(\exp(1)-2).$$

Since $\exp(1) < 2+3/4$ and $-\log(1+u) \geqq -u$, we can write

$$-\log \int_X \exp(\delta\zeta_i)\,dQ_i \geqq -\frac{3}{4}\delta^2$$

which together with (56) yields the desired result.

Let us remark that if $\lim\inf_i \Delta_i = 1$ then (52) does not imply (53). Indeed, let us define

$$\Delta_i = \begin{cases} 1-\exp(-2^k) & \text{if } i=2^k \text{ for some natural } k \\ 0 & \text{in the opposite case.} \end{cases}$$

It is easily seen that in this case

$$\sum_{i=1}^{n} -\log(1-\Delta_i) \geqq \frac{n}{2}$$

whereas

$$\sum_{i=1}^{n} \Delta_i \leqq \frac{\log n}{\log 2}.$$

THEOREM 7. *If*

$$e_{1/2}\left(\prod_{i=m+1}^{m+k} P_i, \prod_{i=m+1}^{m+k} Q_i\right) \geqq e_{1/2}(P^k, Q^k) \tag{57}$$

for every natural m and k, then there exists $\lambda \in [0, 1]$ such that

$$e_{1/2}(P^n, Q^n) = \lambda^{n+o(n)}. \tag{58}$$

PROOF. If we put $e_n = e_{1/2}(P^n, Q^n)$ then the relation

$$\lim_n -\frac{1}{n} \log e_n = \lambda \in [0, 1] \tag{59}$$

is to be proved. But (57) together with (50) imply that $2e_{m+k} \geqq 2e_k 2e_m$ for every natural m, k. By a well known lemma of analysis, this inequality guarantees that the limit in (59) exists Q.E.D.

Th. 7 which has been first stated in [16] can be considered as a generalization of the result (5). Indeed, if the stationarity conditions (3) hold, then (57) holds as well, and (5) can be deduced from Th. 7. We stated this very simply provable theorem in this paper for the sake of completeness. It is to be noted here, however, that in the cited paper of H. Chernoff where (5) was proved a deeper result was established. In this paper not only (5) but also the following explicite formula for $\lambda = \lambda(P, Q)$ was given:

$$\lambda(P, Q) = \inf_{\alpha \in (0, 1)} H_\alpha(P, Q) = \inf_{\alpha \in (0, 1)} H_\alpha(Q, P). \tag{60}$$

Before to bring into an end our investigation of the statistical model where all measures under consideration are probability measures, let us apply the results we have obtained to a concrete example. The urn model which is described below was communicated to the author by A. RÉNYI.

Let an urn contain r_0 red and w_0 white balls. Suppose that we draw a ball from the urn, define

$$\xi_0 = \begin{cases} 1 & \text{if the ball is red} \\ 0 & \text{if the ball is white,} \end{cases}$$

put it back, and add, independently of the ball drawn colour, r_1 red and w_1 white balls. After mixing the balls we draw again a ball, define

$$\xi_1 = \begin{cases} 1 & \text{if it is red} \\ 0 & \text{if it is white} \end{cases}$$

and put it back together with r_2 red and w_2 white new balls. Continuing this process, after the i-th step

$$R_i = \sum_{j=1}^{i} r_j$$

red and

$$W_i = \sum_{j=1}^{i} w_j$$

white balls are added so that the total number of balls in the urn is

$$T_i = \sum_{j=0}^{i} r_j + \sum_{j=0}^{i} w_j .$$

Suppose now that the hypothesis $H: r_0 = a, w_0 = b$ is tested against $K: r_0 = b, w_0 = a$, where $a, b > 0$ are arbitrary integers such that $r_0 + w_0 = a + b$, $a \neq b$. It is supposed that both the hypotheses are of a non-zero a priori probability and that they are tested on the basis of information contained in the vector $(\xi_1, \xi_2, \dots, \xi_n)$. This vector is distributed by P^n or Q^n defined in (1) for

$$P_i = \left(\frac{a+R_i}{T_i}, \frac{b+W_i}{T_i}\right), \quad Q_i = \left(\frac{b+R_i}{T_i}, \frac{a+W_i}{T_i}\right).$$

In this case we have

$$\Delta_i = \frac{|a-b|}{T_i}, \quad H_{1/2}^i = \frac{1}{T_i}\left[\sqrt{(a+R_i)(b+R_i)} + \sqrt{(a+W_i)(b+W_i)}\right].$$

Thus, using Th. 4, we find that the origin composition of the urn can be determined with probability 1 only if or if

$$\sum_{i=1}^{\infty} \frac{1}{T_i} = +\infty \tag{61}$$

or

$$\sum_{i=1}^{\infty} \frac{1}{T_i^2} = +\infty \tag{62}$$

respectively. (Let us notice that here no role plays the fact which is the proportion of red and white balls we are adding to the urn, i.e. which is the ratio $r_i/(r_i + w_i)$, $i = 1, 2, \dots$.) If, for example, $r_i/(r_i + w_i) = 1/2$, $i = 1, 2, \dots$, i.e. if we are adding the same number of balls of both colours, then

$$H^i_{1/2} = \frac{2}{T_i}\sqrt{(a+R_i)(b+R_i)},$$

where $T_i = 2R_i+a+b$ so that

$$H^i_{1/2} = \sqrt{\left(1+\frac{a-b}{T_i}\right)\left(1-\frac{a-b}{T_i}\right)} = 1-\frac{(a-b)^2}{T_i^2}+O\left(\frac{1}{T_i^2}\right).$$

Hence, by Th. 3, we can determine the original composition of the urn with probability 1 iff (62) holds. In another case where $r_i/(r_i+w_i) = 1$ we get

$$H^i_{1/2} = \sqrt{\left(1-\frac{a}{T_i}\right)\left(1-\frac{b}{T_i}\right)}+\frac{\sqrt{ab}}{T_i} = 1-\frac{(\sqrt{a}-\sqrt{b})}{2T_i}+O\left(\frac{1}{T_i^2}\right)$$

so that the original composition of the urn can be determined iff (61) holds. From Th. 5 and (22) we get that for any $\pi\in(0,1)$ the Bayes error $e_\pi(P^n, Q^n)$ converges exponentially to zero iff

$$\sum_{i=1}^{n}\frac{1}{T_i} = O(n) \tag{63}$$

(as we stated in Th. 4, (63) implies (62)). Since T_i is non-decreasing, (63) holds iff $\lim_i T_i < +\infty$. If $\lim_i T_i = T < +\infty$, $\lim_i R_i = R$, and $\lim_i W_i = W$, $T = R+W+ +a+b$, then we might obtain $e_\pi(P^n, Q^n) = \lambda^{n+o(n)}$, where $\lambda\in(0,1)$ is defined by

$$\lambda = \frac{1}{T}[(a+R)^\alpha(b+R)^{1-\alpha}+(b+W)^\alpha(a+W)^{1-\alpha}]$$

for a suitable $\alpha\in(0,1)$. If, for example, $R=W$ (i.e., in particular, if $r_i/(r_i+w_i) = 1/2$ for every i such that $r_i+w_i > 0$), then

$$\lambda = \sqrt{1-\left(\frac{a-b}{T}\right)^2}.$$

5. General case. In this section we shall suppose that P_i, Q_i defining P^n, Q^n by (1) are arbitrary totally finite measures on $(X, \mathscr{X})$. In this case μ defined by (8) is still a totally finite measure uniformly dominating P_i, Q_i so that the densities in (12) and (7) exist. Hence all the concepts of preceeding sections based on the densities can be extended to this more general case. In particular, the α-entropies $H_\alpha(P^n, Q^n)$ and $H^i_\alpha = H_\alpha(P_i, Q_i)$ can be defined by (13) and (31) holds.

The aim of the present section is to investigate the limit property (2) of total variation $V(P^n, Q^n)$. In accordance with the notation employed above, for the sake of simplicity we shall write

$$e(P^n, Q^n) = \frac{1}{4}[P^n(X^n)+Q^n(X^n)-V(P^n, Q^n)] \quad \text{(cf. (9))},$$

where

$$e(P^n, Q^n) = \frac{1}{2}\int_{X^n}\min(p^n, q^n)\,d\mu^n. \tag{64}$$

Our exposition will be very concise and we will restrict ourselves to illustrate by several simple examples that the basic results of the preceding sections concerning the relation

$$\lim_n e(P^n, Q^n) = 0 \tag{65}$$

remain true when we pass to this more general situation. After all, this conclusion as well as the fact that the methods employed above remain applicable is in no way surprising.

The basic fact we shall utilize is the following analogue of Th. 1:

(66)

$$2e(P^n, Q^n) \leqq H_\alpha(P^n, Q^n) \leqq (2R_n)^{\max(\alpha, 1-\alpha)} e(P^n, Q^n)^{\min(\alpha, 1-\alpha)} \quad \text{for every} \quad \alpha \in [0, 1],$$

where R_n stands for $P^n(X^n)+Q^n(X^n)$. The left-hand estimate is obvious, it follows from the inequality $\min(p, q) \leqq p^\alpha q^{1-\alpha}$, $\alpha \in (0, 1)$. To prove the right-hand inequality let us notice first that

$$e(P^n, Q^n) = \frac{R_n}{2} \int_{X^n} \min(u, 1-u)\, d\mu^*, \quad H_\alpha(P, Q) = R_n \int_{X^n} u^\alpha (1-u)^{1-\alpha}\, d\mu^*,$$

where μ^* is a probability measure,

$$\mu^* = \frac{P^n + Q^n}{R_n} \quad \text{on} \quad \mathscr{X}^n,$$

and $u = p^n/(p^n+q^n)$ on X. Now it remains to apply the inequality

$$u^\alpha(1-u)^{1-\alpha} \leqq [\min(u, 1-u)]^{\min(\alpha, 1-\alpha)}$$

holding for all $u \in [0, 1]$ and then the inequality of Jensen:

$$\int_{X^n} [\min(u, 1-u)]^{\min(\alpha, 1-\alpha)}\, d\mu^* \leqq \left[\int_{X^n} \min(u, 1-u)\, d\mu\right]^{\min(\alpha, 1-\alpha)}.$$

Inequality (66) together with (31) gives the following

THEOREM 8. *If* $\limsup_n R_n < +\infty$ *then* (65) *holds iff*

$$\sum_{i=1}^{\infty} -\log H_\alpha^i = +\infty \quad \textit{for some} \quad \alpha \in (0, 1), \tag{67}$$

and $\lim_n e(P^n, Q^n) = 0$ (Exp.) *iff*

$$\sum_{i=1}^{n} -\log H_\alpha^i = O(n) \quad \textit{for some} \quad \alpha \in (0, 1).$$

If $\lim_n e(P^n, Q^n) = 0$ (Exp.), then it is desirable to know something concerning a parameter of this exponential convergence. In this respect the following analogue holds.

THEOREM 9. *If*

$$e\left(\prod_{i=m+1}^{m+k} P_i, \prod_{i=m+1}^{m+k} Q_i\right) \geqq e(P^k, Q^k)$$

for every natural m, k, *then there exists* $\lambda \geqq 0$ *such that*

$$e(P^n, Q^n) = \lambda^{n+o(n)}.$$

An easy modification of a method used by H. CHERNOFF in [1] shows that in the stationary case (3), λ can be evaluated by (60). However, in general case we have considered here, λ need not take on values from [0, 1] only.

Next we shall illustrate by an example that (67) need not be implied by (65) if $\limsup_n R_n = +\infty$. Let $X=[0, 1]$, let $\mathscr{X}$ be the σ-algebra of all Borel measurable subsets of X, and let μ be the Lebesgue measure on $\mathscr{X}$. If we put

$$P_i(E) = i\mu(E), \quad Q_i(E) = \frac{1}{i}\mu(E)$$

for every $E\in\mathscr{X}$, then $H_\alpha(P^n, Q^n)=H_\alpha(P_i, Q_i)=1$ for $\alpha=1/2$ whereas $e(P^n, Q^n) = 1/n!$ so that $\lim_n e(P^n, Q^n)=0$. In this case, however, $R_n = n!+1/n!$ so that $\limsup_n R_n = +\infty$.

REFERENCES

[1] CHERNOFF H.: A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations, *Ann. Math. Stat.*, **23** (1952), 493—507.

[2] VON MISES, R.: Fundamentalsätze der Wahrscheinlichkeitsrechnung, *Math. Zeitschr.*, **4** (1919), 1—97.

[3] KAKUTANI, S.: On equivalence of infinite product measures, *Ann. of Math.*, **49** (1948), 214—226.

[4] HELLINGER, E.: Neue Begründung der Theorie quadratischen Formen von unendlichvielen Veränderlichen, *Journ. für die reine und angew. Math.*, **136** (1909), 210—271.

[5] NEMETZ, T.: *Information theory and the testing of a Hypothesis,* Proc. of Coll. on Inf. Theory, Debrecen, II, Budapest 1969.

[6] VAJDA, I.: Rate of convergence of the information in a sample concerning a parameter. *Czechosl. Math. J.*, **17** (1967), 225—231.

[7] VAJDA, I.: On the statistical decision problems with finite parameter space, *Kybernetika* **3** (1967), 451—466.

[8] HAHN, H.: Über die Integrale des Herrn Hellinger und die Orthogonalinvarianten der quadratischen Formen von unendlich vielen Veränderlichen, *Monatsh für Math. und Phys.*, **23** (1912), 161—224.

[9] BHATTACHARYYA, A.: On some analogues of the amount of information and their use in statistical estimation, *Sankhya* **8**, (1946), 1—14.

[10] RÉNYI, A.: *On measures of entropy and information,* Proc. 4th Berkeley Symp. on Prob. and Stat., Berkeley, Vol. 1, 547—561.

[11] KOOPMANS, S. H.: Asymptotic rate of discrimination for Markov processes, *Ann. Math. Stat.*, **31** (1960), 982—994.

[12] BOCHNER, S.: *Lectures on Fourier integrals,* Princeton 1969.

[13] LOÈVE, M.: *Probability Theory,* D. Van Nostrand, N. Y., 1960.

[14] KAILATH, T.: Comparison of the divergence and the Bhattacharyya distance measure, Presented at the 4th Prague Conference on Information Theory, Stat. Dec. Functions, and Rand. Processes, 1965.

[15] PEREZ, A.: Some estimates of the probability of error in discriminating two stationary random processes, Presented at the Conference on information theory in Dubna, USSR, 1969.

[16] VAJDA, I.: *On the convergence of information contained in a sequence of observations.* Proc. Coll. on Inf. Theory, Debrecen, II, Budapest 1969.

UTIA ČSAV

Prague 2, Vyšehradská 49.

(Received April 8, 1970)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 335—339.

COMBINATORIAL PROPERTIES OF PLANE PARTITIONS

by

G. J. SIMMONS

Introduction. There is a well known "folk theorem" which states that if the points of the plane are arbitrarily partitioned into two sets, then at least one of these sets contains the vertices of some equilateral triangle. Variations on this result have been popular in the problem literature for some time. For example, the same problem as above, except for rectangles, is given and solved in [1], two similar problems for unit side equilateral triangles are discussed by LEO MOSER in [2], and in [3] it is stated as a theorem that if the plane is arbitrarily covered by two closed sets, every triangle can be placed so that its vertices all lie in the same set. In this note we treat two generalizations of the folk theorem; the first of which was posed as a problem by JOHN ANNULIS of the University of New Mexico in May 1969: "Does the conclusion remain valid if instead of two sets an arbitrary but finite number of sets are allowed?"

Analysis for an Equilateral Grid of Points. Theorem 1, which we prove by combinatorial arguments, answers ANNULIS' question in the affirmative. First, for any partitioning of the plane we define a triangle to be monochromatic if all of its vertices are in the same set.

THEOREM 1. *If the points of the plane are arbitrarily partitioned into a finite number, k, of sets then there exists at least one monochromatic equilateral triangle.*

PROOF. Assume the theorem does not hold. Define N_k recursively by the expression

$$N_k = n(k, N_{k-1}+1) \tag{1}$$

where $N_0 = 1$ and $n(k, l)$ is the usual van der Waerden number* [4, 5]. Choose an arbitrary line in the plane and mark off N_k equispaced points which are used to form the base of an equilateral grid E_k of $\frac{N_k(N_k+1)}{2}$ points. For any partitioning of the N_k points on one side of E_k into k sets, at least $N_{k-1}+1$ equispaced points must be assigned to the same set — by the definition of N_k. Form an equilateral grid E_{k-1} based on some $N_{k-1}+1$ of these equispaced points, which are in the same set, say A_k. None of the $\frac{(N_{k-1})(N_{k-1}+1)}{2}$ points not on the base of E_{k-1}

* For k and l arbitrary positive integers, there exists a lower bound $n(k, l)$ such that if $N \geqq n$ (k, l) successive integers are arbitrarily assigned to k classes, then at least one class contains an arithmetic progression, having at least l terms. The quantity $n(k, l)$ is the van der Waerden number.

can be in set A_k by hypothesis, for if any point were in A_k it would complete a monochromatic equilateral triangle. In particular, the N_{k-1} points on the row next above the base in E_{k-1} must be partitioned among the remaining $k-1$ sets. But again, by the definition of N_{k-1}, this must result in at least $N_{k-2}+1$ equispaced points being assigned to the same set, say A_{k-1}. The previous argument also applies to this set of points. If this procedure is repeated $k-1$ times, on the $(k-1)^{st}$ step, since $N_1=1$, at least N_1+1 or 2 points will be assigned to the last remaining set. The point which forms an equilateral triangle, E_1, with two of these points is also a point in the equilateral grids $E_2, E_3, \ldots, E_{k-1}, E_k$ by construction, hence, no matter which set it assigned to it must be a vertex of a monochromatic equilateral triangle, which contradicts the original assumption. ∎

We include the usual proof for $k=2$ (the folk theorem) because of its extreme simplicity. As in the proof of Theorem 1, assume the theorem is not true. For this case, at least two of the vertices of an arbitrary equilateral triangle must be in the same class, say A_1. Let these be points 1 and 2 in the equilateral grid below based on the points 1 and 2:

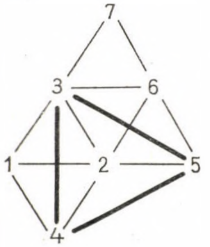


Fig. 1

Then points 3 and 4 must be in Class A_2 if neither of the equilateral triangles 1 2 3 nor 1 2 4 are to be monochromatic, but this requires 5 to be in class A_1 which, in turn, requires 6 to be assigned to class A_2. Now, if 7 is assigned to class A_1, this will complete the monochromatic equilateral triangle 1 5 7 in class A_1 and if 7 is assigned to class A_2 it will complete the monochromatic equilateral triangle 3 6 7 in class A_2. But point 7 must be assigned to one of the two classes.

Theorem 1 required only a finite number of points of the plane arranged in an equilateral grid for its proof, so that there are obviously uncountably many such monochromatic equilateral triangles. Theorem 2 shows that a surprisingly stronger result holds.

THEOREM 2.* *If the points of the plane are arbitrarily partitioned into countably many measurable sets then almost every (in the measure theoretic sense) point of the plane is a vertex of uncountably many monochromatic equilateral triangles whose vertices, furthermore, make up a set of positive measure.*

PROOF. First we recall that a point x is called a point of density for a set A if the outer density of A at x is unity; x is called a point of dispersion for A if the outer density of A at x is zero. It is well known [6] that if A is any set (measurable or not) in R_n, then x is a point of density for A for almost all $x \in A$. Furthermore, if A is measurable, almost all points of $\sim A$ are points of dispersion of A.

Assume Theorem (2) is not true. Let A be an open disc of radius one in the plane and designate the intersections with A of the measurable sets into which the plane is partitioned by $A_1, A_2, \ldots, A_i, \ldots$. Remove from each set A_i the set of points which are not points of density for A_i to form a set B_i. Since only countably many sets of measure zero have been eliminated, almost all points of A are in $\bigcup_i B_i$. Choose an point $x \in \bigcup_i B_i$, then x is a point of density for some B_j and a

* The author is indebted to JULIAN GEVIRTZ for suggesting this line of investigation.

point of dispersion for all B_i, $i \neq j$ from the way in which the B_i were constructed. For an arbitrary $\varepsilon < \frac{1}{2}$, construct a closed disc $S_\delta[x]$ centered on x of radius δ sufficiently small that $S_\delta[x] \subseteqq A$ and

$$\mu(S_\delta[x] \cap B_j) \geqq \pi\delta^2(1-\varepsilon) \tag{2}$$

which is possible since x is a point of density for B_j. Let R_x be the measure preserving operation of rotating the plane through 60° about x as center. Then it must be true that

$$(S_\delta[x] \cap B_j) \cap (R_x(S_\delta[x] \cap B_j)) \tag{3}$$

is a set of measure zero if the initial assumption is to hold, since any point in common to both the right and left parenthesis of Expression (3) would be two of the vertices of a monochromatic equilateral triangle in $(S_\delta[x] \cap B_j)$ with the third vertex being x. Thus, if the intersection given in Expression (3) has positive measure, x is already the vertex of uncountably many monochromatic triangles such that Theorem (2) is satisfied at the point x. Therefore, if the hypothesis is to hold at x, it must be true that after removing at most a set of measure zero corresponding to the intersection given in Expression (3) from $R_x(S_\delta[x] \cap B_j)$ to form

$$R_x(S_\delta[x] \cap B_j)^\dagger = R_x(S_\delta[x] \cap B_j) \sim ((S_\delta[x] \cap B_j) \cap (R_x(S_\delta[x] \cap B_j))) \tag{4}$$

that

$$R_x(S_\delta[x] \cap B_j)^\dagger \subset (S_\delta[x] \cap \bigcup_{\substack{i \\ i \neq j}} B_i) \cup (\bigcup_i (A_i \sim B_i)) \tag{5}$$

But R_x is a measure preserving transformation, hence

$$\mu(R_x(S_\delta[x] \cap B_j)^\dagger = \mu(S_\delta[x] \cap B_j) \leqq \mu(S_\delta[x] \cap \bigcup_{\substack{i \\ i \neq j}} B_i) \tag{6}$$

which is impossible because of the way in which ε was selected. Hence, all of the points of $\bigcup_i B_i$, and therefore all of the points of A except for, at most, a set of measure zero, are vertices of uncountably many monochromatic equilateral triangles whose vertices make up a set of positive measure. ∎

The sets A_i in the preceding proof had to be measurable to guarantee that a point of density for a set A_j was at the same time a point of dispersion for all of the other sets A_i, $i \neq j$.

Analysis for a Square Lattice of Points. A natural question is, "Can these results be extended to other regular tesselations of the plane?" RADO has shown this to be the case for the square lattice* and has given a powerful result [7] which he attributes to G. GRÜNWALD in his elegant analysis of "regular" linear systems** over the complex field.

* Witt purports to have proven an even more general result for homothetic figures [8] in an arbitrary plane partition.

** A system of equations

$$a_{\mu 1}x_1 + a_{\mu 2}x_2 + \ldots + a_{\mu n}x_n = b_\mu \qquad (1 \leqq \mu \leqq m), \tag{7}$$

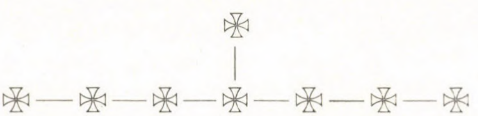
where the $a_{\mu\nu}$, b_μ are complex numbers, is said to be regular with respect to a set of numbers A, if the following condition holds: however we partition A into a finite number of subsets $A_1, A_2, \ldots, A_k$, always at least one of these subsets A_k contains a solution of (7).

Grünwald's Theorem:

"Given any 'configuration' S consisting of a finite number of lattice points of a Euclidean space, and given a distribution of all lattice points of this space into a finite number classes, there is at least one class which contains a configuration S′ of a lattice points which is similar and parallel (homothetic) to S."

Obviously, GRÜNWALD's Theorem includes as a special case a theorem equivalent to Theorem (1) in which the equilateral grid is replaced by a square lattice. In view of the power of GRÜNWALD's Theorem, we shall not prove the counterpart to Theorem (1) for the square lattice, but instead will include the proof for $k=2$, i.e., the case equivalent to the folk theorem since its proof also depends on an extremely simple argument.

LEMMA. *If the points of the plane are arbitrarily partitioned into two sets, and at least one of the sets contains the configuration*



where the points are at unit distance on a square lattice, then at least one of the sets must contain a monochromatic square.

PROOF. Assume the lemma is not true, then it must be possible to assign all of the lattice points to sets without forming a monochromatic square. Making the assignments forced by this assumption leads to a contradiction on the seventh step to prove the lemma. ■

THEOREM 3. *If the points of the plane are arbitrarily partitioned into two sets, then at least one of the sets contains a monochromatic square.*

PROOF. Choose $n(2, 9)$ equispaced collinear points which guarantees at least nine equispaced points in the same class. Take 9 such points and consider the four lattice points marked ○ and □.

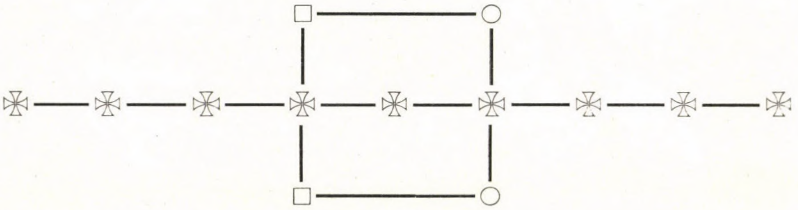


Fig. 2

Both of the points ○ must be in a class different from the class in which points ✠ are assigned or else the lemma applies. But there is only one other class. Similarly, points □ must both be in this other class if the lemma is not to apply, but this forms a monochromatic square. ■

Conclusion Theorem (1) demonstrated that uncountably many monochrome-equilateral triangles existed for an arbitrary partitioning of the plane into finitely many sets, while Theorem (2) showed an even stronger result was true for an arbitrary partitioning into countably many measurable sets. In the form in which they are given here, neither theorem implies the other; however, we have also been unsuccessful in constructing an example to show that either theorem is as strong as possible. The conclusion of Theorem (1) is probably true when countably many sets are permitted in the partition and a weakening of the conclusions of Theorem (2) should make it possible to remove the requirement that the partition sets be measurable. Unfortunately, the arguments used in this paper were inherently finite for Theorem (1) and dependent in an essential manner on the partition sets being measurable in Theorem (2) so that we have been unable to link these theorems.

REFERENCES

[1] Problem 138 Proposed by David L. Silverman, Solved by John E. Ferguson, *Pi Mu Epsilon Journal,* Fall 1963.
[2] GARDNER, M.: *New Mathematical Diversions from Scientific American,* pp. 120—123, Simon and Schuster, New York, 1966.
[3] HADWIGER, H. and DEBRUNNER, H.: *Combinatorial Geometry on the Place,* Theorem 61, p. 25, Holt, Rinehart and Winston, New York, 1964.
[4] VAN DER WAERDEN, B. L.: „Beweis einer Baudetschen Vermutung", *Nieuw Archief voor Wiskunde,* **15** (1927), 212—216.
[5] KHINCHIN, A. Y.: *Three Pearls of Number Theory* (translation), Graylock Press, Rochester, N. Y., pp. 11—17, 1952.
[6] MUNROE, M. E.: *Introduction to Measure and Integration,* Addison-Wesley Co., Reading, Pa., pp. 287—291, 1953.
[7] RADO, R.: Note on Combinatorial Analysis, *Proceedings of the London Mathematical Society,* Second Series, **48** (1945), 122—160.
[8] VON WITT, E.: Ein Kombinatorischer Satz der Elementargeometrie, *Mathematische Nachrichten,* **6** (1951—52), 261—262.

Sandia Laboratories and University of New Mexico

(Received December 11, 1969)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 341—346.

A 3-SPHERE THAT IS NOT 4-POLYHEDRAL

by

D. BARNETTE and G. WEGNER

1. Introduction. In [2] GRÜNBAUM constructs a triangulation $\mathcal{M}$ of the 3-sphere that is not combinatorially equivalent to the boundary complex of any 4-dimensional convex polytope (hereafter to be called a 4-*polytope*). It follows that the dual complex $\mathcal{M}^*$ is also not combinatorially equivalent to the boundary complex of any 4-polytope. We shall show that the 2-skeleton of $\mathcal{M}^*$ is not geometrically realizable, that is, there is no complex combinatorially equivalent to the 2-skeleton of $\mathcal{M}^*$ in which each 2-cell is a convex polygon. We shall also show that the graphs of triangulations of the 3-sphere and the graphs of their duals satisfy all the known necessary conditions to be graphs of 4-polytopes, and yet no 4-polytope has a graph isomorphic to the graph of $\mathcal{M}^*$.

2. Definitions. A *cell complex* $\mathcal{C}$ is a collection of convex polytopes such that every face of a member of $\mathcal{C}$ is a member of $\mathcal{C}$ and the intersection of any two members of $\mathcal{C}$ is a face of both ($\emptyset$ being a face of all members of $\mathcal{C}$, and each member being a face of itself). The *dimension* of a cell complex is the greatest dimension of any of its members. A *Topological d-cell complex* will be a collection $\mathcal{C}$ of k-cells $-1 \leqq k \leqq d$ such that

(i) each k-cell c is homeomorphic to some convex k-dimensional polytope P_c.
(ii) any face of a k-cell in $\mathcal{C}$ is a member of $\mathcal{C}$, where a face of a k-cell is the image of a face of P_c under the homeomorphism.
(iii) the intersection of two members of $\mathcal{C}$ is a face of both.

If $\mathcal{C}$ is a cell complex (topological or otherwise) then $\mathrm{Skel}_k \mathcal{C}$ is the complex consisting of all faces of $\mathcal{C}$ of dimension k or less. Two complexes $\mathcal{C}_1$ and $\mathcal{C}_2$ are *combinatorially equivalent* provided there is a 1—1 correspondence of the faces of $\mathcal{C}_1$ onto the faces of $\mathcal{C}_2$ which preserves incidences, two faces being *incident* if one is a subset of the other. If v is a vertex (i.e. 0-dimensional face) of a complex $\mathcal{C}$ then the *linked complex*, link v, is the set of all faces of $\mathcal{C}$ which are incident to faces which are incident to v but which are not themselves incident to v. A complex $\mathcal{C}_1$ is a *refinement* of $\mathcal{C}_2$ if there is a homeomorphism between them such that the image of any face of $\mathcal{C}_2$ is a subcomplex of $\mathcal{C}_1$.

The *boundary complex* of a d-polytope P is the complex $\mathrm{Skel}_{d-1} P$. A k-complex is d-polyhedral if it is combinatorially equivalent to $\mathrm{Skel}_k P$ for some d-polytope P. A d-cell complex is *simple* if each k-face, $0 \leqq k \leqq d$, is contained in exactly $d+1-k$ different d-faces. We shall use a theorem of GRÜNBAUM [2 pg. 206] that states that every simple d-cell complex is the boundary complex of a simple $(d+1)$-polytope.

A topological cell complex is *geometrically realizable* if it is combinatorially equivalent to a cell complex.

3. The complex $\mathcal{M}$ and its dual. Since giving a description of $\mathcal{M}$ is equivalent to describing $\mathcal{M}^*$ we shall concern ourselves only with the latter. $\mathcal{M}^*$ consists of 8 3-cells of two different combinatorial types (see Fig. 1) and their faces.

A complete description of $\mathcal{M}^*$ is given in Table 1, the facets are denoted by the numbers $1, 2, \ldots, 8$ and the vertices by $A, B, \ldots, Y$.

Table 1

Combinatorial description of the 3-complex $\mathcal{M}^*$

3-cell of $\mathcal{M}^*$	Vertices of the 3-cell (corresponding to the vertices $a, b, \ldots, j$ in Figure 1)	Combinatorial type of the 3-cell
1	*U S C B A W X T H F*	I
2	*U S Q P V A B C L J*	I
3	*N M L Q R J A F B P*	I
4	*N J A F M O Y V U W*	II
5	*O N R T H Y V J P X*	I
6	*N R T H O M L Q S C*	II
7	*O M F W Y H C L B X*	I
8	*U S T X W V P Q R Y*	I

THEOREM 1. $\text{Skel}_2 \mathcal{M}^*$ *is not geometrically realizable.*

PROOF. Suppose $\mathscr{C}$ is a cell complex in E^n which realizes $\text{Skel}_2 \mathcal{M}^*$ geometrically. Let S be any 3-cell in $\mathcal{M}^*$, then its boundary complex is a simple topological cell complex and the corresponding complex in $\mathscr{C}$ is a simple cell complex. This simple cell complex is the boundary complex of some 3-polytopes in E^n.

We shall consider in particular the 3-polytopes 1, 2, and 8, which are all of combinatorial type I. Their intersection is the edge SU of $\mathcal{M}^*$ and, as easily checked from Table 1, the edge SU is equivalently situated in all three of them; it corresponds to the edge *ba* in Figures 1 and 2.

Let any 3-polytope of type I be given. We consider (see Figure 2) the lines L_0, L_1, and L_2 determined respectively by a and b, by c and d, and by g and h.

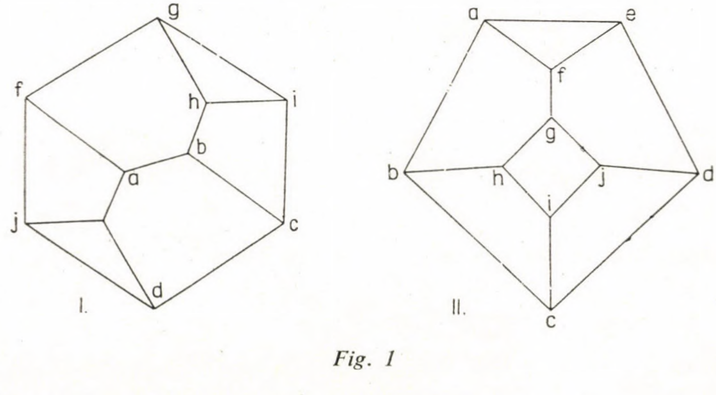


Fig. 1

Since L_0 and L_1 are coplanar, they intersect in a point p_1 (which is, possibly, at infinity). Similarly, L_0 and L_2 intersect in a point p_2. Considering the pentagon *abhgf*, and noting that the line determined by f and g is coplanar with L_1 and therefore intersects L_0 at p_1, we see that the pair a, p_2 separates the pair b, p_1 on the projective) line L_0.

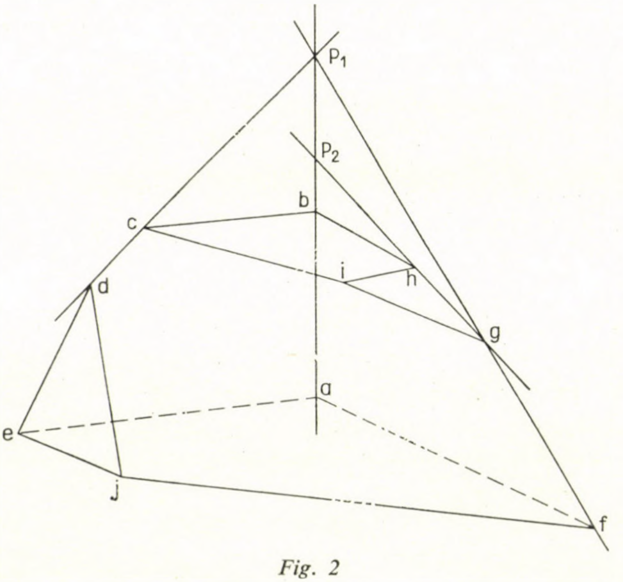


Fig. 2

Now the impossibility of the assumed realization of $\mathrm{Skel}_2 \mathcal{M}^*$ by a cell complex is obvious. Indeed, let us denote by L_0 the line determined by SU, and by Z_1, Z_2, Z_3 the intersections of L_0 with the lines determined respectively by XT, PQ, BC. Applying the above remark in turn to each of the three 3-polytopes corresponding to 8, 2, 1, we see that the pair S, Z_1 separates the pair U, Z_2, the pair S, Z_2 separates U, Z_3, and S, Z_3 separates U, Z_1. This being absurd, the proof of the Theorem is completed.

We now turn to the graph of $\mathcal{M}^*$. We shall need the following theorem of STEINITZ [5]: *A graph is 3-polyhedral if and only if it is planar and 3-connected.* From this it follows that any 2-cell complex homeomorphic to the 2-sphere is 3-polyhedral and any triangulation of the 2-sphere is 3-polyhedral.

If Γ is a circuit in a graph and the only edges joining vertices of Γ are edges of Γ then we say that Γ is *without diagonals.*

LEMMA. *If Γ is a circuit without diagonals with 3, 4 or 5 edges in the graph of a simple 4-polytope P then Γ consists of the edges of some 2-face of P.*

PROOF. We shall give the proof for a circuit of five edges; the proof in the other two cases is quite similar. Since P is simple each pair of consecutive edges on Γ belongs to a 2-face and each three consecutive edges belong to some 3-face of P. Let the vertices of Γ be $v_1, v_2, \ldots, v_5$ in cyclic order. Let F_1 be a 3-face of P containing the edges $v_1 v_2$ and $v_2 v_3$, and let F_2 be a 2-face containing the edges $v_3 v_4$, $v_4 v_5$ and $v_5 v_1$. The intersection of F_1 and F_2 contains v_1 and v_3, but this intersection is either a vertex, an edge, or the face F_1. Since Γ is without diagonals, $F_1 \cap F_2$ is the face F_1.

Let the graph of F_2 be embedded in the plane and assume Γ does not bound a 2-face of P. Then there are edges of the graph of F_2 going inside Γ from vertices of Γ and also edges going outside. Since each vertex is 3-valent there are only two edges going to one of these regions (inside or outside) but this implies that the graph of F_2 is not 3-connected which is a contradiction to STEINITZ' Theorem.

THEOREM 2. *The graph* $\mathcal{G}$ *of* $\mathcal{M}^*$ *is not* 4-*polyhedral.*

PROOF. Suppose P is a 4-polytope whose graph is $\mathcal{G}$. By the lemma above each circuit without diagonals and with 5 or fewer edges in P, bounds a 2-face thus these 2-faces correspond to 2-faces of $\mathcal{M}^*$. If we can show that the remaining 2-faces of $\mathcal{M}^*$ correspond to 2-faces of P we are done because this would imply that $\mathrm{Skel}_2 P$ is a geometric realization of $\mathrm{Skel}_2 \mathcal{M}^*$. The only other 2-cells of $\mathcal{M}^*$ are the 6-sided cells of the 3-cells of type I. Given a cell c of type I all but one 2-face has 5 or fewer sides thus each of these corresponds to a 2-face in P. Consider the 2-faces α and β (see Fig. 3) in c. The corresponding 2-cells in P belong to some 3-face F. The face corresponding to γ must belong to the same facet since only one 2-face in P meets α and β at v_1, v_2 and v_3 and clearly some 2-face in F does just that. Similarly wesee that δ, ε and ζ belong to F. This shows that the graph in Fig. 3 is a subgraph-of the graph of F. Since each vertex in F is 3-valent we see that the graph in Fig. 3 is the graph of F and thus the 6-sided face in Fig. 3 corresponds to the seventh 2-face of F.

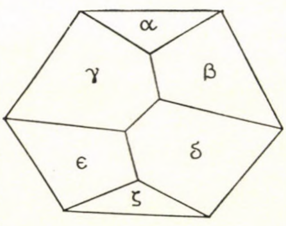


Fig. 3

4. Conditions for 4-polyhedrality. The following necessary conditions for a graph $\mathcal{G}$ to be 4-polyhedral are known:

1. If v is a vertex of a graph $\mathcal{G}$ and V is the set of neighbors of v (i.e. vertices of $\mathcal{G}$ joined to v by edges) then V is contained in a 3-polyhedral subgraph $\mathcal{G}'$ of $\mathcal{G} \sim \{v\}$, and $\mathcal{G}'$ contains a refinement of the complete graph of 4 vertices whose principal vertices are in V [1].
2. If a set of n vertices separates $\mathcal{G}$, then it separates it into at most $(n^2-3n)/2$ components [3].

THEOREM 3. *If* $\mathcal{G}$ *is the graph of a triangulation of the* 3-*sphere, then* $\mathcal{G}$ *satisfies condition* 1.

PROOF. Since the linked complex of a vertex in a 3-manifold is a 2-sphere (See [4, Th. 1]) we conclude that the graph of the linked complex of v is 3-polyhedral. Since each vertex of this graph is a neighbor of v, the result follows because each 3-polyhedral graph contains a refinement of the complete graph on 4 vertices (See [2, p. 201]).

THEOREM 4. *If* $\mathcal{G}$ *is the graph of the dual* T^* *of a triangulation* T *of the* 3-*sphere then* $\mathcal{G}$ *satisfies condition* 1.

PROOF. We begin by showing that link v in T^* is a refinement of the tetrahedron, for each vertex v. Let F be any 3-face of T^* that contains v. Link v in F is a circuit that separates $\beta(F)$ (the boundary of F) into two 2-cells, one of which consists of the 2-cell in $\beta(F)$ that are in link v. We shall call this cell C_F. Two cells C_F and C_H meet only if F and H meet on a 2-face or an edge, and when F and H meet in this way then C_F and C_H meet on an arc or at a vertex. The cells of the form C_F thus form a topological 2-cell complex $\mathscr{C}$. The correspondence $C_F \leftrightarrow F \leftrightarrow v_F$, where v_F is the vertex in T corresponding to F by duality, shows that $\mathscr{C}$ is combinatorially equivalent to the dual of the 3-cell $\mathscr{F}$ of T corresponding to v by duality. Indeed the set of v_F's is the set of vertices of $\mathscr{F}$ and two vertices v_F and v_H are joined by an edge if and only if F and H intersect on a 2-cell, which happens if and only if C_F and C_H meet on an arc.

Since $\mathscr{F}$ is a tetrahedron, link v is a refinement of the tetrahedron. It follows that the graph of $\mathscr{C}$ is a refinement of the graph of the tetrahedron and the neighbors of v are the principal vertices.

For the next theorem we use a theorem by KLEE which is actually condition 2 for 3-polytopes [3]:

If n vertices separate a 3-polyhedral graph then it is separated into $2n-4$ or fewer components.

We shall also use the fact that condition 1 implies that $\mathscr{G}$ is 4-connected (See [1] for a proof).

THEOREM 4. *Let V be a set of n vertices which separates the graph of either T or T^*. Then the number of components of the separated graph is at most $(n^2-3n)/2$.*

PROOF. Let the elements of V be $v_1, \dots, v_n$; let their linked complexes be $\mathscr{L}_1, \mathscr{L}_2, \dots, \mathscr{L}_n$ respectively; and let C_i be the number of components that meet $\mathscr{L}_i$. If n_i is the number of vertices of V in $\mathscr{L}_i$ then by KLEE's condition for $d=3$ we have that $C_i \leqq 2n_i - 4$, thus

$$\sum_{i=1}^{n} C_i \leqq \sum_{i=1}^{n} 2n_i - 4$$

But $\sum_{i=1}^{n} n_i \leqq n^2 - n$ thus

$$\sum C_i \leqq 2(n^2-n)-4n = 2n^2-6n$$

Each component meets the linked complex of at least 4 vertices in V because the graph is 4-connected, thus

$$C \leqq \frac{1}{4}\sum_{i=1}^{n} C_i \leqq \frac{n^2-3n}{2}$$

where C is the number of components.

The above shows that even though the graph of $\mathscr{M}^*$ is not 4-polyhedral, it satisfies all the known conditions for 4-polyhedrality. Although there are other examples of graphs satisfying 1 and 2 which are not 4-polyhedral, this is the first, to the best of the authors' knowledge, which is the graph of a cellular decomposition of the 3-sphere.

Acknowledgement: The authors are indebted to BRANKO GRÜNBAUM for allowing us to use his version of the proof of Theorem 1, which he wrote up for use in a seminar at the University of Washington in 1968.

REFERENCES

[1] BARNETTE, D.: A necessary condition for d-polyhedrality, *Pac. J. Math.* 23 (1967) 435—440.
[2] GRÜNBAUM, B.: *Convex Polytopes,* Wiley, 1967.
[3] KLEE, V.: A Property of d-Polyhedral Graphs, *J. Math. Mech.* 13 (1964) 1039—1042.
[4] MOISE, E. E.: Affine Structures in 3-manifolds, V, (The Triangulation Theorem and Hauptvermutung), *Ann. of Math.* 56 (1952) 96—114.
[5] STEINITZ E. and RADEMACHER, H.: *Vorlesungen Über Die Theorie Der Polyeder,* Springer, Berlin 1934.

University of California, Davis; University of Göttingen, Göttingen, Germany

(Received May 5, 1970)

ON HYPERMATRICES WITH BLOCKS COMMUTABLE IN PAIRS IN THE THEORY OF MOLECULAR VIBRATIONS

by

B. GELLAI

Introduction

Matrix formalism is routinely used for the solution of problems in modern chemistry. Here we consider, in particular, the problem of molecular vibration of the form [1]

$$|\mathbf{GF}-\lambda\mathbf{E}| = 0, \tag{1}$$

where **G** is the inverse of the matrix of kinetic energy depending on the interatomic distances and mass of the molecule, **F** is the matrix of potential energy determined by the force constants.

The problem, mentioned above, can be treated as an eigenvalue problem since the eigenvalues of the **GF** matrix are proportional to the individual frequencies of the molecular vibrations, or as an "inverse eigenvalue problem" if the force constants are to be determined from the elements of the **G** matrix and from the eigenvalues.

The solution of either problem becomes difficult in the case of polyatomic molecules, since the order of the matrices increases with the number of atoms involved. Efforts have been made therefore to split the given problem into a set of smaller problems. A known method in chemistry for this is the construction of symmetry coordinates using some group theoretical considerations, in terms of which the matrix of vibrational problems is reduced to the maximum extent made possible by the molecular symmetry [1].

For molecules, having a "good" symmetry, the **GF** matrix has in some cases a structure such that it can be reduced in terms of pure matrix theory.

In this paper a method based on EGERVÁRY's theorem [2] will be described for the complete reduction of the **GF** matrix which consists of blocks commutable in pairs.

Application of the method will be shown in the case of methyl halide molecules.

1. § Description of the method

The following notation is used:

$\mathbf{A}=[a_{ij}]$	is the matrix composed of the scalars a_{ij}
$[\mathbf{A}_{ij}]$	is the hypermatrix composed of the blocks $\mathbf{A}_{ij}$
$\langle a_1, a_2, \ldots, a_n\rangle$	is the diagonal matrix composed of the scalars a_i

$\langle \mathbf{A}^{(l)} \rangle_1^n = \langle \mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \ldots, \mathbf{A}^{(n)} \rangle$	is the hyperdiagonal matrix of order n composed of the square matrices $\mathbf{A}^{(l)}$
$\mathbf{A}^*$	is the transpose of matrix $\mathbf{A}$
$\mathbf{E}_n$	is the unit matrix of order n
$\boldsymbol{u}, \boldsymbol{v}$	are column vectors
$\boldsymbol{u}^*, \boldsymbol{v}^*$	are row vectors
$\mathbf{A} \cdot \times \mathbf{B} = [\mathbf{A} \cdot b_{ij}]$	is the direct product of the matrices $\mathbf{A}$ and $\mathbf{B}$

$$\boldsymbol{u} \cdot \times \boldsymbol{v} = \begin{bmatrix} uv_1 \\ uv_2 \\ \vdots \\ uv_n \end{bmatrix} \ldots \text{ is the direct product of vectors } \boldsymbol{u} \text{ and } \boldsymbol{v}.$$

$\mathbf{T} = [\boldsymbol{t}_{kl}] = [\boldsymbol{t}_{11}, \boldsymbol{t}_{12}, \ldots, \boldsymbol{t}_{1n}; \boldsymbol{t}_{21}, \boldsymbol{t}_{22}, \ldots, \boldsymbol{t}_{2n}; \ldots]$ is a matrix partitioned in the column vectors $\boldsymbol{t}_{kl}$.

Let $[\mathbf{A}_{ij}]$ $(i, j = 1, 2, \ldots, n)$ be a hypermatrix of order $n \times m$ with blocks $\mathbf{A}_{ij} = p_{ij}(\mathbf{A})$, where $\mathbf{A}$ is a symmetrical matrix of order m and $p_{ij}(x)$ are polynomials of the real variable x, subject only to the restriction $p_{ij}(x) = p_{ji}(x)$. If the spectral decomposition of matrix $\mathbf{A}$ is given by

$$\mathbf{A} = \mathbf{W} \langle \lambda_1, \lambda_2, \ldots, \lambda_m \rangle \mathbf{W}^*,$$

then the blocks $\mathbf{A}_{ij}$ decompose to

$$\mathbf{A}_{ij} = p_{ij}(\mathbf{A}) = \mathbf{W} \langle p_{ij}(\lambda_1), p_{ij}(\lambda_2), \ldots, p_{ij}(\lambda_m) \rangle \mathbf{W}^*.$$

Thus, $[\mathbf{A}_{ij}]$ can be factorised as

$$[\mathbf{A}_{ij}] = (\mathbf{W} \cdot \times \mathbf{E}_n) \cdot \mathbf{P} \cdot \langle \tilde{\mathbf{A}}^{(1)}, \tilde{\mathbf{A}}^{(2)}, \ldots, \tilde{\mathbf{A}}^{(m)} \rangle \cdot \mathbf{P}^* \cdot (\mathbf{W}^* \cdot \times \mathbf{E}_n), \tag{2}$$

where $\tilde{\mathbf{A}}^{(k)} = [p_{ij}(\lambda_k)]$ and $\mathbf{P}$ is the permutation matrix which transforms the sequence of ordered pairs

$$(11)(12)\ldots(1m)(21)(22)\ldots(2m)\ldots(n1)(n2)\ldots(nm)$$

into the sequence

$$(11)(21)\ldots(n1)(12)(22)\ldots(n2)\ldots(1m)(2m)\ldots(nm)[2].$$

On multiplying the left and right hand side of [2] by $\mathbf{P}^* \cdot (\mathbf{W}^* \cdot \times \mathbf{E}_n)$ and $(\mathbf{W} \cdot \times \mathbf{E}_n) \cdot \mathbf{P}$, respectively, we obtain

$$\mathbf{P}^* \cdot (\mathbf{W}^* \cdot \times \mathbf{E}_n) \cdot [\mathbf{A}_{ij}] \cdot (\mathbf{W} \cdot \times \mathbf{E}_n) \cdot \mathbf{P} = \langle \tilde{\mathbf{A}}^{(k)} \rangle_1^m. \tag{3}$$

The transformation matrix $(\mathbf{W} \cdot \times \mathbf{E}_n)\mathbf{P} \cdot$ can be written in the form

$$\mathbf{T} = [\boldsymbol{t}_{kl}] = [\boldsymbol{w}_k \cdot \times \boldsymbol{e}_l] \qquad (k = 1, 2, \ldots, m; \; l = 1, 2, \ldots, n), \tag{4}$$

where $\boldsymbol{w}_k$ is the column k of matrix $\mathbf{W}$, that is the eigenvector corresponding to the eigenvalue λ_k of the matrix $\mathbf{A}$, while $\boldsymbol{e}_l$ is the lth unit vector of order n. If l runs over the values $1, 2, \ldots, n$, the direct products in (4), for fixed k, form the kth "block-column" of the matrix $\mathbf{T}$.

If the matrix $[\mathbf{A}_{ij}]$ is bordered by a row and a column vector in the following manner

$$\left[\begin{array}{c|c} c & \boldsymbol{w}_p^*\cdot\times\boldsymbol{v}^* \\ \hline \boldsymbol{w}_p\cdot\times\boldsymbol{u} & [\mathbf{A}_{ij}] \end{array}\right], \tag{5}$$

where c is constant, $\boldsymbol{w}_p$ is the pth characteristic vector of $\mathbf{A}$, $\boldsymbol{v}^*$ and $\boldsymbol{u}$ are vectors of order n, then the transformation matrix $\mathbf{T}$ — which is bordered by unit vectors — transforms the form (5) into

$$\left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & [\boldsymbol{w}_k^*\cdot\times\boldsymbol{e}_l^*] \end{array}\right]\cdot\left[\begin{array}{c|c} c & \boldsymbol{w}_p^*\cdot\times\boldsymbol{v}^* \\ \hline \boldsymbol{w}_p\cdot\times\boldsymbol{u} & [\mathbf{A}_{ij}] \end{array}\right]\cdot\left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & [\boldsymbol{w}_k\cdot\times\boldsymbol{e}_l] \end{array}\right] = \tag{6}$$

$$= \left[\begin{array}{c|c} c & (\boldsymbol{w}_p^*\cdot\times\boldsymbol{v}^*)\cdot[\boldsymbol{w}_k\cdot\times\boldsymbol{e}_l] \\ \hline [\boldsymbol{w}_k^*\cdot\times\boldsymbol{e}_l^*]\cdot(\boldsymbol{w}_p\cdot\times\boldsymbol{u}) & [\boldsymbol{w}_k^*\cdot\times\boldsymbol{e}_l^*]\cdot[\mathbf{A}_{ij}]\cdot[\boldsymbol{w}_k\cdot\times\boldsymbol{e}_l] \end{array}\right].$$

Performing the multiplication for the lowest block on the left we get

$$[\boldsymbol{w}_k^*\cdot\times\boldsymbol{e}_l^*]\cdot(\boldsymbol{w}_p\cdot\times\boldsymbol{u}) = \begin{array}{r} 1) \\ \\ \\ k) \\ \\ \\ \\ m) \end{array}\left[\begin{array}{c} (\mathbf{w}_1^*\mathbf{w}_p)\cdot(\mathbf{e}_l^*\mathbf{u}) \\ \hline \vdots \\ \hline (\mathbf{w}_k^*\mathbf{w}_p)(\mathbf{e}_1^*\mathbf{u}) \\ \vdots \\ (\mathbf{w}_k^*\mathbf{w}_p)\cdot(\mathbf{e}_l^*\mathbf{u}) \\ \vdots \\ (\mathbf{w}_k^*\mathbf{w}_p)(\mathbf{e}_n^*\mathbf{u}) \\ \hline \vdots \\ \hline (\mathbf{w}_m^*\mathbf{w}_p)(\mathbf{e}_l^*\mathbf{u}) \end{array}\right]\begin{array}{l} \\ \\ 1 \\ \\ l = \\ \\ n \\ \\ \\ \end{array}\left[\begin{array}{c} \delta_{1p}u_l \\ \vdots \\ \delta_{kp}u_l \\ \vdots \\ \delta_{mp}u_l \end{array}\right]. \tag{7}$$

The product of $\mathbf{T}^*$ and $(\boldsymbol{w}_p \cdot \times \boldsymbol{u})$ is non zero only if $k=p$ and this product is precisely the product of the pth „block-row" of $\mathbf{T}^*$ and the vector $(\boldsymbol{w}_p \cdot \times \boldsymbol{u})$. (In fact, for $p \neq k$ on the right hand side of (7) we have the product of two eigenvectors corresponding to different eigenvalues of $\mathbf{A}$, which is equal to zero.) If l runs over the values $1, 2, \ldots, n$, we obtain the vector $\boldsymbol{u}$. Consequently, the hyperdiagonal matrix (6) is bordered by vectors partitioned into m parts, the pth of which is only aiffering from zero, therefore the bordering row and column vectors can be written ds $(\boldsymbol{e}_p^* \cdot \times \boldsymbol{v}^*)$ and $(\boldsymbol{u} \cdot \times \boldsymbol{e}_p)$ respectively. Thus (6) becomes

$$\left[\begin{array}{c|c} c & \boldsymbol{e}_p^* \cdot \times \boldsymbol{v}^* \\ \hline \boldsymbol{u} \cdot \times \boldsymbol{e}_p & \langle \tilde{\mathbf{A}}^{(k)} \rangle_1^m \end{array}\right], \tag{8}$$

where $\langle \tilde{\mathbf{A}}^{(k)} \rangle_1^m$ is the hyperdiagonal matrix defined by (3).

Without loss of generality, we may put $p=1$ choosing an appropriate numbering of the eigenvalues. In this case the hyperdiagonal matrix (8) will have the form

$$\left[\begin{array}{cc|cccc} c & \boldsymbol{v}^* & & & & \\ \boldsymbol{u} & \tilde{\mathbf{A}}^{(1)} & & & 0 & \\ \hline & & \tilde{\mathbf{A}}^{(2)} & & & \\ & 0 & & \ddots & & \\ & & & & \tilde{\mathbf{A}}^{(m)} & \end{array}\right]. \tag{9}$$

In the solution of a given vibrational problem it frequently occurs that the hypermatrices have blocks $\mathbf{A}_{ij}$ with a cyclic structure, that is, blocks, whose elements are related as

$$c_{ij} = \begin{cases} c_{j-i} & \text{if } j \geqq i \\ c_{n+j-i} & \text{if } j < i. \end{cases}$$

These matrices are uniquely determined by their first row:

$$\mathbf{C}(c_0 c_1 c_2 \ldots c_{n-1}).$$

It is well known, that any cyclic matrix of order n can be written as a maximum $(n-1)$-th order polynomial of the primitive cyclic matrix $\mathbf{\Omega} = \mathbf{C}(0\ 1\ 0 \ldots 0)$ of order n. The eigenvalues of $\mathbf{\Omega}$ are the nth roots of 1 and the components of their eigenvectors are the powers of these nth roots of 1 [3].

Specifically, if the cyclic matrix of order 3 is symmetrical

$$\mathbf{C}(c_0 c_1 c_1), \tag{10}$$

its spectral decomposition is

$$\mathbf{C}(c_0 c_1 c_1) = \mathbf{W} \langle \lambda_0, \lambda_1, \lambda_2 \rangle \mathbf{W}^*,$$

where $\lambda_0 = c_0 + 2c_1$, $\lambda_1 = \lambda_2 = c_0 - c_1$ and

$$(11) \qquad \mathbf{W} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

2. § Application

The vibrational problem of a given molecule can be formulated in terms of *internal coordinates*, i.e. coordinates determined by the changes in the interatomic distances and in the angles betwen chemical bonds, which are the most physically significant set for use in describing the potential energy of the molecule [1].

For the methyl halide molecules the internal coordinates are the changes in the distances R, r_i $(i=1, 2, 3)$ and in the angles α_i, β_i $(i=1, 2, 3)$. (See the methyl iodide molecule CH_3J in Fig. 1.)

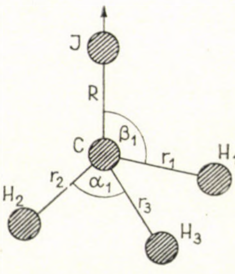


Fig. 1

For the present purpose we write down the **F** matrix only, since entirely analogous arguments can be applied to the **G** matrix and eventually to the secular equation (1). The **F** matrix in terms of internal coordinates has the form [4]:

$$(12) \qquad \mathbf{F} = [\mathbf{F}_{ij}] = \left[\begin{array}{c|ccc|ccc|ccc} f_{11} & f_{12} & f_{12} & f_{12} & f_{13} & f_{13} & f_{13} & f_{14} & f_{14} & f_{14} \\ \hline & f'_{22} & f_{22} & f_{22} & f'_{23} & f_{23} & f_{23} & f'_{24} & f_{24} & f_{24} \\ & f_{22} & f'_{22} & f_{22} & f_{23} & f'_{23} & f_{23} & f_{24} & f'_{24} & f_{24} \\ & f_{22} & f_{22} & f'_{22} & f_{23} & f_{23} & f'_{23} & f_{24} & f_{24} & f'_{24} \\ \hline & & & & f'_{33} & f_{33} & f_{33} & f'_{34} & f_{34} & f_{34} \\ & & & & f_{33} & f'_{33} & f_{33} & f_{34} & f'_{34} & f_{34} \\ & & \mathbf{F}_{ij}=\mathbf{F}_{ji} & & f_{33} & f_{33} & f'_{33} & f_{34} & f_{34} & f'_{34} \\ \hline & & & & & & & f'_{44} & f_{44} & f_{44} \\ & & & & & & & f_{44} & f'_{44} & f_{44} \\ & & & & & & & f_{44} & f_{44} & f'_{44} \end{array}\right] \qquad i, j = 1, 2, 3, 4.$$

Using the notation (10) for the blocks $\mathbf{F}_{ij}$, (12) can be written as (5):

$$(13) \qquad \left[\begin{array}{c|ccc} f_{11} & & \boldsymbol{w}_0^* \cdot \times \boldsymbol{u}^* & \\ \hline & \mathbf{C}(f'_{22}\; f_{22}\; f_{22}) & \mathbf{C}(f'_{23}\; f_{23}\; f_{23}) & \mathbf{C}(f'_{24}\; f_{24}\; f_{24}) \\ \boldsymbol{w}_0 \cdot \times \boldsymbol{u} & & \mathbf{C}(f'_{33}\; f_{33}\; f_{33}) & \mathbf{C}(f'_{34}\; f_{34}\; f_{34}) \\ & & & \mathbf{C}(f'_{44}\; f_{44}\; f_{44}) \end{array}\right]$$

where w_0 is the first column vector of matrix (11) and

$$u = \begin{bmatrix} \sqrt{3}\cdot f_{12} \\ \sqrt{3}\cdot f_{13} \\ \sqrt{3}\cdot f_{14} \end{bmatrix}.$$

Applying the method, described in § 1., matrix (13) can be transformed by making use of the matrix

$$(14) \qquad T = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \mathbf{W}\cdot\times\mathbf{E}_3 \end{array}\right] \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \mathbf{P} \end{array}\right] = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & [w_k\cdot\times e_l] \end{array}\right]$$

$$\mathbf{TT}^* = \mathbf{E}, \qquad (k, l = 1, 2, 3),$$

to the hyperdiagonal form:

$$(15) \qquad \left[\begin{array}{cc|c|c} f_{11} & u^* & & \\ u & \tilde{\mathbf{F}}^{(0)} & 0 & \\ \hline & & \tilde{\mathbf{F}}^{(1)} & \\ \hline & 0 & & \tilde{\mathbf{F}}^{(2)} \end{array}\right],$$

where

$$\tilde{\mathbf{F}}^{(k)} = [p_{ij}(\omega_k)] = \left[f'_{ij} + 2f_{ij}\cos\frac{2k\pi}{3}\right],$$

$k = 0, 1, 2$; $i, j = 2, 3, 4$ and ω_k are the cube roots of 1. Thus the given vibrational problem of order 10 has been reduced to a problem of order 4 and two identical problems of order 3.

Let us see now, whether the matrix (15) is the completely reduced form.

The methyl molecules CH_3X belong to the symmetry point group C_{3v}, composed of the following operations: two rotations by $\frac{2\pi}{3}$ and $\frac{4\pi}{3}$ about the axis coinciding with the $C-X$ bond, three reflections through the vertical planes σ_i $(i = 1, 2, 3)$ each passing through one of the $C-H_i$ bonds and the identity operation E. This operations are called the *symmetry operations* which carry the molecule into a configuration equivalent to its initial configuration [1].

The internal coordinates, introduced above, separate into sets which do not mix with oneanother, thus, the members of each set *transform only among themselves*, consequently the matrices, representing the symmetry operations are in diagonal block form, that is, the matrix representation of the symmetry point group C_{3v} is reducible [1].

It is known from representation theoretical consideration [4] that the C_{3v} symmetry point group has four 1-dimensional and three 2-dimensional irreducible representations, consequently the completely reduced form of matrix **R**, representing any symmetry operation of the molecule, contains four 1-dimensional blocks and three 2-dimensional blocks. That is

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}^{(1)} & & & & & & \\ & \mathbf{R}^{(1)} & & & & & \\ & & \mathbf{R}^{(3)} & & & 0 & \\ & & & \mathbf{R}^{(1)} & & & \\ & & 0 & & \mathbf{R}^{(3)} & & \\ & & & & & \mathbf{R}^{(1)} & \\ & & & & & & \mathbf{R}^{(3)} \end{bmatrix}.$$

The matrix of potential energy commutes with all group representation matrices and if it is partitioned to correspond with the diagonal blocks of **R**, the commutable rule will be valid for the individual blocks too, thus by Schur's lemma we find, by merely permuting rows and columns, that in the final form the **F** matrix contains one 4-dimensional block and two 3-dimensional blocks [4]. Consequently the form (15) is *completely reduced form.*

Alternatively, it can be shown, that applying the transformation with the matrix **T** for the 10-th order vector of the internal coordinates, we obtain the well known symmetry coordinates of methyl halide molecules [5], including the redundant coordinate the removal of which requires some chemical considerations.

Note. The methyl halide molecules have a „good" symmetry, therefore the matrix of the vibrational problem consists of blocks of a special cyclic structure, however, the method, described above, can be applied also to molecules having a "less good" symmetry. Attempts at the extension of the method to such molecules as well as the development of the method for the removal of redundant coordinates are in progress.

Acknowledgement. Thanks are due to Professor P. Rózsa and to A. Lee for helpful advice.

LITERATURE

[1] Wilson, E. B., Decius, J. C. and Cross, P.: *Molecular Vibrations,* McGraw—Hill, New York (1955).
[2] Egerváry, E.: *Acta Scient. Mathematicarum,* Tomus XV. Fasc. 1. 211—222 (1953).
[3] Aitken, A. C.: *Determinants and Matrices.* Oliver and Boyd Edinburgh and London, New York: Interscience Publishers, Inc.
[4] Маянц, Л. С.: Теория расчёт колебаний молекул. Изд. Академии Наук СССР, Москва 1960.
[5] Aldous, J., Mills, I. M.: *Spectrochim. Acta* **18** 1073 (1962).

Central Research Institute for Physics of the Hungarian Academy of Sciences, Budapest

(Received May 14, 1970.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 355—359.

A PROPERTY OF CONDITIONAL ENTROPY

by

J. KÖRNER

§. 1

Let $\{\xi_i\}_{i=1}^{\infty}$ be a discrete, memoryless and stationary information source the letters of which are taken from a finite or countably infinite alphabet X. According to Shannon's well-known theorem if $N(n, \lambda)$ denotes the minimum number of sequences of source symbols of length n with total probability greater than or equal to $1-\lambda$ (where $0 \leqq \lambda < 1$), then

$$\lim_{n \to \infty} \frac{1}{n} \log_2 N(n, \lambda) = \mathsf{H}(\xi),$$

where $\mathsf{H}(\xi)$ is the Shannon-entropy of the random variable ξ. Now, $N(n, \lambda)+1$ codewords are sufficient to code the blocks of length n of the source with an error probability less than λ.

The aim of the present paper is to prove a similar proposition for the conditional entropy.

§. 2

The conditional entropy $\mathsf{H}(\xi | f(\xi))$ can be interpreted as the remaining uncertainty concerning ξ if we are given $f(\xi)$, a function of the random variable ξ. Allow for this interpretation we give a coding procedure which could be called *complementary coding*. The coding of the source is realized in two steps. At the first step we consider instead of the original source $\{\xi_i\}_{i=1}^{\infty}$ the source $\{f(\xi_i)\}_{i=1}^{\infty}$, i.e., we are only interested in a "rough observation" of the source originally given. So we have the contracted alphabet $f(X)$. We are given the exact form of this „rough" representation of all the sequences of source symbols.

The second step is what we call a complementary coding. We want to code the source such a way that if we know the complementary codeword and the "rough" representation of an arbitrary source sequence, then we have a coding of the original source with an error probability less than λ. To attain this, at the second step we need not distinguish again two source sequences having different "rough" representations. We denote by $M(n, \lambda)+1$ the number of complementary codewords needed to code the original source $\{\xi_i\}_{i=1}^{\infty}$ with an error probability less than λ. We obtain the following result which is an analogon of Shannon's theorem:

$$\lim_{n \to \infty} \frac{1}{n} \log_2 M(n, \lambda) = \mathsf{H}(\xi | f(\xi)),$$

where $\mathsf{H}(\xi | f(\xi))$ is the conditional entropy of the random variable ξ.

§. 3

Now we turn to a mathematically rigorous treatment of our subject. Let $\{\xi_i\}_{i=1}^{\infty}$ be again a discrete, memoryless and stationary information source with letters taken from the finite or countably infinite alphabet X. We are also given a function f on X partitioning X.

Let us consider on X the following two relations:

$a \perp b$ if and only if $f(a) \neq f(b)$, that means the "rough" representations of a and b differ.

$a | b$ if and only if $f(a) = f(b)$ and $a \neq b$.

Let us extend these relations to the set X^n i.e. to the n-length sequences of letters from X. For two sequences $a = a_1 a_2 \dots a_n$ and $b = b_1 b_2 \dots b_n$

$a \perp b$ holds if and only if there exists an i with $a_i \perp b_i$.

$a | b$ holds if and only if $a \perp b$ does not hold and $a \neq b$. That means $a|b$ if and only if we have either $a_i = b_i$ or $a_i | b_i$ for every $i = 1, 2, \dots, n$ and for at least one i, we have $a_i | b_i$.

Now we pass over to an exact definition of $M(n, \lambda)$.

We call a "*good set*" a set which does not contain any pair of elements a, b, with $a|b$. A "*good decomposition*" of a set $A \subset X^n$ is then a decomposition of A whose components are all "good sets". We define the number $[A]$ for any set A as the minimum number of components in a "good" decomposition of A. As usual, let us denote by $\|A\|$ the number of elements of an arbitrary set A. We have

$$[A] = \min \left\{ r; A = \bigcup_{i=1}^{r} A_i, A_i \cap A_j = \emptyset \text{ for } i \neq j \text{ and } a, b \in A_i \Rightarrow a|b \text{ does not hold} \right\}$$

Let us define $M(n, \lambda)$ by

$$M(n, \lambda) = \min_{P(\xi_1 \xi_2 \dots \xi_n \in A) \geqq 1 - \lambda} [A].$$

Now we can establish our

THEOREM: *For every discrete, memoryless and stationary information source* $\{\xi_i\}_{i=1}^{\infty}$ *and for every* λ *satisfying* $0 \leqq \lambda < 1$ *the limit* $\lim_{n \to \infty} \frac{1}{n} \log_2 M(n, \lambda)$ *exists and we have*

$$\lim_{n \to \infty} \frac{1}{n} \log_2 M(n, \lambda) = \mathsf{H}(\xi | f(\xi)).$$

Remark 1: One may think that only a special class of conditional entropies is treated here, because $\mathsf{H}(\zeta|\eta)$ is defined for two arbitrary random variables ζ and η while in our case we have assumed that $\eta = f(\zeta)$. However, every conditional entropy can be reduced to this special form, since $\mathsf{H}(\zeta|\eta) = \mathsf{H}((\zeta, \eta|\eta)$; here η is a function of (ζ, η).

Remark 2: The theorem mentioned in the introduction is a special case of ours, where f is a constant function. Thus $\mathsf{H}(\xi | f(\xi)) = \mathsf{H}(\xi)$. However, this paper does not give any new proof for that case. We turn now to the

PROOF of our theorem: Let $F_n \subset X^n$ consist of all sequences $x = x_1 x_2 \dots x_n \in X^n$ with property $\left| -\frac{1}{n} \sum_{i=1}^{n} \log_2 p(x_i) - \mathsf{H}(\xi) \right| \leqq \varepsilon$ ($\varepsilon > 0$ is an arbitrary positive number), i.e.

$$(1) \qquad F_n = F_n(\varepsilon) = \left\{ x_1 x_2 \dots x_n ; \left| -\frac{1}{n} \sum_{i=1}^{n} \log_2 p(x_i) - \mathsf{H}(\xi) \right| \leqq \varepsilon \right\}$$

We denote by G_n the set of equivalence classes of f with the property

$$\left| -\frac{1}{n} \sum_{i=1}^{n} \log_2 p(f^{-1}(f(x_i))) - \mathsf{H}(f(\xi)) \right| \leqq \tau$$

($\tau > 0$ is an arbitrary positive number), where $f^{-1}(f(x_i))$ is evidently the equivalence class of x_i, i.e.

$$(2) \quad G_n = G_n(\tau) = \left\{ f^{-1}(f(x_1 x_2 \dots x_n)); \left| -\frac{1}{n} \sum_{i=1}^{n} \log_2 p(f^{-1}(f(x_i))) - \mathsf{H}(f(\xi)) \right| \leqq \tau \right\}$$

thus G_n is the analogon of F_n for $f(\xi)$ and $\tau > 0$.

It follows from the weak law of large numbers that we have $p(F_n) \to 1$ for every $\varepsilon > 0$ if $n \to \infty$ and also $p(G_n) \to 1$ for every $\tau > 0$, since the random variables

$$\eta_i = -\log_2 p(x) \quad \text{if} \quad \xi_i = x$$

are independent and identically distributed with the common expected value $\mathsf{H}(\xi)$ and the random variables

$$\zeta_i = -\log_2 p(f^{-1}(f(x))) \quad \text{if} \quad f(\xi_i) = f(x)$$

are independent and identically distributed with common expected value $\mathsf{H}(f(\xi))$.

Let $L_n(\varepsilon, \tau)$ be the intersection of $F_n(\varepsilon)$ and $G_n(\tau)$, i.e. $L_n = F_n \cap G_n$. Thus, for all $\varepsilon > 0$ and $\tau > 0$ we have $p(L_n) \to 1$. As the source is memoryless, i.e. $p(x) = p(x_1 x_2 \dots x_n) = \prod_{i=1}^{n} p(x_i)$, from (1) follows

$$(3) \qquad 2^{-n[\mathsf{H}(\xi)+\varepsilon]} \leqq p(x) \leqq 2^{-n[\mathsf{H}(\xi)-\varepsilon]} \quad \text{if} \quad x \in F_n.$$

Likewise, from (2), we get

$$(4) \qquad 2^{-n[\mathsf{H}(f(\xi))+\tau]} \leqq p(f^{-1}(f(x))) \leqq 2^{-n[\mathsf{H}(f(\xi))-\tau]} \quad \text{if} \quad x \in G_n.$$

Since $L_n = F_n \cap G_n$, the bounds (3) hold for all $x \in L_n$. If we denote by $y(x) = y_{\varepsilon,\tau}(x)$ the set $L_n \cap f^{-1}(f(x))$ for all $x \in L_n$, we obtain from (4)

$$(5) \qquad p(y(x)) \leqq 2^{-n[\mathsf{H}(f(\xi))-\tau]} \quad \text{if} \quad x \in L_n.$$

Since for every $x \in L_n$ we have $y(x) \subset F_n$, from (3) follows

$$(6) \qquad p(y(x)) \geqq 2^{-n[\mathsf{H}(\xi)+\varepsilon]} \cdot \|y(x)\|.$$

Comparing (5) and (6) we obtain for the number of elements of the set $y(x)$ the following upper bound:

$$\|y(x)\| \leqq 2^{n[\mathsf{H}(\xi) - \mathsf{H}(f(\xi)) + \varepsilon + \tau]}.$$

Since $\mathsf{H}(\xi)-\mathsf{H}(f(\xi)) = \mathsf{H}(\xi, f(\xi))-\mathsf{H}(f(\xi)) = \mathsf{H}(\xi|f(\xi))$, this can be written also as

$$\|y(x)\| \leqq 2^{n[\mathsf{H}(\xi|f(\xi))+\varepsilon+\tau]} \quad \text{for every} \quad \varepsilon, \tau > 0. \tag{7}$$

As $\lim_{n\to\infty} p(L_n)=1$, we get for every sufficiently large n and $0 \leqq \lambda < 1$ that $M(n,\lambda) \leqq$ $\leqq [L_n]$. To obtain an upper bound for $[L_n]$ we give a suitable „good" decomposition of L_n.

As we know, $f(\xi)$ defines a partition of L_n. Two sequences $a_1 a_2 \dots a_n$ and $b_1 b_2 \dots b_n$ belong to the same class if and only if $f(a_1)\dots f(a_n)=f(b_1)\dots f(b_n)$. We choose a class of this partition the number of elements of which is maximum. Let us choose to every element of this maximal class certain elements from the other classes of $f(\xi)$. More precisely: to different elements of the maximal class we choose different elements from every equivalence class, respectively. If in course of this procedure all elements of a certain class are exhausted, we neglect that class further on. So, to every element in the maximal class we make correspond a "good" set of elements of L_n, i.e. a set, where for every a and b $a|b$ does not hold. Since the number of elements of the chosen class was maximum, every element of L_n is contained in one of our "good" sets. Therefore we have got a "good" decomposition of L_n. The number of components of this decomposition is equal to the number of elements of a maximal class, which is not greater than $2^{n[\mathsf{H}(\xi|f(\xi))+\varepsilon+\tau]}$, according to (7). So,

$$M(n, \lambda) \leqq [L_n] = [L_n(\varepsilon, \tau)] \leqq 2^{n[\mathsf{H}(\xi|f(\xi))+\varepsilon+\tau]} \quad \text{for every} \quad \varepsilon, \tau > 0,$$

whence

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log_2 M(n, \lambda) \leqq \mathsf{H}(\xi|f(\xi))+\varepsilon+\tau \quad \text{for every} \quad 0 \leqq \lambda < 1 \text{ and } \varepsilon, \tau > 0.$$

Hence

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log_2 M(n, \lambda) \leqq \mathsf{H}(\xi|f(\xi)) \tag{8}$$

holds, too.

We have now to prove an inequality in the opposite direction. We do it again through the set L_n.

If $E \subset X^n$, we can state $[E] \geqq [E \cap L_n]$, since every "good" decomposition of E generates a "good" decomposition of $E \cap L_n$. If $\delta > 0$, then for n sufficiently large we get $p(L_n) \geqq 1-\delta$, so if $p(E) \geqq 1-\lambda$, we also have $p(E \cap L_n) \geqq 1-\lambda-\delta$.

On the other hand we estimate $p(E \cap L_n)$ by one of the "good" decompositions with minimum number of components. Substituting the probabilities of the components of such a decomposition by their maximum possible value, we get

$$p(E \cap L_n) \leqq [E \cap L_n] \cdot \max \{p(A); A \subset E \cap L_n, (a, b \in A, a \neq b \Rightarrow a \perp b)\}.$$

The inequality

$$\max \{p(A);\ A \subset E \cap L_n, A \text{ is a "good" set}\} \leqq$$
$$\leqq \max_{x \in L_n} p(x) \cdot \max \{\|A\|;\ A \subset E \cap L_n,\ A \text{ is a "good" set}\}$$

is trivial.

From this we deduce

(9) $$p(E\cap L_n) \leqq [E\cap L_n]\cdot \max_{x\in L_n} p(x)\cdot \max\{\|A\|;\ A \text{ is a „good" set}\}.$$

As $x\in L_n$, we get from (3)

$$p(x) \leqq 2^{-n[\mathsf{H}(\xi)-\varepsilon]}.$$

Let N_n denote the number of the equivalence classes of f contained in L_n; then N_n is an upper bound for $\max\{\|A\|;\ A \text{ is a "good" set}\}$, since if A contained two different elements of the same equivalence class, for these elements we would have $a|b$. Hence we have

(10) $$p(E\cap L_n) \leqq 2^{-n[\mathsf{H}(\xi)-\varepsilon]}\cdot N_n\cdot [E\cap L_n].$$

Now, because of the inequalities $1 \geqq p(G_n) \geqq N_n\cdot 2^{-n[\mathsf{H}(f(\xi))+\tau]}$, (where the second one follows from (2)), we get for N_n the bound

(11) $$N_n \leqq 2^{n[\mathsf{H}(f(\xi))+\tau]}.$$

Finally, the inequalities (10)—(11) result in

$$p(E\cap L_n) \leqq 2^{-n[\mathsf{H}(\xi|f(\xi))-\varepsilon-\tau]}\cdot [E\cap L_n],$$

whence

$$[E\cap L_n] \geqq (1-\lambda-\delta)\cdot 2^{n[\mathsf{H}(\xi|f(\xi))-\varepsilon-\tau]}.$$

Consequently

$$M(n,\lambda) = \min_{\mathsf{P}(\xi_1\ldots\xi_n\in E)\geqq 1-\lambda} [E] \geqq (1-\lambda-\delta)\cdot 2^{n[\mathsf{H}(\xi|f(\xi))-\varepsilon-\tau]},$$

thus

$$\frac{1}{n}\log_2 M(n,\lambda) \geqq \frac{1}{n}\log_2(1-\lambda-\delta)+\mathsf{H}(\xi|f(\xi))-\varepsilon-\tau.$$

Since the first summand on the right hand side tends to 0, if $n\to\infty$, we get

$$\varliminf_{n\to\infty}\frac{1}{n}\log_2 M(n,\lambda) \geqq \mathsf{H}(\xi|f(\xi))-\varepsilon-\tau \quad \text{for all} \quad \varepsilon,\tau>0$$

i.e.

(12) $$\varliminf_{n\to\infty}\frac{1}{n}\log_2 M(n,\lambda) \geqq \mathsf{H}(\xi|f(\xi)).$$

From (8) and (12) we obtain

$$\lim_{n\to\infty}\frac{1}{n}\log_2 M(n,\lambda) = \mathsf{H}(\xi|f(\xi)),$$

what we wanted to prove.

Mathematical Institute of the Hungarian Academy of Sciences

(Received May 18, 1970.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 361—365.

ON A GENERALIZATION OF FARKAS THEOREM

by

L. REJTŐ

To the memory of A. Rényi

1. Let H be an arbitrary finite set with elements $x_1, \ldots, x_n$, and let $f_1, \ldots, f_k$ be real valued functions on H, and $c_1, \ldots, c_k$ real numbers. We seek the conditions for the existence of a σ-algebra and a finite measure μ for which

$$\int_H f_i(x)\mu(dx) = c_i \quad \text{for} \quad i=1, \ldots, k.$$

Let

$$A = \begin{pmatrix} f_1(x_1) & f_1(x_2) \ldots f_1(x_n) \\ \vdots & \vdots & \vdots \\ f_k(x_1) & f_k(x_2) \ldots f_k(x_n) \end{pmatrix}$$

and

$$\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \quad c = \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix}.$$

We look for a necessary and sufficient condition for the existence of a vector μ, with nonnegative components, for which $A\mu = c$.

A. RÉNYI and J. CSIMA dealt with this problem. A necessary and sufficient condition was given by J. CSIMA [2] for the case when the elements of A and the components of c are nonnegative. A. RÉNYI recognized that the problem is equivalent to the FARKAS theorem.

THEOREM (Farkas) *Let A be an $n \times k$ matrix and let $c \in R^k$. (R^k denotes the k-dimensional Euclidean space.) For the existence of a vector $\mu \in R^n$ with nonnegative components, such that $A\mu = c$, it is necessary and sufficient that $c\lambda \geqq 0$ provided that the components of $A^*\lambda$ are nonnegative.* (A^* denotes the transpose of A). (See [3], pp. 124—125.)

In the case of finite H, the FARKAS theorem gives a complete answer of the problem.

The next question is due to A. RÉNYI. What can we say about the general case, when the elements of H, and the number of functions on H are not finite?

Let $\mathscr{X}$ be an arbitrary set and the functions $F_y(x)$, $y \in Y$ be real valued functions on $\mathscr{X}$. Let $\varphi(y)$ be a real valued function on the set of indices Y. What are the conditions for the existence of a finite regular Borel measure μ, such that

$$\int_{\mathscr{X}} F_y(x)\mu(dx) = \varphi(y) \quad \text{for all} \quad y \in Y?$$

The purpose of this paper is to give a necessary and sufficient condition for the case when Y is a Hilbert space.

2. Let E be a locally convex topological vector space, and let E' be the set of continuous linear functionals on E. The elements of E' are real valued functions if E is a vector space over the real number and they are complex valued functions if E is a vector space over the complex numbers.

Definition: We define the weak* topology on E', by a local base of an $f_0 \in E'$. A local base is:

$$U(f_0;\, x_1, \dots, x_n;\, \varepsilon) = \{f\colon\, |f(x_k) - f_0(x_k)| < \varepsilon, \quad k = 1 \dots n\}$$

where $x_1, \dots, x_n$ are arbitrary elements of E, and $\varepsilon > 0$.

Definition: Suppose that X is a nonempty subset of a locally convex topological vectorspace E, and that μ is a nonnegative finite regular Borel measure on X. A point x in E is said to be represented by μ if $f(x) = \int\limits_X f(t)\mu(dt)$ for all $f \in E'$.

Let $\mathcal{X}'$ be the space of real valued linear functionals on $\mathcal{X}$. Let us consider the following problem: find a linear functional G on $\mathcal{X}''$ such that

$$G\bigl(F_y(x)\bigr) = \varphi(y) \quad \text{for all} \quad y \in Y.$$

To answer this problem we have:

THEOREM ([4] pp. 31). *Given a normed linear space $\mathcal{X}$, a collection of elements $\{x_\alpha;\, \alpha \in \mathcal{Y}\} \in \mathcal{X}$ and a collection of real numbers $\{c_\alpha;\, \alpha \in \mathcal{Y}\}$, a necessary and sufficient condition for the existence of a bounded linear functional G such that $G(x_\alpha) = c_\alpha$ for all $\alpha \in \mathcal{Y}$, and $\|G\| \leqq M$ is that the inequality*

$$\Bigl\|\sum_{\alpha \in \pi} \beta_\alpha c_\alpha\Bigr\| \leqq M \Bigl\|\sum_{\alpha \in \pi} \beta_\alpha x_\alpha\Bigr\|$$

holds for each finite subset π of $\mathcal{Y}$ and for every choice of the real numbers β_α.

If we suppose that a measure ν and a σ-algebra $\mathfrak{S}$ exist on $\mathcal{X}$ and $\{F_y(x)\colon y \in Y\} \subset L_2(\mathcal{X};\nu)$ then the above theorem is applicable, i.e. there exists a measure μ such that

$$\int\limits_{\mathcal{X}} F_y(x)\mu(dx) = \varphi(y),$$

but the measure μ is not necessarily finite.

The next theorem gives a necessary and sufficient condition for the case when Y is a Hilbert space.

THEOREM A. *Let $\mathcal{X}$ be an arbitrary set and $\{Y;\, \mathfrak{S};\, \lambda\}$ a σ-finite measure space $\lambda \not\equiv 0$. Let $\varphi(y) = \varphi_y$ and $F(x;y) = F_x(y) = F_y(x)$ be real valued functions on $\mathcal{Y}$ and $\mathcal{X} \times \mathcal{Y}$ respectively. Suppose that $\varphi(y)$ and the functions $F_x(y)$, are elements of $L_2(\mathcal{Y}, \lambda)$, for all fixed x, further suppose that the set $G = \{F_x(y)\colon\, x \in \mathcal{X}\}$ is weak* compact. Then there exist a σ-algebra and a finite regular measure μ on $\mathcal{X}$, such that*

$$\int\limits_{\mathcal{X}} F_y(x)\mu(dx) = \varphi_y \quad \textit{for all} \quad y \in \mathcal{Y},$$

if and only if the inequality

$$\int_{\mathcal{Y}} f(y) F_x(y) \lambda(dy) \geqq 0, \quad \text{where} \quad f(y) \in L_2(\mathcal{Y}, \lambda)$$

implies

$$\int_{\mathcal{Y}} f(y) \varphi(y) \lambda(dy) \geqq 0.$$

We first prove another theorem and Theorem A follows as a special case of this.

THEOREM B. *Suppose that P is a weak* compact subset of a real Hilbert space $\mathcal{H}$, and that q is a fixed element of $\mathcal{H}$. Then there exist a σ-algebra and a finite regular measure μ on P such that*

$$\int_P (p, y) \mu(dp) = (q, y) \quad \text{for all} \quad y \in \mathcal{H}$$

if and only if the inequality

$$(p, x) \geqq 0 \quad \text{for all} \quad p \in P$$

implies

$$(q, x) \geqq 0.$$

To prove Theorem B we need the following lemmas.

LEMMA 1. (see [1], 5 p.) *Suppose that Y is a compact subset of a locally convex topological vector space E. A point x in E is in the closed convex hull X of Y if and only if there exists a probability measure μ on Y which represents x.*

LEMMA 2. *Let F be a compact subset of a topological vectorspace E. Suppose that the cone L having the base $\overline{\mathcal{C}(F)}$, (which denotes the closed convex hull of F), is closed. Let K be the closure of the convex hull of the cone with the base F. Then $K = L$.*

PROOF: Since $F \subset \overline{\mathcal{C}(F)}$, the cone L contains the cone with the base F. It is easy to verify that the cone with base $\mathcal{C}(F)$, equals the convex hull of the cone with base F. Thus the cone with base $\mathcal{C}(F)$ is a susbet of L. Hence L is closed an $K \subset L$. Since K is a convex closed set, we have

$$\overline{\mathcal{C}(F)} \subset K.$$

Hence $L \subset K$. The proof of Lemma 2. is complete.

PROOF of Theorem B. We first prove the sufficiency of the conditions. Let

$$H = \{x\colon (x, p) \geqq 0 \ \text{ for all } \ p \in P\},$$
$$L = \{y\colon (y, x) \geqq 0 \ \text{ for all } \ x \in H\},$$

and K be the convex hull of the cone with base P and $\overline{K}$ be the closure of K. We prove that $\overline{K} = L$. L is a nonempty convex cone, since $q \in L$ and $P \subset L$. Further if $y_1; y_2 \in L$, $c > 0$, $1 > \lambda > 0$ then $\lambda y_1 + (1-\lambda) y_2 \in L$; $c y_1 \in L$. Hence $K \subset L$. If L is a closed set then $\overline{K} \subset L$ holds. We shall prove that L is weak* closed, and from this follows that L is a closed set in the Hilbert space $\mathcal{H}$. It is enough to show that $\mathcal{H}-L$ is weak* open. It is known that the conjugate space of $\mathcal{H}$ is $\mathcal{H}$. The sets

$$U(x;\, y_1, \dots, y_n;\, \varepsilon) = \{z\colon\ |(y_i, x-z)| < \varepsilon \quad i=1\dots n\}$$

form a local base at x in the weak* topology, for arbitrary $y_1, \dots, y_n$ and $\varepsilon>0$. Let v be an element of $\mathcal{H}-L$. Then there exists an $x_0\in\mathcal{H}$ such that

$$-\delta=(v, x_0)<0 \qquad (\delta>0).$$

Let us consider the following neighbourhood of v

$$U\left(v;\, x_0;\, \frac{\delta}{2}\right) = \left\{z: |(x_0, v-z)| < \frac{\delta}{2}\right\}.$$

Let z be an arbitrary element of $U\left(v;\, x_0;\, \frac{\delta}{2}\right)$. Then the scalar product (z, x_0) will be negative since

$$-\frac{3}{2}\delta < (x_0, z) < -\frac{\delta}{2}.$$

Hence $z\notin L$, and L is weak* closed. To prove the converse, let us suppose that there exists an element of L, y_0, such that $y_0\notin\overline{K}$. In a Hilbert space a closed convex set M and a point $x\notin M$, can be separated, such that there exists an element a of $\mathcal{H}$, such that

$$(a, x)<0$$

and

$$(a, y)\geqq 0 \quad \text{for all } y\in M.$$

Hence, there exists an element z_0 such that

$$(y_0, z_0)<0$$

and

$$(x, z_0)\geqq 0 \quad \text{for all } x\in\overline{K}.$$

It follows that $(z_0, p)\geqq 0$ for all $p\in P$, and therefore $z_0\in H$. Hence $(y_0, z_0)>0$, which is a contradiction. This implies that $\overline{K}=L$. It follows from Lemma 2, that L is equal to the cone generated by the closure of the convex hull of P. It is known that $q\in L$. Hence there exists a real number β, $0<\beta<+\infty$, such that βq is an element of the closure of the convex hull of P. P is weak* compact. From Lemma 1 it follows that $q\in\overline{\mathscr{C}(P)}$ if and only if there exists a finite regular measure μ on P which represents q. In our case this means that $(q, y) = \int\limits_P (p, y)\mu(dp)$ for all $y\in\mathcal{H}$.

The necessity of the condition is easy to verify. If $\int\limits_P (p, y)\mu(dp)=(q, y)$ for all $y\in\mathcal{H}$ and if there exists an x such that $(p, x)\geqq 0$ for all $p\in P$, then

$$(q, x) = \int\limits_P (p, x)\,\mu(dp).$$

Hence $(q, x)\geqq 0$, since μ is a nonnegative measure.

PROOF of Theorem A:
Theorem B implies that:

$$\int_{\mathscr{Y}} \varphi(y) f(y) \lambda(dy) = \int_{\mathscr{X}} \left\{ \int_{\mathscr{Y}} F(x;y) f(y) \lambda(dy) \right\} \mu(dx)$$

for all $f(y) \in L_2(Y, \lambda)$.

It is known that μ is a finite regular measure. Let $\mu(\mathscr{X}) = \alpha$. Then

$$\int_{\mathscr{X}} \left\{ \int_{\mathscr{Y}} f(y) \left[F(x;y) - \alpha\varphi(y) \right] \lambda(dy) \right\} \mu(dx) = 0.$$

The order of integration can be interchanged, by using the theorems of Tonelli and Fubini (see [5] p. 194). Therefore

$$\int_{\mathscr{Y}} f(y) \left\{ \int_{\mathscr{X}} \left[F(x;y) - \alpha\varphi(y) \right] \mu(dx) \right\} \lambda(dy) = 0$$

for all $f(y) \in L_2(Y, \lambda)$. From this it is obvious that

$$\int_{\mathscr{X}} F(x;y) \mu^*(dx) = \varphi(y)$$

where

$$\mu^* = \frac{1}{\alpha} \mu.$$

Remark: The condition of Theorem A, (that $\{\mathscr{Y}; \mathfrak{S}; \lambda\}$ is a σ-finite measure space) was used when we applied the theorems Tonelli and Fubini.

REFERENCES

[1] PHELPS, R.: *Lectures on Choquet's theorem.* Van Nostrand, Princeton, N. J. 1966.
[2] Combinatorial Theory and its Application. Colloquia Mathematica Soc. J. B. IV.
[3] Черников: Линейные неравенства. Издательство ,,Наука" Москва 1968.
[4] HILLE—PHILIPS: *Functional Analysis and Semi-Groups.* Coll. Publ. Amer. Math. Soc. Vol. 31. New York 1957.
[5] DUNFORD—SCHWARTZ: *Linear operators* (part I) Interscience Publishers, New York 1958.

Mathematical Institute of the Hungarian Academy of Sciences

(Received June 20, 1970)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 367—374.

ON EXPANSIONS IN ORTHOGONAL POLYNOMIALS

by

G. FREUD

I. Introduction

Let us denote by $w(x)$ a nonnegative weight function with finite support $[-1, +1]$, so that for some $m_0 > 0$

(1) $$w(x) \geqq m_0(1-x^2)^{1/2} \qquad (-1 \leqq x \leqq 1).$$

Let $p_\nu(w; x) = \gamma_\nu(w)x^\nu + \ldots$ be the sequence of orthonormal polynomials with respect to the weight $w(x)$, let

(2) $$f(x) \sim \sum a_\nu(f) p_\nu(w; x)$$

be the orthogonal expansion of a function $f \in \mathscr{L}(w)$, and finally let $s_\nu(w; f; x)$ be the partial sum of (formal) degree $\nu - 1$ of (2). In what follows we give an estimate of $\|f - s_\nu(w; f)\|_C$ under the condition

(3) $$f^{(r)} \in \mathrm{Lip}\, \varrho,$$

where r is some nonnegative integer and $0 < \varrho \leqq 1$ (see Theorem I below). If $r = 0$, i.e. $f \in \mathrm{Lip}\, \varrho$ we need concerning $w(x)$ the assumption that

(4) $$\int_{-1}^{+1} (1-x^2)^{-1+\varrho}\, w(x)\, dx < \infty.$$

For the case of Legendre polynomials, i.e. $w(x) \equiv 1$ our estimate is a refinement of a result of P. K. SUETIN [8] (see also T. H. GRONWALL [6], A. S. DZAFAROV [3]). It is a counterpart of a theorem due to J. SHOHAT where $w(x) \geqq m \geqq 0$ was assumed.[1] The statements of J. SHOHAT are weaker than ours, but for $r = 0$ no additional condition like (4) is needed. If $r \geqq 1$ our result covers that of SHOHAT's without exceptions. The simplicity of proof may be of interest. In Theorem II (part III) we give a refinement of the main result and Theorem III (part IV) covers the most important weights for which our result applies.

II. The main theorem

LEMMA. *Under the condition* (1) *we have*

(5) $$\sum_{\nu=0}^{n-1} p_\nu^2(w; x) \leqq \frac{2}{\pi m_0}\, n(1-x^2)^{-1} \qquad (-1 < x < 1)$$

[1] See for details remark to Theorem IV. 5.1. in the book [4] of the author.

PROOF. The polynomials with respect to the weight $W(x) = m_0\sqrt{1-x^2}$ are $p_\nu(W; x) = \sqrt{\frac{2}{\pi m_0}}\, U_\nu(x)$, where $U_\nu(\cos\Theta) = \frac{\sin(\nu+1)\Theta}{\sin\Theta}$, so that $|U_\nu(x)| \leqq (1-x^2)^{-1/2}$. We conclude from $w(x) \geqq W(x)$ by [4] Theorem I. 4. 2 and formula I. (4. 7)

$$\sum_{\nu=0}^{n-1} p_\nu^2(w; x) \leqq \sum_{\nu=0}^{n-1} p_\nu^2(W; x) \leqq \frac{2}{\pi m_0} n(1-x^2)^{-1}$$

Q.e.d.

THEOREM I. *Under the conditions* (1), (3) *and either* (4) *or* $r \geqq 1$ *we have*

$$|f(\xi) - s_{n-1}(w; f; \xi)| \leqq C(r, \varrho) n^{-r-\varrho}\left(m_0^{-1/2} n^{1/2} + \sqrt{|p_{n-1}(w; \xi) p_n(w; \xi)|}\right) \tag{6}$$

$$(-1 \leqq \xi \leqq 1;\ n = 1, 2, \ldots)$$

PROOF. By a refinement of the standard theorem on polynomial approximation (see S. A. TELIAKOVSKI [10] and J. G. GOPENGAUS [5]) there exists a polynomial $\pi_{n-1}(x)$ whose degree does not exceed $n-1$ and for which we have

$$|f(x) - \pi_{n-1}(x)| \leqq A(r, \varrho)(1-x^2)^{1/2(r+\varrho)} n^{-r-\varrho} \qquad (-1 \leqq x \leqq 1). \tag{7}$$

Let

$$K_n(w; x, \xi) = \sum_{\nu=0}^{n-1} p_\nu(w; \xi) p_\nu(w; x)$$

be the Dirichlet kernel of the expansion (2). We have then by (7) and the Schwarz inequality

$$|s_n(w; f; \xi) - \pi_{n-1}(\xi)| = \left|\int_{-1}^{+1} K_n(w; x, \xi)[f(x) - \pi_{n-1}(x)] w(x)\, dx\right| \leqq \tag{8}$$

$$\leqq A(r, \varrho) n^{-\varrho-r} \int_{-1}^{+1} |K_n(w; x, \xi)| (1-x^2)^{1/2(\varrho+r)} w(x)\, dx \leqq$$

$$\leqq A(r, \varrho) n^{-\varrho-r} \left\{\int_{-1}^{+1} [K_n(w; x, \xi)]^2 (1-x^2) w(x)\, dx \int_{-1}^{+1} (1-x^2)^{\varrho+r-1} w(x)\, dx\right\}^{1/2}.$$

The second integral in curly brackets is bounded as a consequence of (4) resp. of $r \geqq 1$ and $\varrho > 0$. The first integral we estimate as follows:

$$\int_{-1}^{+1} [K_n(w; x, \xi)]^2 (1-x) w(x)\, dx =$$

$$= \int_{-1}^{+1} [K_{n+1}(w; x, \xi) - p_n(w; \xi) p_n(w; x)] (1-x) K_n(w; x; \xi) w(x)\, dx =$$

$$= (1-\xi) K_n(w; \xi, \xi) - p_n(w; \xi) \int_{-1}^{+1} p_n(w; x)(1-x) \sum_{\nu=0}^{n-1} p_\nu(w; \xi) p_\nu(w; x) w(x)\, dx =$$

$$= (1-\xi) K_n(w; \xi, \xi) + p_n(w; \xi) p_{n-1}(w; \xi) \int_{-1}^{+1} x p_{n-1}(w; x) p_n(w; x) w(x)\, dx.$$

We have by [4] Lemma I. 7. 2.

$$0 < \int_{-1}^{+1} x p_{\nu-1}(w; x) p_\nu(w; x) w(x)\, dx \leqq 1$$

so that

$$\text{(9.a)} \quad \int_{-1}^{+1} [K_n(w; x, \xi)]^2 (1-x) w(x)\, dx \leqq (1-\xi) \sum_{\nu=0}^{n-1} p_\nu^2(w; \xi) + |p_{n-1}(w; \xi) p_n(w; \xi)|$$

and by symmetry

$$\text{(9.b)} \quad \int_{-1}^{+1} [K_n(w; x, \xi)]^2 (1+x) w(x)\, dx \leqq (1+\xi) \sum_{\nu=0}^{n-1} p_\nu^2(w; \xi) + |p_{n-1}(w; \xi) p_n(w; \xi)|.$$

From (9.a) and (9.b) we have by our Lemma 1

$$\text{(10)} \quad \int_{-1}^{+1} [K_n(w; x, \xi)]^2 (1-x^2) w(x)\, dx \leqq$$

$$\leqq 2 \min \left\{ \int_{-1}^{+1} [K_n(w; x, \xi)]^2 (1-x) w(x)\, dx, \int_{-1}^{+1} [K_n(w; x, \xi)]^2 (1+x) w(x)\, dx \right\} \leqq$$

$$\leqq 2(1-\xi^2) \sum_{\nu=0}^{n-1} p_\nu^2(w; \xi) + 2|p_{n-1}(w; \xi) p_n(w; \xi)| \leqq \frac{4}{\pi m} n + 2|p_{n-1}(w; \xi)|\, |p_n(w; \xi)|.$$

From (7), (8) and (10) we obtain

$$\text{(11)} \quad |f(\xi) - s_n(w; f; \xi)| \leqq A(r, \varrho) n^{-r-\varrho} \cdot$$

$$\cdot \left\{ 1 + \sqrt{\left[\left[\frac{4}{\pi m} n + 2|p_{n-1}(w; \xi)|\, |p_n(w; \xi)|\right] \int_{-1}^{+1} (1-x^2)^{\varrho+r-1} w(x)\, dx\right]} \right\}$$

and (11) implies (6), Q.e.d.

III. Refinement of the main theorem

LEMMA 2. *If for a fixed* $x_1 \in [-1, +1]$

$$\text{(12)} \quad w(x) \geqq m_1 (1-x)^\gamma \qquad (x \in [x_1, 1])$$

then we have uniformly in $[x_1 + \varepsilon, 1]$ *for every* $\varepsilon > 0$ $(x_1 + \varepsilon < 1)$

$$\text{(13)} \quad (1-x) \sum_{\nu=0}^{n-1} p_\nu^2(w; x) = \begin{cases} O(n) & \gamma \leqq 1/2 \\ O(n^{2\gamma}) & \gamma > 1/2 \end{cases}$$

PROOF. The orthonormal polynomials belonging to the weight

$$W_1(x) = \begin{cases} m_1(1-x)^\gamma & x \in [x_1, 1] \\ 0 & x \notin [x_1, 1] \end{cases}$$

are

$$p_\nu(W_1; x) = \frac{1}{\sqrt{m_1}} \left(\frac{2}{1-x_1}\right)^{\frac{\gamma+1}{2}} p_\nu^{(\gamma, 0)} \left[1 - \frac{2}{1-x_1}(1-x)\right],$$

where $p_\nu^{(\gamma, 0)}(x)$ are the orthonormal Jacobi polynomials corresponding to the weight $(1-x)^\gamma$. From standard estimates on Jacobi polynomials we obtain[2]

$$p_\nu^{(\gamma, 0)}(x) = O(1) \min \{n^{\gamma+1/2}, (1-x)^{-\gamma/2-1/4}\}$$

uniformly in $[-1+\varepsilon, +1]$. We conclude

(14) $$p_\nu(W_1; x) = O(1) \min \{n^{\gamma+1/2}, (1-x)^{-\gamma/2-1/4}\}$$

uniformly in $[x_0+\varepsilon, +1]$.

From (12) we get using (14) and Theorem I. 4. 2 of [ ]

$$(1-x) \sum_{\nu=0}^{n-1} p_\nu^2(w; x) \leqq (1-x) \sum_{\nu=0}^{n-1} p_\nu^2(W_1; x) =$$

$$= O(1) \min [(1-x)n^{2\gamma+2} n(1-x)^{-\gamma+1/2}] = \begin{cases} O(n) & (\gamma \geqq 1/2) \\ O(n^{2\gamma}) & (\gamma < 1/2) \end{cases}$$

uniformly in $x \in [x_0+\varepsilon, 1]$, Q.e.d.

THEOREM II. *Let $w(x)$ satisfy the conditions of Lemma 2, and let either* (3) *be satisfied for $r \geqq 1$ or if it is satisfied for $r=0$, then let further*

(15) $$\int_{-1}^{+1} (1-x)^{\varrho-1} w(x)\, dx < \infty;$$

under this assumptions we have uniformly in $\xi \in [x_0+\varepsilon, +1]$ for every $0 < \varepsilon < 1-x_0$

(16) $$|f(\xi) - s_n(w; f; \xi)| = O(n^{-r-\varrho})\left[n^\tau + \sqrt{|p_{n-1}(w; \xi) p_n(w; \xi)|}\right],$$

where $\tau = 1/2$ for $\gamma \leqq 1/2$ and $\tau = \gamma$ for $\gamma > 1/2$.

PROOF. Let $\Pi_{n-1}(x)$ be the same polynomial as in (7), then we have by (8)

$$s_n(w; f; \xi) - \pi_{n-1}(\xi) =$$

$$= O(n^{-r-\varrho}) \left\{ \int_{-1}^{+1} [K_n(w; x, \xi)]^2 (1-x) w(x)\, dx \int_{-1}^{+1} (1-x)^{\varrho+r-1} w(x)\, dx \right\}^{1/2},$$

[2] The estimate of $p_\nu^{(\gamma, 0)}$ is also a consequence of Lemma 3 below.

so that by (9) and Lemma 2 we have

$$s_n(w;f;\xi)-\pi_{n-1}(\xi) = O(n^{-r-\varrho})\sqrt{n^{2\tau}+|p_{n-1}(w;\xi)|\,|p_n(w;\xi)|}\,.$$

Combining this estimation with (7), we obtain (16), Q.e.d.

By symmetry an analogous theorem holds asserting a uniform estimate $f(\xi)-s_n(w;f;\xi)$ in $[-1, x_0-\varepsilon]$. We leave the formulation to the reader.

IV. An application to a class of weight functions

Both Theorem I and Theorem II can clearly be combined with any estimate of $|p_n(w;x)|$. In this last part of the paper we give just one result of this kind.

LEMMA 3. *If the weight function* $0\leqq w(x)\in\mathscr{L}$ $(x\in[-1,+1])$ *satisfies the following two conditions:*

a) *for suitable* $\gamma\geqq-\frac{1}{2}$ *we have in* $[1-2\delta, 1]$

$$0<m_1\leqq(1-x)^{-\gamma}w(x)\in \mathrm{Lip}\ 1/2; \tag{17}$$

b) *we have*

$$\int_{-1}^{+1}\frac{dx}{(1+x)w(x)}<\infty; \tag{18}$$

then we have

$$|p_n(w;x)|\leqq O(1)\begin{cases} n^{1/2+\gamma} \\ (1-x)^{-1/4-\gamma/2}\end{cases} (x\in[1-\delta, 1]). \tag{19}$$

PROOF. Let

$$W(x)=\begin{cases} w(x) & (x\in[1-2\delta, 1]) \\ w(1-2\delta)(2-2\delta)^{1/2}(1+x)^{-1/2} & (x\in[-1, 1-2\delta]).\end{cases} \tag{20}$$

We estimate $p_n(W;x)$ first. Let

$$W^*(t)=|t|W(1-2t^2)\qquad(-1\leqq t\leqq+1) \tag{21}$$

then by an elementary transformation

$$p_{2n}(W^*;t)=p_n(W;1-2t^2). \tag{22}$$

By (20), (21) and (17) we have

$$W^*(t)=|t|^{1+2\gamma}(1-t^2)^{-1/2}\eta(t)$$

where $\eta(t)\in\mathrm{Lip}\ 1/2$ in $[-1, +1]$.

Applying a theorem of P. K. SUETIN [9] we obtain from this

$$|p_{2n}(W^*;t)|=O(1)\begin{cases} n^{1/2+\gamma} \\ |t|^{-1/2-\gamma}\end{cases} (|t|\leqq 1) \tag{23}$$

From (23) and (22) we get

$$p_n(W;x)=O(1)\begin{cases} n^{1/2+\gamma} \\ (1-x)^{-1/4-\gamma/2}\end{cases} (|x|\leqq 1) \tag{24}$$

We make the transition from (24) to (19) with the aid of a classical argument of J. KOROUS [7].

Expanding $p_n(w; x)$ in terms of $\{p_\nu(W; x)\}$ we obtain

$$p_n(w; x) = p_n(W; x) \int_{-1}^{+1} p_n(w; \xi) p_n(W; \xi) W(\xi)\, d\xi + \int_{-1}^{+1} p_n(w; \xi) K_{n-1}(W; x, \xi) W(\xi) d\xi.$$

By orthogonality we have

$$\int_{-1}^{+1} p_n(w; \xi) K_{n-1}(W; x, \xi) w(\xi)\, d\xi = 0.$$

From the two last formulas we have by subtraction (see also (20))

$$p_n(w; x) = p_n(W; x) \int_{-1}^{+1} p_n(w; \xi) p_n(W; \xi) W(\xi)\, d\xi + \tag{25}$$
$$+ \int_{-1}^{1-2\delta} p_n(w; \xi) K_{n-1}(W; x, \xi) [W(\xi) - w(\xi)]\, d\xi.$$

The integral in the first term can be estimated in consequence of (20) and (24) by

$$\int_{-1}^{+1} |p_n(w; \xi)| (1-\xi)^{-1/4+\gamma/2} (1+\xi)^{-1/2}\, d\xi = \tag{26}$$
$$= O(1) \left\{ \int_{-1}^{+1} p_n^2(w; \xi) w(\xi)\, d\xi \int_{-1}^{+1} \frac{d\xi}{w(\xi)(1+\xi)(1-\xi)^{1/2-\gamma}} \right\}^{1/2} =$$
$$= O(1) \left\{ \int_{1-2\delta}^{1} \frac{d\xi}{w(\xi)(1-\xi)^{1/2-\gamma}} + \int_{-1}^{1-2\delta} \frac{d\xi}{w(\xi)(1+\xi)} \right\}^2,$$

and this expression is bounded as a consequence of (17) and (18). Taking (24) in account, we obtain for the first term in (25)

$$p_n(W; x) \int_{-1}^{+1} p_n(w; \xi) p_n(W; \xi) W(\xi)\, d\xi = O(1) \begin{cases} n^{1/2+\gamma} \\ (1-x)^{-1/4-\gamma/2} \end{cases} \quad (|x| \leqq 1) \tag{27}$$

We turn to the estimation of the second term in (25). We apply the Christoffel—Darboux formula

$$K_{n-1}(W; x, \xi) = \frac{\gamma_{n-1}(W)}{\gamma_n(W)} \frac{p_{n-1}(W; \xi) p_n(W; x) - p_n(W; \xi) p_n(W; x)}{x - \xi}.$$

For $\xi \in [-1, 1-2\delta]$ and $x \in [1-\delta, 1]$, i.e. $|x-\xi| \geqq \delta$ we have using Lemma I. 7. 2 from [4] and (24)

$$|K_{n-1}(W; x, \xi)| \leqq 2\delta^{-1} [|p_n(W; x)|\, |p_{n-1}(W; \xi)| + |p_{n-1}(W; x)|\, |p_n(W; \xi)|] =$$
$$= O(1)[|p_n(W; x)| + |p_{n-1}(W; x)|] (1-\xi)^{-1/4-\gamma/2}.$$

So that by (26) and a repeated use of (24)

$$(28)\qquad \left|\int_{-1}^{1-2\delta} p_n(w;\xi)K_{n-1}(W;x,\xi)W(\xi)\,d\xi\right| = O(1)\begin{cases} n^{1/2+\gamma} \\ (1-x)^{-1/4-\gamma/2} \end{cases} \quad (x\in[1-\delta,1])$$

and (applying Schwarz's inequality in the third line)

$$(29)\qquad \left|\int_{-1}^{1-2\delta} p_n(w;\xi)K_{n-1}(W;x,\xi)w(\xi)\,d\xi\right| =$$

$$= O(1)[|p_n(W;x)|+|p_{n-1}(W;x)|]\int_{-1}^{+1}|p_n(w;\xi)|\,w(\xi)\,d\xi =$$

$$= O(1)[|p_n(W;x)|+|p_{n-1}(W;x)|] = O(1)\begin{cases} n^{1/2+\gamma} \\ (1-x)^{-1/4-\gamma/2} \end{cases} \quad (x\in[1-\delta,1])$$

From (25), (27), (28) and (29) follows (19), Q.e.d.

THEOREM III. *Let the weight function $w(x)$ satisfy the conditions of Lemma 3 and let*

$$(28)\qquad w(x)\geqq m_2 \qquad [x_0\leqq x\leqq 1-\delta],$$

then provided that $f(x)$ satisfies the same conditions as in Theorem II, we have uniformly in $\xi\in[x_0+\varepsilon,1]$ for every $0<\varepsilon<1-x$.

$$f(\xi)-s_n(w;f;\xi) = O(n^{-r-\varrho+\sigma})$$

where $\sigma = 1/2$ if $\gamma\leqq 0$ and $\sigma = 1/2+\gamma$ if $\gamma>0$.

We remark that for $w(x) = (1-x)^\gamma(1+x)^\beta$ and $\gamma\geqq 0$ our estimate is best possible in order (see S. A. AGAHANOV—G. I. NATANSON [1]) and that under much more restrictive conditions on the weight function, a similar but more precise statement was proved by V. M. BADKOV [2].

LITERATURE

[1] Агаханов, С. А. — Намансон, Г. И.: Ппиближение функции суммами фурье — Якоби, *Доклады А. Н. СССР* **166** (1966) 9—10.
[2] Бадков, В. М.: Приблжение функции частными суммами ряда фурье по обобщенным многочленам Якоби. *Матем. Замитки*, **3** (1968) 671—682.
[3] Джафаров, Ариф С.: Приближение функции на сфере аналогами сумм фейера и Валле—Пуссена, *Известия А. Н. Аз. ССР Сер. физ-мехн. и матем. науки*, **3** (1965) 15—20.
[4] FREUD, G.: *Orthogonale Polynome*, Birkhäuser Verlag Basel 1969.
[5] Гопенгауз, И. З.: К теории А. Ф. Симана о приближении функции многочленами на конечном отрейке. *Матем. Заметки*, **1** (1967) 163—172.
[6] GRONWALL, T. H.: Über die Laplacesche Reihe *Mathem. Annalen* **74** (1913) 213—270.

[7] KOROUS, J.: O rozvaji funkci jedné reálné promenné v radu jistych ortogonálnich polinomu. *Rozpravy České Akademie* (2) **48** (1938), 1—12.
[8] Суетин, П. К.: О представлении непрерывных и дифференцируемых функции рядами фурье но многочленам. *Лежанбра, Доклады А. Н. СССР* **158** (1964) 1275—1277.
[9] Суетин, П. К.: Некостопые свойсгва многочленов, ормогоналньных на сеъменте, *Сибипский Матем. Журпал* **10** (1969) 653—670.
[10] Теляковски, С. А.: Дбе теоремы о приближении функций алгебрагескими многогленам, *Мат. Сбор*. **70** (1966) 252—265.

Mathematical Institute of the Hungarian Academy of Sciences

(Received June 22, 1970)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 375—380.

A CONSTRUCTION OF BROWNIAN MOTION PROCESS IN r-DIMENSION

by

A. FÖLDES

To the memory of A. Rényi

The present paper deals with an explicit construction (in the form of infinite series) of the Brownian motion process. This construction is not the only possible way for proving the existence of Brownian motion process but seems to be favourable for applications.

Let $\{\psi_n\}_{n=0}^{\infty}$ be a complete orthonormal system in $L_2[0, 1]$, and $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ a sequence of independent, standard normally distributed random variables. Then

$$x(\omega, t) = \sum_{n=0}^{\infty} \xi_n(\omega) \int_0^t \psi_n(x)\,dx$$

is a Brownian motion process. It is easier to verify this result for special orthonormal systems. For the trigonometric system it was done by N. WIENER. For arbitrary complete orthonormal system the theorem was proved by ITÔ and NISIO ([3]) in 1968. Here we generalize their result for r-dimension, using partly their method, partly the method of LAMPERTI and RÉNYI (see [1], [2], [3]).

Definition: The stochastic process $x(\omega, \vec{t})$ is an r-dimensional Brownian motion process, if the following conditions hold:

1., $\mathsf{P}(x(\omega, \vec{0}) = 0) = 1$

2., $\Delta\vec{h}_1\, x(\omega, \vec{t}_1),\ \Delta\vec{h}_2\, x(\omega, \vec{t}_2), \ldots, \Delta\vec{h}_r\, x(\omega, \vec{t}_r)$

are independent normally distributed random variables with expectation 0 and variances $V_1, V_2, \ldots, V_r$, if the r-dimensional intervals figuring in the differences above are disjoint, and $V_1, V_2, \ldots, V_r$ denotes the r-dimensional measure of the intervals. Here the meaning of the notation $\Delta\vec{h}x\,(\omega, \vec{t})$ is the following:

$$\Delta\vec{h}\, x(\omega, \vec{t}) = \Delta h_n(\Delta h_{n-1}(\ldots(\Delta h_1\, x(\omega, t_1 \ldots t_n))\ldots))$$

and

$$\Delta h_i\, x(\omega, t_1 \ldots t_n) =$$

$$= x(\omega, t_1, \ldots, t_{i-1}, t_i + h_i, t_{i+1}, \ldots, t_n) - x(\omega, t_1 \ldots t_{i-1}, t_i, t_{i+1} \ldots t_n)$$

3., $x(\omega, \vec{t})$ is a continuous function of r-variables with probability 1.

THEOREM: *Let $\{\psi_n\}_{n=0}^{\infty}$ be a complete orthonormal system in $L_2\left(\bigtimes_{k=1}^{r} [0, 1]\right)$. Further let $\xi_0, \xi_1, \ldots, \xi_n, \ldots$ be a sequence of independent standard normally dis-*

tributed random variables in a probability space $(\Omega, \mathcal{F}, \mathsf{P})$. *In this case*

$$(1)\qquad x(\omega, \vec{t}) = \sum_{n=0}^{\infty} \xi_n(\omega) \int_{\vec{0}}^{\vec{t}} \psi_n(\vec{u})\, d\vec{u}$$

is a Brownian motion process.

PROOF. For the sake of simplicity we prove the theorem for two dimension. First of all we show that (1) converges with probability 1 ([1]).

Let

$$c_n(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} \psi_n(u_1, u_2)\, du_1\, du_2 \qquad (n=0, 1, 2 \dots)$$

Then

$$(2)\qquad x(\omega, t_1, t_2) = \sum_{n=0}^{\infty} c_n(t_1, t_2) \xi_n(\omega)$$

Clearly (2) is an orthonormal series in $L_2(\Omega, \mathcal{F}, \mathsf{P})$ for $\xi_n(\omega)$'s are independent and standard normally distributed. Denote by $e_{t_1, t_2}(x, y)$ the characteristic function of the interval $[0, t_1] \times [0, t_2]$. Clearly the Fourier expansion of $e_{t_1, t_2}(x, y) \in \in L_2([0, 1] \times [0, 1])$ is convergent in the norm of L_2, i.e.

$$e_{t_1,t_2}(x, y) = \sum_{n=0}^{\infty} \Big(\int_0^1 \int_0^1 e_{t_1,t_2}(u_1, u_2) \psi_n(u_1, u_2)\, du_1\, du_2 \Big) \psi_n(x, y) =$$

$$= \sum_{n=0}^{\infty} \Big(\int_0^{t_1} \int_0^{t_2} \psi_n(u_1, u_2)\, du_1\, du_2 \Big) \psi_n(x, y) = \sum_{n=0}^{\infty} c_n(t_1, t_2) \psi_n(x, y).$$

Using Parseval relation we get

$$\int_0^1 \int_0^1 e_{t_1,t_2}(x, y) e_{s_1,s_2}(x, y)\, dx\, dy = \sum_{n=0}^{\infty} c_n(t_1, t_2) c_n(s_1, s_2)$$

But clearly the left-hand side of the above relation is

$$\int_0^{\min(t_1, s_1)} \int_0^{\min(t_2, s_2)} 1\, dx\, dy = \min(t_1, s_1) \min(t_2, s_2)$$

so we get

$$(3)\qquad \sum_{n=0}^{\infty} c_n(t_1, t_2) c_n(s_1, s_2) = \min(t_1, s_1) \min(t_2, s_2).$$

In particularly: $\displaystyle\sum_{n=0}^{\infty} c_n^2(t_1, t_2) = t_1 t_2.$

It follows that the orthonormal series (2) converges in norm, thus the Parseval relation holds.

Especially $\mathsf{E}\big(x^2(\omega, t_1, t_2)\big) = t_1 t_2.$

On the other hand applying the three series theorem for (2) we have

$$\sum_{n=0}^{\infty} \mathsf{E}(c_n(t_1, t_2)\xi_n(\omega)) = 0,$$

$$\sum_{n=0}^{\infty} \mathsf{D}^2(c_n(t_1, t_2)\xi_n(\omega)) = \sum_{n=0}^{\infty} \mathsf{E}(c_n^2(t_1, t_2)\xi_n^2(\omega)) = \sum_{n=0}^{\infty} c_n^2(t_1, t_2) = t_1 t_2 .$$

Hence (2) converges with probability 1. Now we turn to the proof of properties 1—3.

1. It is clear from the definition.

2. We show that the differences taking on disjoint rectangles are independent normally distributed random variables with zero expectation and variances being equal to the measures of the rectangles. For the sake of simplicity we prove it for two rectangles (by rectangle we mean the Cartesian product of two intervals). It is enough to verify the following identity ([2]).

(4) $$\mathsf{E}(\exp[i\{\lambda_1 \Delta k_1 \Delta h_1 x(\omega, s_1, s_2) + \lambda_2 \Delta k_2 \Delta h_2 x(\omega, u_1, u_2)\}]) =$$
$$= \exp\left\{-\frac{1}{2}[(k_1 h_1)\lambda_1^2]\right\} \exp\left\{-\frac{1}{2}[(k_2 h_2)\lambda_2^2]\right\}$$

Namely, here the left-hand side is the characteristic function of the random vector $(\Delta k_1 \Delta h_1 x(\omega, s_1, s_2), \Delta k_2 \Delta h_2 x(\omega, u_1, u_2))$, on the right-hand side stands what we should get if every coordinates were independent normally distributed random variables with zero expectation and variances $k_1 h_1$, $k_2 h_2$ respectively.

To verify (4) we need the following lemma ([3])

LEMMA: *Let $\eta_1, \eta_2, \dots, \eta_n, \dots$ be a sequence of independent random variables, for which $\sum_{n=1}^{\infty} \eta_n$ converges with probability* 1. *In this case.*

$$\mathsf{E}\left(\exp\left[i \sum_{n=1}^{\infty} \eta_n\right]\right) = \prod_{n=1}^{\infty} \mathsf{E}(\exp i\eta_n).$$

So

(5) $$\mathsf{E}(\exp[i\{\lambda_1 \Delta k_1 \Delta h_1 x(\omega, s_1, s_2) + \lambda_2 \Delta k_2 \Delta h_2 x(\omega, u_1, u_2)\}]) =$$
$$= \mathsf{E}\left(\exp\left[i\left\{\lambda_1 \sum_{n=0}^{\infty} \Delta k_1 \Delta h_1 c_n(s_1, s_2)\xi_n + \lambda_2 \sum_{n=0}^{\infty} \Delta k_2 \Delta h_2 c_n(u_1, u_2)\xi_n\right\}\right]\right) =$$
$$= \mathsf{E}\left(\exp\left[i \sum_{n=0}^{\infty} \xi_n\{\lambda_1 \Delta k_1 \Delta h_1 c_n(s_1, s_2) + \lambda_2 \Delta k_2 \Delta h_2 c_n(u_1, u_2)\}\right]\right) =$$
$$= \prod_{n=0}^{\infty} \mathsf{E}(\exp i[\xi_n\{\lambda_1 \Delta k_1 \Delta h_1 c_n(s_1, s_2) + \lambda_2 \Delta k_2 \Delta h_2 c_n(u_1, u_2)\}]) =$$
$$= \prod_{n=0}^{\infty} \exp\left(-\frac{1}{2}\{\lambda_1 \Delta k_1 \Delta h_1 c_n(s_1, s_2) + \lambda_2 \Delta k_2 \Delta h_2 c_n(u_1, u_2)\}^2\right) =$$
$$= \exp\left(-\frac{1}{2}\sum_{n=0}^{\infty}\{\lambda_1 \Delta k_1 \Delta h_1 c_n(s_1, s_2) + \lambda_2 \Delta k_2 \Delta h_2 c_n(u_1, u_2)\}^2\right).$$

By (5) it is enough to prove that

a)
$$\sum_{n=0}^{\infty} \{\Delta k_1\, \Delta h_1\, c_n(s_1, s_2)\}^2 = k_1 h_1$$

$$\sum_{n=0}^{\infty} \{\Delta k_2\, \Delta h_2\, c_n(u_1, u_2)\}^2 = k_2 h_2$$

b)
$$\sum_{n=0}^{\infty} \{\Delta k_1\, \Delta h_1\, c_n(s_1, s_2)\}\{\Delta k_2\, \Delta h_2\, c_n(u_1, u_2)\} = 0$$

are valid.

We can verify a., and b., by the repeated application of (3)

a., It is enough to prove that

$$\sum_{n=0}^{\infty} \{\Delta k\, \Delta h\, c_n(s_1, s_2)\}^2 = kh.$$

This can be seen as follows:

$$\sum_{n=0}^{\infty} \{\Delta k\, \Delta h\, c_n(s_1, s_2)\}^2 = \sum_{n=0}^{\infty} \{\Delta k[\Delta h\, c_n(s_1, s_2)]\}^2 =$$

$$= \sum_{n=0}^{\infty} \{\Delta k[c_n(s_1+h, s_2) - c_n(s_1, s_2)]\}^2 = \sum_{n=0}^{\infty} [\Delta k\, c_n(s_1+h, s_2)]^2 + \sum_{n=0}^{\infty} [\Delta k\, c_n(s_1, s_2)]^2 -$$

$$-2\sum_{n=0}^{\infty} \Delta k\, c_n(s_1+h, s_2)\, \Delta k\, c_n(s_1, s_2) =$$

$$=(s_1+h)[(s_2+k)+s_2-2s_2]+s_1[(s_2+k)+s_2-2s_2]-2s_1[(s_2+k)+s_2-s_2-s_2]=hk.$$

b.,
$$\sum_{n=0}^{\infty} [\Delta k_1\, \Delta h_1\, c_n(s_1, s_2)]\,[\Delta k_2\, \Delta h_2\, c_n(u_1, u_2)] =$$

$$= \sum_{n=0}^{\infty} \{\Delta k_1 [c_n(s_1+h_1, s_2) - c_n(s_1, s_2)]\}\{\Delta k_2 [c_n(u_1+h_2, u_2) - c_n(u_1, u_2)]\}.$$

Since the rectangles are disjoint we may suppose that $s_2+k_1 < u_2$. (In case $s_1+h_1 < u_1$ we choose the opposite order of operations: $\Delta h \Delta k$).

Making the prescribed operations, clearly we get for fixed n 16 products which we order in groups by 4.

E.g.:

$$\sum_{n=0}^{\infty} \Delta k_1\, c_n(s_1+h_1, s_2)\, \Delta k_2\, c_n(u_1+h_2, u_2) =$$

$$= \min(s_1+h_1, u_1+h_2)\,[(s_2+k_1)-(s_2+k_1)+s_2-s_2] = 0$$

In the other 3 groups we take the minimum of various expressions, but the expressions being in the brackets are the same because we always compare the same points: $s_2+k_1, s_2, u_2+k_2, u_2$. More precisely we compare s_2+k_1 with u_2+k_2 and with

u_2 and the result of these two comparison is s_2+k_1 which we have to count once with + once with − sign. In the same way we have to compare s_2 with u_2+k_2 and u_2 respectively, where we get s_2 as the minimum again once with + and once with − sign. Herewith our statement is proved.

3. The proof is based on the extension of P. LÉVY's theorem ([3]). Let E be a separable real Banach space. Denote by $\mathscr{B}$ the σ-algebra of Borel sets of E, and by $\mathscr{P}$ the set of all probability measures defined on $(E, \mathscr{B})$. The basic probability space is denoted by $(\Omega, \mathscr{F}, \mathsf{P})$. An E-valued random variable X is a map of Ω into E which is measurable with respect to $(\mathscr{F}, \mathscr{B})$. The probability law μ_X of X is a probability measure in $(E, \mathscr{B})$ defined by $\mu_X(B)=\mathsf{P}(X\in B)$ $B\in\mathscr{B}$. It is well known that every $\mu\in\mathscr{P}$ is tight ([4]) i.e.: for every $\varepsilon>0$ there exists a compact set $K_\mu\subset E$, for which $\mu(K_\mu)>1-\varepsilon$. A subset $\mathscr{M}$ of $\mathscr{P}$ is called uniformly tight, if for every $\varepsilon>0$ there exists a compact set $K\subset E$ for which $\mu(K)>1-\varepsilon$ for every $\mu\in\mathscr{M}$. Let $X_n(\omega)$ $n=1, 2, \dots$ be a sequence of independent E-valued random variables, $S_n=\sum_{i=1}^{n} X_i$, and μ_n the probability law of S_n.

Then the following theorem holds:

THEOREM (the extension of P. LÉVY's theorem) (see [3]). *The following conditions are equivalent.*

a., S_n *converges with probability* 1.

b., S_n *converges stochastically.*

c., μ_n *converges in Prohorov metric.*

If S_n $(n=1, 2, \dots)$ are symmetrically distributed then also holds:

THEOREM: a., b., c., *are equivalent to the following:* d., $\{\mu_n\}_{n=0}^{\infty}$ *is uniformly tight.*

We shall apply the latter theorem to prove the uniform convergence of (2). (Almost everywhere.) The proof is analogous of the one-dimensional case.

Let E be the space of continuous functions on the unit-square, with the norm $\|f\| = \sup_{0\le t, s\le 1} |f(t, s)|$ i.e. $E = C([0, 1]\times[0, 1])$ In this case E is a separable Banach space where the convergence in norm is equivalent with the uniform convergence.

Let

$$S_n(\omega, t_1, t_2) = \sum_{k=1}^{n} \xi_k(\omega)\int_0^{t_1}\int_0^{t_2} \psi_k(u_1, u_2)\, du_1\, du_2 = \sum_{k=1}^{n} c_k(t_1, t_2)\xi_k(\omega).$$

Since $c_k(t_1, t_2)$ is clearly continuous, for fixed ω $S_n(\omega, t_1, t_2)$ is a continuous function. Therefore it is enough to prove that the E-valued random variable $\mathbf{S}_n(\omega)$ mapping every ω to $S_n(\omega, t_1, t_2)$ converges (in norm) with probability 1.

Since $\xi_k(\omega)$-s are symmetrical, the same holds for $\mathbf{S}_n(\omega)$, so we only have to prove the uniform tightness of the sequence $\{\mu_{\mathbf{S}_n}\}$ i.e. for every $\varepsilon>0$ there exists a $K=K(\varepsilon)$ compact set in E for which

$$\mathsf{P}(\mathbf{S}_n(\omega)\in K) > 1-\varepsilon \tag{6}$$

uniformly in n.

We define the set K as follows:

(7)
$$K = K(\delta) = K(\varepsilon) =$$
$$= \left\{ \begin{array}{l} f;\ f \in C([0,1]\times[0,1])\ f(0,0)=0, \\ \text{for every fixed } y\ |f(x_1,y)-f(x_2,y)| \leqq L|x_1-x_2|^\beta \text{ whenever } |x_1-x_2|<\delta(\varepsilon) \\ \text{for every fixed } x\ |f(x,y_1)-f(x,y_2)| \leqq L|y_1-y_2|^\beta \text{ whenever } |y_1-y_2|<\delta(\varepsilon) \\ 0<\beta<\frac{1}{4} \end{array} \right\}$$

It is easy to prove the compactness of $K(\varepsilon)$. Clearly the elements of $K(\varepsilon)$ are equicontinuous and because of $f(0,0)=0$ they are equibounded too. Therefore by the Arzela—Ascoli theorem one can choose a uniformly convergent subsequence which consists of continuous functions, what (by Weierstrass theorem) converges to a continuous function. So $K(\varepsilon)$ is conditionally compact and closed hence it is compact.

In (7) $L>0$ is arbitrary but fixed.

The fact that $\delta(\varepsilon)$ can be chosen in a suitable way to ε becomes clear on the way of the proof.

Proving (6) we get that in (7) we have to choose β as $\beta\in(0,\frac{1}{4})$. The basic idea of the proof is a decomposition of the unit square by diadic rationals. Further we need Chebyshev inequality and the following two inequalities. We do not give the details.

1.
$$\mathsf{E}[|S_n(\omega,t_1,t_2)-S_n(\omega,u_1,u_2)|^2] =$$
$$= \mathsf{E}\left[\left|\sum_{k=0}^{n} c_k(t_1,t_2)\xi_k(\omega)-c_k(u_1,u_2)\xi_k(\omega)\right|^2\right] \leqq \sum_{k=0}^{\infty}[c_k(t_1,t_2)-c_k(u_1,u_2)]^2 =$$
$$= \sum_{k=0}^{\infty} c_k^2(t_1,t_2)+\sum_{k=0}^{\infty} c_k^2(u_1,u_2)-2\sum_{k=0}^{\infty} c_k(t_1,t_2)c_k(u_1,u_2) =$$
$$= t_1t_2+u_1u_2-2\min(t_1,u_1)\min(t_2,u_2) = V(I_{\bar t}\circ I_{\bar u})$$

where $I_{\bar t} = [0,t_1]\times[0,t_2]$, $\circ$ means the symmetric difference, and V the Lebesgue measure of the interval standing behind it.

2.
$$\mathsf{E}[|S_n(\omega,t_1,t_2)-S_n(\omega,u_1,u_2)|^4] = 3\mathsf{E}[|S_n(\omega,t_1,t_2)-S_n(\omega,u_1,u_2)|^2]^2 \leqq$$
$$\leqq 3V^2(I_{\bar t}\circ I_{\bar u})$$

So we proved that (2) converges uniformly with probability 1 and as it consists of continuous functions its limit function is also continuous.

REFERENCES

[1] Rényi, A.: *Foundation of probability* (In print).
[2] Lamperti, J.: *Probability: A Survey of the Mathematical Theory.* W. A. Benjamin, Inc., New York, Amsterdam 1966.
[3] Itô, K.—Nisio, M.: On the convergence of sums of indepedent Banach space valued random variables. *Osaka Journal of Mathematics* **5** (1968), 35—48.
4] Prohorov, Yu V.: Convergence of random processes and limit theorems in probability theory. *Theor. Probability Appl.* **1** (1956), 157—214.

Mathematical Institute of the Hungarian Academy of Sciences

(Received June 29, 1970.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 381—385.

ÜBER DIE POLYMERDEGRADATION

von

P. BÁRTFAI

Diese Arbeit schließt sich zur Arbeit [1] von T. KELEN, F. TÜDŐS, GY. GALAMBOS, und P. BÁLINT an, sie haben ein mathematisches Modell für die Polymerdegradation durch Elimination gegeben. Wir beschränken uns hier nur auf die mathematische Untersuchung dieses Modells und beschäftigen uns nicht mit seinem chemischen Hintergrund. Wir werden hier eine ausführlichere und vollständigere Verhandlung gewähren und einige Fehler verbessern.

Bezeichnen wir die Zustände einer Monomereinheit im Polymermolekul mit den Ziffern 0, 1 und 2, und es sei $\eta_n(t)=j$, wenn die n-te Monomereinheit im Zeitpunkt t im j-ten Zustand ist. Es ist notwendig zu bemerken, daß den Werten $\eta_n(t)=1$ und 2, chemisch derselbe Zustand entspricht, aber der eine aktiviert die $(n+1)$-ste Monomereinheit, der andere nicht, darum mußten wir diese unterscheiden.

Das mathematische Modell der Degradation ist das folgende (siehe [1]): $\eta_n(t)$ $(n=1, 2, \ldots;\ t\geqq 0)$ sei ein zweidimensionaler stochastischer Prozess mit den Eigenschaften

1. $\eta_n(t)$ kann nur die Werte 0, 1 und 2 annehmen;
2. $\eta_n(0)=0$ $(n=1, 2, \ldots)$;
3. für $0\leqq s<t$ und für die Werte $n=1, 2, \ldots$ (mit der Ergänzung $\eta_0(t)\equiv 0$) gilt;
4. $\eta_n(t)$ ist unter der Bedingung $\eta_n(t)>0$ unabhängig von der Zufallsveränderlichen $\eta_k(s)$ $(k<n;\ s\leqq t)$ und vom Ereignis $\eta_n(s)=0$ $(s<t)$, ferner es sei $\mathsf{P}(\eta_n(t)=1|\eta_n(t)>0)=\delta$, wobei δ eine Konstante ist;
5. aus $\eta_n(s)=j$ folgt $\eta_n(t)=j$ für $t\geqq s$ und $j=1, 2$.

Wir werden in dem Modell die folgenden Fragen untersuchen:

a) wie viele 0 Elemente stehen in der Folge $\eta_1(t), \eta_2(t), \ldots, \eta_N(t)$?

b) wie viele aus positiven Elementen bestehende Sequenzen der Länge k werden in der Folge $\eta_1(t), \eta_2(t), \ldots, \eta_N(t)$ sein?

c) wie viele aus 0 Elementen bestehende Sequenzen der Länge k werden in der Folge $\eta_1(t), \eta_2(t), \ldots, \eta_N(t)$ sein?

Es sei noch bemerkt, daß die Folge $\eta_1(t), \eta_2(t), \ldots$ keine Markoffsche Kette bildet, nämlich die Zufallsveränderliche $\eta_n(t)$ hängt davon stark ab, *wann* das vor ihr stehende 1 Element entstand, letzteres hängt aber auch von $\eta_{n-2}(t)$ ab. (Die Markoffsche Eigenschaft dieser Folge ist in der Arbeit [1] ausgenützt, so ist der Antwort auf die Frage b) irrtümlich.)

Obwohl die Folge $\eta_1(t), \eta_2(t), \ldots$ keine Markoffsche Kette bildet, enthält sie doch sogenannte Markoffsche Elemente, und zwar die Elemente $\eta_k(t)=0$. Nämlich es gilt

$$\mathsf{P}(\eta_n=j\,|\,\eta_{n-1}=0, \eta_{n-2}=i_{n-2}, \ldots, \eta_1=i_1) = \mathsf{P}(\eta_n=j\,|\,\eta_{n-1}=0).$$

Wenn wir eine Linie nach allen 0 Elementen in der Folge $\eta_1(t), \eta_2(t), \dots$ ziehen, so verteilen wir die Folge auf Abschnitte, und bezeichnen die Anzahl der in dem n-ten Abschnitt stehenden positiven Elemente mit ξ_n. So sind die Zufallsveränderlichen $\xi_1, \xi_2, \dots$ unabhängig und haben gleiche Verteilung. Es gelang uns dadurch die Probleme a)—c) auf die Bestimmung der Verteilung von ξ_1 zu reduzieren. Im ersten Teil dieser Arbeit bestimmen wir die erzeugende Funktion von ξ_1, dann — im zweiten Teil — geben wir Antworten auf die Probleme a)—c).

I.

Wir werden zuerst eine Rekursionsformel für die Funktion

$$f_k(t,\tau) = \mathsf{P}\big(\eta_1(t)>0, \eta_2(t)>0, \dots, \eta_k(t)>0, \eta_{k+1}(\tau)=0\big) \qquad (0 \leqq \tau \leqq t) \tag{1}$$

aufschreiben. Mit der Hilfe des Satzes über die vollständige Wahrscheinlichkeit ergibt sich einfach

$$f_{k+1}(t,\tau) = -(1-\delta)\int_0^\tau \left[\frac{d}{dx} f_k(t,x)\right] e^{-\alpha x} e^{-\beta(\tau-x)}\,dx +$$
$$+(1-\delta)e^{-\alpha\tau}\big(f_k(t,\tau)-f_k(t,t)\big)+\delta e^{-\alpha\tau} f_{k+1}(t,0).$$

Durch partielle Integration können wir die Rekursionsformel umformen:

$$f_{k+1}(t,\tau) = (1-\delta)(\beta-\alpha)e^{-\beta\tau}\int_0^\tau f_k(t,x)e^{(\beta-\alpha)x}\,dx + \tag{2}$$
$$+(1-\delta)\,[e^{-\beta\tau}f_k(t,0)-e^{-\alpha\tau}f_k(t,t)]+\delta e^{-\alpha\tau}f_{k+1}(t,0), \qquad (k=0,1,2,\dots)$$

und

$$f_0(t,\tau) = e^{-\alpha\tau}. \tag{3}$$

Führen wir die erzeugende Funktion $G(t,\tau;z) = \sum_{k=0}^{\infty} f_k(t,\tau)z^k$ ein; multiplizieren wir (2) mit z^{k+1} und summieren für die Werte $k=0,1,2,\dots$, so ergibt sich

$$G(t,\tau;z)-e^{-\alpha\tau} = (1-\delta)(\beta-\alpha)e^{-\beta\tau}z\int_0^\tau G(t,x;z)e^{(\beta-\alpha)x}\,dx + \tag{4}$$
$$+(1-\delta)ze^{-\beta\tau}G(t,0;z)-(1-\delta)ze^{-\alpha\tau}G(t,t;z)+\delta e^{-\alpha\tau}\big(G(t,0;z)-1\big).$$

Durch Differenzieren erhält man eine lineare Differenzialgleichung für $G(t,\tau;z)$

$$G'_\tau(t,\tau;z)+\big(\beta-(1-\delta)(\beta-\alpha)ze^{-\alpha\tau}\big)G(t,\tau;z) = \tag{5}$$
$$=(\beta-\alpha)e^{-\alpha\tau}\big[1-(1-\delta)zG(t,t;z)+\delta\big(G(t,0;z)-1\big)\big].$$

Die Lösung von (5) ist

$$G(t,\tau;z)=\exp\left\{-\beta\tau-(1-\delta)\frac{\beta-\alpha}{\alpha}ze^{-\alpha\tau}\right\}\times \tag{6}$$

$$\times\left[C+(\beta-\alpha)\left(1-(1-\delta)zG(t,t;z)+\delta\left(G(t,0;z)-1\right)\right)\right]\times$$

$$\times\int_0^\tau \exp\left\{(\beta-\alpha)\left[y+\frac{1-\delta}{\alpha}ze^{-\alpha y}\right]\right\}dy,$$

wobei C durch Einsetzung $\tau=0$ bestimmt werden kann, und zwar

$$C=G(t,0;z)e^{(1-\delta)\frac{\beta-\alpha}{\alpha}z}. \tag{7}$$

Anderseits aus (2) folgt

$$f_{k+1}(t,0)=f_k(t,0)-f_k(t,t),$$

oder zur erzeugenden Funktion übergegangen

$$G(t,0;z)=\frac{1-zG(t,t;z)}{1-z}. \tag{8}$$

Gesetzt man in (6) $\tau=t$ und schreibt man in (6) den Ausdruck (8) von $G(t,0;z)$, so kann $G(t,t;z)$ aus dieser Gleichung ausgedrückt werden. Halber der Kürze sei $G(t,t;z)=G(t,z)$, und führen wir die Bezeichnung

$$W(z)=\exp\left\{-\beta t-(1-\delta)\frac{\beta-\alpha}{\alpha}ze^{-\alpha t}\right\}\times \tag{9}$$

$$\times\left[e^{(1-\delta)\frac{\beta-\alpha}{\alpha}z}+(\beta-\alpha)(1-z+\delta z)\int_0^t \exp\left\{(\beta-\alpha)\left(y+(1-\delta)\frac{z}{\alpha}e^{-\alpha y}\right)\right\}dy\right]$$

ein, so gilt

$$G(t,z)=\frac{W(z)}{1-z\left(1-W(z)\right)}. \tag{10}$$

Dadurch gelang es uns die erzeugende Funktion der Wahrscheinlichkeiten $f_k(t,t)=$ $=\mathsf{P}(\eta_1(t)>0,\ldots,\eta_k(t)>0,\eta_{k+1}(t)=0)=\mathsf{P}(\xi_1=k)$ zu bestimmen.

SATZ 1. *Es gilt*

$$G(t,z)=\sum_{k=0}^{\infty}\mathsf{P}(\xi_1=k)z^k=\frac{W(z)}{1-z\left(1-W(z)\right)},$$

wobei $W(z)$ *durch* (9) *definiert ist.*

II.

Beschäftigen wir uns zuerst mit dem Verhältnis der 0 und 1 Elemente. ν_n bezeichnet die Anzahl der 0 Elemente in der Folge $\eta_1(t),\eta_2(t),\ldots,\eta_n(t)$.

SATZ 2. *Die folgenden drei Behauptungen gelten:*

(11) $$\lim_{n\to\infty} \frac{\nu_n}{n} = W(1)$$

mit Wahrscheinlichkeit 1;

(12) $$\lim_{n\to\infty} \frac{\mathsf{M}(\nu_n)}{n} = W(1);$$

die Verteilung von $\frac{\nu_n}{n}$ *ist asymptotisch normal mit dem Erwartungswert* $W(1)$ *und mit der Streuung* $\frac{\sigma}{\sqrt{n}}(W(1))^{3/2}$, *wobei*

$$\sigma^2 = \mathsf{D}^2(\xi_1) = \frac{1-W(1)-2W'(1)}{(W(1))^2}.$$

BEWEIS. (11) folgt aus dem Gesetz der großen Zahlen, nämlich

$$\lim_{n\to\infty} \frac{\nu_n}{n} = \lim_{n\to\infty} \frac{\nu_n}{\sum_{1}^{\nu_n}(\xi_i+1)} = \lim_{n\to\infty} \frac{n}{\sum_{1}^{n}(\xi_i+1)} = \frac{1}{1+\mathsf{M}(\xi_1)} = \frac{1}{1+G_z'(t,1)} = W(1).$$

(12) und die dritte Behauptung folgen aus die Sätze der Theorie des wiederkehrenden Prozesses (siehe z. B. [2], S. 354 und S. 359). Die Streuung von ξ_1 kann durch zweimaliges Differenzieren von $G(t, z)$ bestimmt werden.

SATZ 3. *Bezeichnet* $M_k(n)$ *die Anzahl der aus* 0 *Elementen bestehenden Sequenzen der Länge genau* k *in der Folge* $\eta_1(t), \ldots, \eta_n(t)$, *so gilt mit Wahrscheinlichkeit* 1

(13) $$M_k = \lim_{n\to\infty} M_k(n) = W(1)(1-e^{-\alpha t})^2 e^{-(k-1)\alpha t}.$$

BEWEIS. Wir setzen

$$\varepsilon_j = \begin{cases} 1, & \text{falls} \quad \xi_j<0,\ \xi_{j+1}=0, \ldots, \xi_{j+k-1}=0,\ \xi_{j+k}>0, \\ 0 & \text{sonst,} \end{cases}$$

so ist

(14) $$M_k = \lim_{n\to\infty} M_k(n) = \lim_{n\to\infty} \frac{\sum_{j=1}^{n-k} \varepsilon_j}{\sum_{j=1}^{n}(\xi_j+1)}.$$

Da $\varepsilon_1, \varepsilon_2, \ldots$ eine $(k+2)$-abhängige Folge bildet, das Gesetz der großen Zahlen anwendbar ist, und da $\mathsf{M}(\varepsilon_j) = \mathsf{P}(\varepsilon_j=1) = (1-e^{-\alpha t})^2 e^{-(k-1)\alpha t}$ ist, folgt die Behauptung des Satzes aus (14).

SATZ 4. *Bezeichnet* $C_k(n)$ *die Anzahl der aus positiven Elementen bestehenden Sequenzen der Länge genau* k *in der Folge* $\eta_1(t), \ldots, \eta_n(t)$, *so gilt mit Wahrscheinlichkeit* 1

(15) $$C_k = \lim_{n\to\infty} C_k(n) = W(1) f_k(t, t).$$

BEWEIS. Wenn $\lim_{n\to\infty} C_k(n) = C_k$ existiert, so gilt mit Wahrscheinlichkeit 1

$$C_k = \lim_{n\to\infty} \frac{r_k(n)}{\sum_1^n (\xi_i+1)} = \lim_{n\to\infty} \frac{\frac{1}{n} r_k(n)}{\frac{1}{n}\sum_1^n (\xi_i+1)} = W(1) f_k(t, t),$$

wo $r_k(n)$ die Häufigkeit des Ereignesses $\xi_i = k$ in der Folge $\xi_1, \xi_2, \dots, \xi_n$ bezeichnet. Die Existenz von C_k folgt aus den unteren und oberen Schätzungen mit den Ausdrücken $\frac{r_k(n)}{\sum_1^n (\xi_i+1)}$ bzw. $\frac{r_k(n)}{\sum_1^{n-1} (\xi_i+1)}$, und daraus, daß beide zum selben Grenzwert mit Wahrscheinlichkeit 1 streben.

Bemerkung. Die Funktionen $f_k(t, t)$ sind auch für kleine Werte von k ziemlich kompliziert, aber für konkrete Werte von α und β ist es möglich sie explicite auszurechnen mit der Hilfe der Rekursionsformel (2).

LITERATUR

[1] KELEN, T., TÜDŐS, F., GALAMBOS, GY., BÁLINT, P.: Polimerek eliminációs mechanizmusú degradációja, IV. *Magyar Kémiai Folyóirat,* **76** (1970), 16—23.

[2] FELLER, W.: *An Introduction to Probability Theory and its Applications,* Vol. II., J. Wiley, New York, 1966.

Mathematisches Institut der Ungarischen Akademie der Wissenschaften, Budapest

(Eingegangen: 6 Juli 1970.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 387—391.

НИЖНЯЯ ОЦЕНКА В ТЕОРИИ СПЛАЙН-ПРИБЛИЖЕНИЙ

В. ПОПОВ и Г. ФРАЙД (G. FREUD)

Обозначим через $\mathscr{S}(n, m)$ класс всех сплайн-функций порядка (n, m) на отрезке $[0, 1]$. Напомним, что $s(x) \in \mathscr{S}(n, m)$, если $s(x) \in \mathscr{C}^{(m-1)}[0, 1]$ и существует такое множество $n+1$ точек интервала $[0, 1]$: $0 = x_0 \leqq x_1 \leqq \ldots \leqq x_n = 1$, что на отрезке $[x_{i-1}, x_i]$ дункция $s(x)$ является многочленом m-той степени.

Рассмотрим наилучшее приближение $E^{(m)}_{n, L_p}(f)$ функции $f(x)$ сплайн-функциями из $\mathscr{S}(n, m)$ в норме L_p, $p > 0$:

$$E^{(m)}_{n, L_p}(f) = \inf_{s(x) \in \mathscr{S}(n, m)} \|f(x) - s(x)\|_{L_p[0, 1]}.$$

Известно, что если $f(x) \in \mathscr{C}^K$, то приближение в норме $\mathscr{C}$

$$E^{(k+1)}_{n, C}(f) = O\left[n^{-k}\omega_2\left(f^{(k)}, \frac{1}{n}\right)\right], \tag{1}$$

где $\omega_2(f^{(k)}, \delta)$ обозначает второй модуль непрерывности функции $f^{(k)}(x)$ (модуль гладкости Зигмунда).

С другой стороны, если $f^{(k)}(x)$ является абсолютно непрерывной на отрезке $[0, 1]$, то [1]

$$E^{(k+1)}_{n, L_p}(f) = O[n^{-k}\omega_2(f^{(k)}, \lambda_n)], \tag{2}$$

где

$$\lim_{n \to \infty} n\lambda_n = 0.$$

Здесь мы покажем, что условие абсолютной непрерывности существенно для (2). Точнее мы покажем:

A. *Для каждого выпуклого модуля непрерывности* $\omega(\delta)$, *удовлетворяыщему условию* $\lim\limits_{\delta \to 0} \dfrac{\delta}{\omega(\delta)} = 0$, *существует функция* $f(x)$ *такая, что*

$$\omega(f^{(k)}, \delta) \leqq c\omega(\delta)$$

и кроме того

$$\limsup_{n \to \infty} \frac{n^k E^{(m)}_{n, L_p}(f)}{\omega\left(\frac{1}{n}\right)} > c_1 > 0 \qquad (m = 1, 2, \ldots). \tag{3}$$

B. *Существует функция* $f(x) \in \mathscr{C}^{(k)}[0, 1]$ *такая, что* $\omega_2(f^{(k)}, \delta) \leqq c\delta$ *и*

$$\limsup_{n \to \infty} n^{k+1} E^{(m)}_{n, L_p}(f) > c_2 > 0 \qquad (m = 1, 2, \ldots).$$

C. *Для любой последовательчости* $\varepsilon_n \to 0, \varepsilon_n > 0$ *существует функция* $f(x) \in \mathscr{C}^{(k)}$ *такая, что* $f^{(k)}(x) \in Lip\ 1$ *и*

$$\limsup_{n\to\infty} \frac{n^{k+1}}{\varepsilon_n} E^{(m)}_{n, L_p}(f) > c_3 > 0 \qquad (m=1, 2, \ldots).$$

Пусть $0<p\leqq 1$ фиксированное число* и пусть c_r $(r=1, 2, \ldots)$ числа, зависящие только от p, N, k.

Рассмотрим функцию $\varphi(x) = e^{-\frac{1}{1-x^2}}$ в интервале $[-1, +1]$ и продолжим ее с периодом 2 на $(-\infty, +\infty)$. Отметим, что каждого m и $p>0$ имеем

$$E_{m, L_p}(\varphi) = \inf_{P\in H_m} \left\{\int_0^1 |\varphi(x) - P(x)|^p\, dx\right\}^{1/p} = \delta_{mp}^{1/p} > 0, \tag{4}$$

где H_m обозначает совокупность всех алгебраических многочленов степени не больше m.

Из (4) сразу следует, что для функции

$$\varphi(lx) \qquad (l=1, 2, 3, \ldots)$$

имеет место

$$\int_{\frac{i}{l}}^{\frac{i+1}{l}} |\varphi(lx) - P(x)|^p\, dx \geqq \frac{\delta_{mp}}{l} \qquad (i=0, \pm 1, \pm 2, \ldots) \tag{5}$$

для каждого $l>0$, $p>0$ и $P(x) \in H_m$.

Пусть $\varepsilon>0$ такое, что $9\varepsilon < \delta_{mp}$.

Рассмотрим сначала тот случал, когда $\lim\limits_{\delta\to 0} \dfrac{\delta}{\omega(\delta)} = 0$.

Будем предполагать, что $m \geqq k+1$; если $m<k+1$, то испорльзуем, что $E_n^m \geqq E_n^{k+1}$.

Найдем последовательность $\{A_i\}$ таких чётных натуральных чисел, для которых выполнены следующие условия:

$$A_{i+1}\omega\left(\frac{1}{A_{i+1}}\right) > 2A_i\omega\left(\frac{1}{A_i}\right) \tag{I}$$

(это можно сделать, так как $\lim\limits_{\delta\to 0} \dfrac{\delta}{\omega(\delta)} = 0$),

$$2\omega\left(\frac{1}{A_{i+1}}\right) < \omega\left(\frac{1}{A_i}\right), \tag{II}$$

$$A_i > 4A_{i-1}, \tag{III}$$

$$E^{(m)}_{\frac{A_l}{2}, L_p}\left(\varphi(A_i x)\right) \leqq \frac{\varepsilon\omega\left(\frac{1}{A_l}\right)}{lA_l^k} \qquad (i<l) \tag{IV}$$

---

* Исловия $p\leqq 1$ не влияет общноеть, потому что $E_{m;\, L_p}(\varphi) \geqq E_{m,\, L_1}(\varphi)$ при $p>1$.

(это можно сделать, так как φ бесконечно дифференцируемая функция (ср. (I)) и $m > k+1$, $l < A_l$), следовательно

$$\text{(V)} \qquad \sum_{i=l+1}^{\infty} \frac{\omega\left(\frac{1}{A_i}\right)}{A_i^k} < 2\frac{\omega\left(\frac{1}{A_{l+1}}\right)}{A_{l+1}^k} < \frac{\varepsilon}{A_l^k}\,\omega\left(\frac{1}{A_l}\right).$$

Рассмотрим функцию

$$F(x) = \sum_{i=1}^{\infty} \frac{\varphi(A_i x)\,\omega\left(\frac{1}{A_i}\right)}{A_i^k}.$$

Очевидно

$$F(x) \in C^k[0, 1] \quad \text{и} \quad F^{(k)}(x) = \sum_{i=1}^{\infty} \varphi^{(k)}(A_i x)\,\omega\left(\frac{1}{A_i}\right). \qquad \text{[см. (II)]}$$

Докажем, что

$$\omega(F^{(k)}(x), \delta) \leqq c\omega(\delta).$$

Рассмотрим

$$F_N^{(k)}(x) = \sum_{l=1}^{N} \varphi^{(k)}(A_i x)\,\omega\left(\frac{1}{A_i}\right).$$

Имеем:

$$\text{(6)} \qquad |F_N^{(k+1)}(x)| = \left|\sum_{i=1}^{N} A_i \varphi^{(k+1)}(A_i x)\,\omega\left(\frac{1}{A_i}\right)\right| \leqq c_4 A_N \omega\left(\frac{1}{A_N}\right). \qquad \text{[см. (I)]}$$

С другой стороны

$$\text{(7)} \qquad \left|\sum_{l=N+1}^{\infty} \varphi^{(k)}(A_i x)\,\omega\left(\frac{1}{A_i}\right)\right| \leqq c_5\,\omega\left(\frac{1}{A_{N+1}}\right). \qquad \text{[см. (II)]}$$

Из (6) и (7) следует:

$$|F^{(k)}(x+h) - F^{(k)}(x)| \leqq c_6\left[A_N\,\omega\left(\frac{1}{A_N}\right)h + \omega\left(\frac{1}{A_{N+1}}\right)\right].$$

Выбирая N так, что $A_N \leqq \frac{1}{h} < A_{N+1}$ и принимая во внимание, что из выпуклости ω следует $\frac{\omega(\delta_1)}{\delta_1} > \frac{\omega(\delta_2)}{\delta_2}$, если $\delta_1 < \delta_2$, получаем:

$$\omega(F^{(k)}, \delta) \leqq c\omega(\delta).$$

Оценим теперь $E^{(m)}_{\frac{A_l}{2}, L_p}(F(x))$: ввиду (IV) и (V)

$$\text{(8)} \qquad [E^{(m)}_{\frac{A_l}{2}, L_p}(F)]^p \geqq \left[E^{(m)}_{\frac{A_l}{2}, L_p}\left(\frac{\varphi(A_l x)\,\omega\left(\frac{1}{A_l}\right)}{A_l^k}\right)\right]^p - \left[\frac{2\varepsilon\omega\left(\frac{1}{A_l}\right)}{A_l^k}\right]^p.$$

Пусть $S_{\frac{A_l}{2}, m}(x)$ произвольная сплайн-функция порядка $\left(\frac{A_l}{2}, m\right)$. Из вида $\varphi(A_l x)$ сразу следует, что существует не менее чем $\frac{A_l}{3}$ интервалов $\left[\frac{i}{A_l}, \frac{i+1}{A_l}\right]$ таких, что внутри них нет узлов из $S_{\frac{A_l}{2}, m}(x)$.

Тогда из (5) следует, что

$$\int_0^1 |\varphi(A_l x) - S_{\frac{A_l}{2}}(x)|^p\,dx \geqq \frac{\delta_{mp}}{3} \tag{9}$$

для камдой сплайн-функции $S_{\frac{A_l}{2},m}(x)$ порядка $\left(\frac{A_l}{2}, m\right)$.

Из (8) и (9) мы полущаем неравенство

$$[E^{(m)}_{\frac{A_l}{2}, L_p}(F)]^p \geqq \frac{\delta_{mp} - 9\varepsilon}{3A_l^{kp}}\left[\omega\left(\frac{1}{A_l}\right)\right]^p,$$

т. е. (3).

В случае теоремы **B** пусть $\varepsilon > 0$ достаточно малая и выбераем последовательность $\{A_i\}$ натуральных чисел так, чтобы были выполнены условия ($\|\varphi\| = \max|\varphi(x)|$)

$$4A_i \leqq A_{i+1}, \tag{I'}$$

$$E^{(m)}_{\frac{A_l}{2}, L_p}\big(\varphi(A_i x)\big) \leqq \frac{\varepsilon\|\varphi\|}{A_l^{k+1}}, \qquad i<l,\ m \geqq k+1, \tag{II'}$$

$$\sum_{i=l+1}^{\infty} \frac{1}{A_i^{k+1}} \leqq \frac{\varepsilon}{A_l^{k+1}}. \tag{III'}$$

Рассмотрим функцию

$$F(x) = \sum_{i=1}^{\infty} \frac{\varphi(A_i x)}{A_i^{k+1}}.$$

Очевидно

$$F^{(k)}(x) = \sum_{i=1}^{\infty} \frac{\varphi^{(k)}(A_i x)}{A_i}.$$

Покажем, что $F^{(k)}(x) \in \mathscr{Z}$, где $\mathscr{Z}$ обозначает класс Зигмунда.

Действительно, при $\|\varphi^{k+2}\| = \max|\varphi^{(k+2)}(x)|$ имеем

$$|F^{(k)}(x+h) - 2F^{(k)}(x) + F^{(k)}(x-h)| \leqq$$

$$\leqq \left|\sum_{i=1}^{N} A_i \|\varphi^{(k+2)}\| h^2\right| + 4\sum_{i=N+1}^{\infty} \frac{\|\varphi^{(k)}\|}{A_i} \leqq c_7\left(A_N h^2 + \frac{1}{A_{N+1}}\right).$$

Выберая N так, что $A_N \leqq \frac{1}{h} < A_{N+1}$, мы получаем, что $F^{(k)}(x) \in \mathscr{Z}$.

То, что

$$E^{(m)}_{\frac{A_l}{2}, L_p}(f) \geqq \frac{c}{A_l}$$

доказывается также, как и выше.

Наконец мы доказаем утверждение **C**.

Выберем подпоследовательность ε_{n_i} такую, что $\sum_{i=1}^{\infty} \varepsilon_{A_i} < \infty$ и пусть A_i удовлетворяет условиям (I′) – (III′).

Рассмотрим функцию

$$F(x) = \sum_{i=1}^{\infty} \frac{\varphi(A_i x)}{A_i^{k+1}} \varepsilon_{A_i}.$$

Тогда очевидно

$$|F^{(k)}(x+h) - F^{(k)}(x)| \leqq ch \sum_{i=1}^{\infty} \varepsilon_{A_i} = c'h$$

и также, как и выше, получаем, что

$$E^{(m)}_{\frac{A_l}{2}, L_p}(F) \geqq c^* \frac{\varepsilon_{A_l}}{A_l^{k+1}}.$$

ЛИТЕРАТУРА

[1] Г. Фрайд—В. А. Попов: Некоторые вопросы, связанные с аппроксимацией сплайн-функциями и многочленами, *Studia Sci. Math. Hungar.* **5** (1970), 161—171.

Институт Математики Академии Наук Болгарии, София

Институт Математики Академии Наук Венгрии, Будапешт

(Поступила 10-ого июля 1970 г.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 393—397.

SOME REMARKS ON RANDOM NUMBER TRANSFORMATION

by

A. SZÉP

In Memoriam Professor A. Rényi

The problems of random number generation and transformation are of basic importance in the theory of Monte Carlo Methods.

In our paper we shall deal with the latter ones.

The general problem is the following: Let $\{\Omega, \mathcal{A}, \mathsf{P}\}$ be a probability field, $\xi=\xi(\omega)$ a real-valued random variable, defined on Ω, $G(x)$ a prescribed distribution function.

We have to find a transformation T such that the random variable $\eta=T\xi$ should possess the distribution function $G(x)$:

$$\mathsf{P}(\eta<x) = \mathsf{P}(T\xi<x) = G(x)$$

For solving this problem we have to tell something more about the special character of T. From the practical point view those T-s are of interest, which depend only on the "actual value" of ξ.

The random number generator of a computer gives one pseudo-random number after the other and we can start only from these numbers. At first one may try to find an appropriate function $\varphi(x)$, that is $T\xi=\varphi(\xi)$. It can be seen easily, that this is not always possible.

It is well known [1], that if $F(x)$ is the distribution function, of ξ and $G(x)$ is strictly monotone, then $\varphi(x)=G^{-1}(F(x))$ is an appropriate transforming function.

This result can hardly be applied because of the long computing time of the inverse function $G^{-1}(x)$. (However, there are some exceptions.)

A. Békéssy has raised the following question: Try to find such T-s, for which possibly a great part of the values of $\xi(\omega)$ remains unchanged, or, more precisely, the (probability) measure of the set $\{\omega\colon T\xi(\omega)\neq\xi(\omega)\}$, is small. In the sequel we shall consider random transformations, too. This means that we take two random numbers and, depending on the value of the second one, we change, or reject or accept the first one. Such procedures are very familiar in the applications of Monte Carlo methods [1].

In what follows let ξ be a random variable with distribution function $F(x)$. Let $G(x)$ be another distribution function. We have to find a real valued function $T(t, \omega)$, defined on $R\times\Omega$ (as a matter of fact, $T(t, \omega)$ is a stochastic process), such that

1. $\eta=T(\xi(\omega), \omega)$ is measurable
2. $\mathsf{P}(\eta<x)=G(x)$
3. $\mathsf{P}(\xi\neq\eta)$ is minimal, provided this minimum exists.

LEMMA. *Let A be any Borel subset of fhe real line. Then*

$$P(\xi \neq \eta) \geqq P(\xi \in A) - P(\eta \in A)$$

PROOF.
$$P(\xi \neq \eta) = P(\{\xi \neq \eta\} \cap \{\xi \in A\}) + P(\{\xi \neq \eta\} \cap \{\xi \notin A\}) \geqq$$
$$\geqq P(\{\xi \neq \eta\} \cap \{\eta \in A\}) \geqq P(\{\xi \in A\} \cap \{\eta \notin A\}) =$$
$$= P(\xi \in A) - P(\{\xi \in A\} \cap \{\eta \in A\}) \geqq P(\xi \in A) - P(\eta \in A) \quad \text{q.e.d.}$$

THEOREM 1. *Let ξ, F, G denote the same as before. Let us assume F and G to be absolutely continuous, $F' = f$, $G' = g$. If $T(t, \omega)$ is a transforming function with property* 1., 2., *then*

$$P(\xi(\omega) \neq T(\xi(\omega), \omega)) \geqq \frac{1}{2} \int\limits_{+\infty}^{-\infty} |f(x) - g(x)|\, dx. \tag{1}$$

PROOF: We apply the lemma to the set $A = \{x: f(x) > g(x)\}$.

$$P(\xi(\omega) \neq T(\xi(\omega), \omega) = P(\xi \neq \eta) \geqq P(\xi \in A) - P(\eta \in A) =$$

$$= \int\limits_A f(x)\, dx - \int\limits_A g(x)\, dx = \int\limits_A (f(x) - g(x))\, dx = \int\limits_{-\infty}^{+\infty} |f(x) - g(x)|_+\, dx =$$

$$= \int\limits_{-\infty}^{+\infty} \left[|f(x) - g(x)|_+ - \frac{1}{2}(f(x) - g(x)) \right] dx = \frac{1}{2} \int\limits_{-\infty}^{+\infty} |f(x) - g(x)|\, dx$$

Here we have considered that

$$\int\limits_{-\infty}^{+\infty} f(x)\, dx = \int\limits_{-\infty}^{+\infty} g(x)\, dx = 1,$$

hence,

$$\int\limits_{-\infty}^{+\infty} (f(x) - g(x))\, dx = 0$$

$|a|_+$ denotes the positive part of the real number a,

$$|a|_+ - \frac{a}{2} = \frac{|a|}{2} \quad \text{q.e.d.}$$

THEOREM 2. *By an appropriate choice of $T(t, \omega)$ (1) is valid with the sign of equality.*

PROOF: Let $\zeta(\omega)$ be a random variable, independent of ξ, uniformly distributed in [0, 1]. We define the following events:

$$B = \{\omega: f(\xi(\omega)) \leqq g(\xi(\omega))\}$$
$$C = \{\omega: f(\xi(\omega))\zeta(\omega) \leqq g(\xi(\omega)) < f(\xi(\omega))\}$$
$$D = \{\omega: f(\xi(\omega))\zeta(\omega) > g(\xi(\omega))\}$$

Clearly B, C and D is a complete system of events, i.e. they are mutually exclusive and their union is Ω. Let $A=\{u\colon u\in R,\ f(u)>g(u)\}$.

Consider the equation

$$(2)\qquad \int_{-\infty}^{t} |f(u)-g(u)|_+\,du = \int_{-\infty}^{\varphi} |g(u)-f(u)|_+\,du$$

Both integrals are continuous, increasing fnnctions of the upper limit of in tegration,

$$\int_{-\infty}^{+\infty} |f-g|_+\,du = \int_{-\infty}^{+\infty} |g-f|_+\,du \left(=\frac{1}{2}\int_{-\infty}^{+\infty} |f-g|\,du\right).$$

Thus we can express from (2) φ as a function of t (generally not uniquely). If we define

$$\varphi(t) = \inf\left\{\varphi\colon \int_{-\infty}^{t} |f-g|_+\,du = \int_{-\infty}^{\varphi} |g-f|_+\,du\right\},$$

then $\varphi(t)$ will be increasing.

Define $T(t,\omega)$ as follows:

$$T(t,\omega) = \begin{cases} t & \text{if } t\notin A \\ t & \text{if } t\in A \text{ and } f(t)\zeta(\omega)\leqq g(t) \\ \varphi(t) & \text{if } t\in A \text{ and } f(t)\zeta(\omega) > g(t) \end{cases}$$

Put

$$T(\xi(\omega),\omega) = \eta(\omega)$$

We prove that

$$\mathsf{P}(\eta\neq\xi) = \frac{1}{2}\int_{-\infty}^{+\infty} |f-g|\,du.$$

Applying the well-known formula for the conditional density function [2], we have

$$\mathsf{P}(\xi\neq\eta) = \int_{-\infty}^{+\infty} \mathsf{P}(\xi\neq\eta\,|\,\xi=u)f(u)\,du = \int_A \mathsf{P}(\xi\neq\eta\,|\,\xi=u)f(u)\,du \leqq$$

$$\leqq \int_A \left(1-\frac{g(u)}{f(u)}\right) f(u)\,du = \int_A (f(u)-g(u))\,du = \int_{-\infty}^{+\infty} |f-g|_+\,du = \frac{1}{2}\int_{-\infty}^{+\infty} |f-g|\,du$$

We prove that for any real x

$$\mathsf{P}(\eta<x) = \mathsf{P}\big(T(\xi(\omega),\omega)<x\big) = G(x).$$

Denote $c(x)$ the characteristic function of A (i.e. $c(x)=0$ or 1, according to $x\notin A$ or $x\in A$). Consider the event $A_x=\{\eta<x\}$. Clearly

$$\mathsf{P}(A_x)=\mathsf{P}(A_x\cap B)+\mathsf{P}(A_x\cap C)+\mathsf{P}(A_x\cap D) \tag{3}$$

(i)
$$\mathsf{P}(A_x\cap B)=\mathsf{P}(\{\eta<x\}\cap\{\xi\notin A\})=\mathsf{P}(\{\xi<x\}\cap\{\xi\notin A\})=$$
$$=\int\limits_{\substack{n\notin A\\ n<x}} f(u)\,du=\int\limits_{-\infty}^{x}(1-c(u))f(u)\,du$$

(ii)
$$\mathsf{P}(A_x\cap C)=\int\limits_{-\infty}^{+\infty}\mathsf{P}(A_x\cap C\,|\,\xi=u)f(u)\,du$$

We calculate the conditional probability under the integral sign:

$$\mathsf{P}(A_x\cap C\,|\,\xi=u)=\mathsf{P}\left(\{\eta<x\}\cap\left\{\zeta<\frac{g(u)}{f(u)}\right\}\middle|\,\xi=u\right)=$$
$$=\mathsf{P}\left(\{\eta<x\}\,|\,\{\xi=u\}\cap\left\{\zeta<\frac{g(u)}{f(u)}\right\}\right)\mathsf{P}\left(\zeta<\frac{g(u)}{f(u)}\right)$$

Here

$$\mathsf{P}\left(\eta<x\,|\,\{\xi=u\}\cap\left\{\zeta<\frac{g(u)}{f(u)}\right\}\right)=\begin{cases}1 & \text{if}\quad \xi\in A,\ \zeta<x\\ 0 & \text{elsewhere}\end{cases}$$

and

$$\mathsf{P}\left(\zeta<\frac{g(u)}{f(u)}\right)=\frac{g(u)}{f(u)}$$

Hence

$$\mathsf{P}(A_x\cap C\,|\,\xi=u)=\begin{cases}c(u)\dfrac{g(u)}{f(u)}, & \text{if}\quad u<x\\ 0 & \text{otherwise,}\end{cases}$$

thus

$$\mathsf{P}(A_x\cap C)=\int\limits_{-\infty}^{x}c(u)\frac{g(u)}{f(u)}\cdot f(u)\,du=\int\limits_{-\infty}^{x}c(u)g(u)\,du$$

(iii)
$$\mathsf{P}(A_x\cap D)=\mathsf{P}\left(\{\eta<x\}\cap\{\xi\in A\}\cap\left\{\zeta\geqq\frac{g(\xi)}{f(\xi)}\right\}\right)=$$
$$=\int\limits_{-\infty}^{+\infty}\mathsf{P}\left(\{\eta<x\}\cap\{\xi\in A\}\cap\left\{\zeta\geqq\frac{g(\xi)}{f(\xi)}\right\}\middle|\,\xi=u\right)f(u)\,du.$$

The conditional probability in the last integral vanishes for $u\notin A$. Suppose now $u\in A$, then

$$\mathsf{P}\left(\{\eta<x\}\cap\left\{\zeta\geqq\frac{g(\xi)}{f(\xi)}\right\}\middle|\,\xi=u\right)=\begin{cases}1-\dfrac{g(u)}{f(u)}, & \text{if}\quad \varphi(u)<x\\ 0, & \text{if}\quad \varphi(u)\geqq x.\end{cases}$$

Thus

$$\mathsf{P}(A_x \cap D) = \int\limits_{\varphi(u)<x} c(u)\left[1-\frac{g(u)}{f(u)}\right] f(u)\,du =$$

$$= \int\limits_{\varphi(u)<x} |f(u)-g(u)|_+\,du = \int\limits_{-\infty}^{x} |g(u)-f(u)|_+\,du = \int\limits_{-\infty}^{x} (1-c(u))\,(g(u)-f(u))\,du.$$

Here we have made use of the definition of φ.

Substituting the results into (3), we get

$$\mathsf{P}(\eta<x) = \int\limits_{-\infty}^{x} [1-c(u)]f(u)\,du + \int\limits_{-\infty}^{x} c(u)g(u)\,du + \int\limits_{-\infty}^{x} [1-c(u)]\,[g(u)-f(u)]\,du =$$

$$= \int\limits_{-\infty}^{x} g(u)\,du = G(x) \quad \text{q.e.d.}$$

Note that both Theorems 1 and 2 can be reformulated for the case of discrete random variables:

THEOREM 1′. *Let* $\{p_i\}_{i=-\infty}^{+\infty}, \{q_j\}_{j=-\infty}^{+\infty}$ *discrete probability distributions,* ξ *a discrete random variable,* $\mathsf{P}(\xi=k)=p_k$ $k=0, \pm 1, \dots,$

$$\eta = T(\xi(\omega), \omega), \quad \mathsf{P}(\eta=l)=q_l, \qquad l=0, \pm 1, \dots,$$

then

$$\mathsf{P}(\xi(\omega)\neq\eta(\omega)) \geqq \frac{1}{2}\sum_{j=-\infty}^{+\infty} |p_j-q_j| \tag{4}$$

THEOREM 2′. *By an appropriate choice of* $T(t, \omega)$, (4) *is valid with the sign of equality.*

REFERENCES

[1] HAMMERSLEY, J. M.: *Monte Carlo Methods.* Wiley-Methuen, 1964.
[2] RÉNYI, A.: *Probability theory.* North-Holland Publ. Comp. 1970.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received July 20, 1970.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 399—407.

ON THE ORDER OF CONVERGENCE OF FINITE-DIFFERENCE APPROXIMATIONS TO SOLUTIONS OF NON-SELFADJOINT ELLIPTIC BOUNDARY VALUE PROBLEMS

by

L. VEIDINGER

RIVKIND obtained in [1], [2] error estimates for finite-difference approximations to the solution of the Dirichlet problem for the general second-order elliptic equation in a bounded open region R. In these papers, however, it is assumed that the solution of the Dirichlet problem has bounded first derivatives and square integrable second derivatives in R. In general, these assumptions are satisfied only if the boundary of the region R is sufficiently smooth or if R is convex (see [3], [4]). In the present paper we shall obtain error estimates without these assumptions in the two-dimensional case. In deriving our estimates we shall use a discrete analogue of an inequality discovered by LADYŽENSKAJA and URAL'CEVA (see [5], p. 176).

1. Let R be a bounded open plane region whose boundary C consists of a finite number of piecewise-analytic simple closed curves. Denote by A_i $(i=1, 2, \ldots, n)$ the corners of C, i.e. those points on C where distinct analitic curves meet.

We consider the boundary value problem

$$Lu(x, y) = g(x, y), \quad (x, y) \in R, \tag{1}$$

$$u(x, y) = \varphi(x, y), \qquad (x, y) \in C,$$

where

$$Lu \equiv \frac{\partial}{\partial x}\left[a(x, y)\frac{\partial u}{\partial x}\right] + \frac{\partial}{\partial x}\left[b(x, y)\frac{\partial u}{\partial y}\right] + \frac{\partial}{\partial y}\left[b(x, y)\frac{\partial u}{\partial x}\right] +$$

$$+ \frac{\partial}{\partial y}\left[c(x, y)\frac{\partial u}{\partial y}\right] + d(x, y)\frac{\partial u}{\partial x} + e(x, y)\frac{\partial u}{\partial y} + f(x, y)u.$$

Let the coefficients of the operator L and the right-hand side $g(x, y)$ be analytic in an open region G containing the closure of R in its interior. Let $\varphi(x, y)$ be continuous on C and analytic on each analytic portion of C. Suppose that at all points of R

$$a\xi^2 + 2b\xi\eta + c\eta^2 \geqq \alpha(\xi^2 + \eta^2) \qquad (\alpha = \text{const} > 0) \tag{2}$$

for all real ξ, η.

Suppose the infinite plane of the region R is subdivided by two families of parallel lines into a square net. Let the lines of the net be $x = mh$ and $y = nh$ $(m, n = 0, \pm 1, \pm 2, \ldots)$. The points (mh, nh) will be called the nodes of the net. The smallest squares bounded by four lines of the net will be called meshes of the net. Denote by S_h the set of all nodes of the plane. Let R^* be the union of all meshes contained in R and let C^* be the boundary of R^*. Let R_h^* consist of all the interior nodes of R^* and let C_h^* be the net boundary of R_h^*. Let $\bar{R}_h^* = R_h^* \cup C_h^*$. The points of intersection of the net lines with the boundary C form the set C_h. We define

$$V_x(P) = h^{-1}[V(E) - V(P)], \quad V_{\bar{x}}(P) = h^{-1}[V(P) - V(W)],$$

$$V_y(P) = h^{-1}[V(N) - V(P)], \quad V_{\bar{y}}(P) = h^{-1}[V(P) - V(S)],$$

where $V = V(P)$ is any real-valued function defined on $\bar{R}_h^*$, $E = (x_P + h, y_P)$, $N = (x_P, y_P + h)$, $W = (x_P - h, y_P)$ $S = (x_P, y_P - h)$ are the four neighbours of the node $P = (x_P, y_P)$.

The solution u of the problem (1) is approximated by the solution U of the finite-difference problem

$$\text{(3)} \qquad \begin{aligned} L_h U(P) &= g(P), \qquad P \in R_h^*, \\ U(P) &= \varphi(P'), \qquad P \in C_h^*, \end{aligned}$$

where

$$L_h U = 0{,}5[(aU_x)_{\bar{x}} + (aU_{\bar{x}})_x + (bU_y)_{\bar{x}} + (bU_{\bar{y}})_x + (bU_x)_{\bar{y}} + (bU_{\bar{x}})_y + (cU_{\bar{y}})_y +$$
$$+ (cU_y)_{\bar{y}} + d(U_x + U_{\bar{x}}) + e(U_y + U_{\bar{y}})] + fU$$

and P' is the point of C_h closest to P.

If $V = V(P)$ and $W = W(P)$ are any two functions defined on $\bar{R}_h^*$, then we define

$$(V, W) = h^2 \sum_{P \in R_h^*} V(P)W(P), \quad \|V\| = (V, V)^{\frac{1}{2}}$$

and

$$H_h(V, W) = 0{,}5h^2 \sum_{P \in R_h^*} \{a(P)[V_x(P)W_x(P) + V_{\bar{x}}(P)W_{\bar{x}}(P)] +$$
$$+ b(P)[V_x(P)W_y(P) + V_y(P)W_x(P) + V_{\bar{x}}(P)W_{\bar{y}}(P) + V_{\bar{y}}(P)W_{\bar{x}}(P)] +$$
$$+ c(P)[V_y(P)W_y(P) + V_{\bar{y}}(P)W_{\bar{y}}(P)] - d(P)W(P)[V_x(P) + V_{\bar{x}}(P)] -$$
$$- e(P)W(P)[V_y(P) + V_{\bar{y}}(P)] - 2f(P)V(P)W(P)\}.$$

Here we have put

$$\text{(4)} \qquad \begin{aligned} &V_x(P) = W_x(P) = 0, \text{ if } P \in C_h^*,\ E \notin \bar{R}_h^*, \quad V_{\bar{x}}(P) = W_{\bar{x}}(P) = 0, \text{ if } P \in C_h^*,\ W \notin \bar{R}_h^*, \\ &V_y(P) = W_y(P) = 0, \text{ if } P \in C_h^*,\ N \notin \bar{R}_h^*, \quad V_{\bar{y}}(P) = W_{\bar{y}}(P) = 0, \text{ if } P \in C_h^*,\ S \notin \bar{R}_h^*. \end{aligned}$$

We define

$$\|\delta V\| = \{h^2 \sum_{P \in R_h^*} [(V_x(P))^2 + (V_y(P))^2]\}^{\frac{1}{2}},$$

where the conditions (4) are valid. We can extend the definition of V onto the whole set S_h putting $V(P) = 0$ for $P \notin \bar{R}_h^*$. Then we define

$$\|\delta V\|_1 = \{h^2 \sum_{P \in S_h} [(V_x(P))^2 + (V_y(P))^2]\}^{\frac{1}{2}}.$$

2. LEMMA 1. *u is an analytic function of x as well as of y in R.*

For a proof, see [6], p. 213.

LEMMA 2. *u is analytic on C, excluding the corners.*

For a proof of a more general result, see [7].

LEMMA 3. *Let $A_i=(x_{A_i}, y_{A_i})$ be a corner of C, with interior angle $\pi\alpha_i$ $(0<\alpha_i<2)$. Let*

$$x^* = k_{A_i}x+l_{A_i}y, \quad y^* = m_{A_i}x+n_{A_i}y \tag{5}$$

be a linear transformation which transforms the operator L into the normal form at the point A_i. Let $r_{A_i} = [(x-x_{A_i})^2+(y-y_{A_i})^2]^{\frac{1}{2}}$. If the transformation (5) transforms the angle $\pi\alpha_i$ into an angle $\pi\alpha_i^*$ $(0<\alpha_i^*<2)$, then for $\alpha_i^*\neq\frac{1}{m}$ (m an integer)

$$u(x, y) = u_1(x, y)+O\big(r_{A_i}^{\frac{1}{\alpha_i^*}}\big), \tag{6}$$

where $u_1(x, y)$ and its partial derivatives of all orders remain bounded when $(x, y)\to A$ in R while

$$u(x, y) = u_1(x, y)+O\big(r_{A_i}^{\frac{1}{\alpha_i^*}}|\log r_{A_i}|\big), \tag{7}$$

when $\alpha_i^*=\frac{1}{m}$ $(m=1, 2, \ldots)$. These relations may be indefinitely formally differentiated.

This lemma follows from the results of KONDRAT'EV (see [3]).

LEMMA 4. *Let $V=V(P)$ be any function defined at the nodes which vanishes outside $\bar{R}_h^*$. Then for h sufficiently small*

$$\max_{P\in\bar{R}_h^*} |V(P)| < c_1 |\log h|^{\frac{1}{2}} \|\delta V\|_1, \tag{8}$$

where c_1 is a positive constant depending only on the region R.

For a proof, see [8].

LEMMA 5. *Let $Z=Z(P)$ be any function defined on $\bar{R}_h^*$ which vanishes on C_h^*. Then for h sufficiently small*

$$\|Z\| \leqq c_2(\text{mes } R)^{\frac{1}{2}} \|\delta Z\|, \tag{9}$$

where $c_2 = \frac{1}{j\sqrt{\pi}}+\varepsilon$, i is the smallest positive root of the Bessel-function $J_0(x)$ and ε is any positive real number.

PROOF. If $v\in \mathring{W}_2^{(1)}(R)$*, then we define

$$\|v\|_{L_2(R)} = \Big[\iint_R v^2\,dx\,dy\Big]^{\frac{1}{2}}, \quad \|\nabla v\|_{L_2(R)} = \Big[\iint_R (v_x^2+v_y^2)\,dx\,dy\Big]^{\frac{1}{2}}.$$

* We denote by $\mathring{C}^{(1)}(R)$ the class of differentiable functions with compact support in R and by $\mathring{W}_2^{(1)}(R)$ the completion of $\mathring{C}^{(1)}(R)$ with respect to the norm of the space $W_2^{(1)}(R)$.

Divide each mesh of the net into two triangles by means of a diagonal in a fixed direction. Denote by $\bar{z}(x, y)$ the function which is linear in each triangle, coincides with $Z(P)$ at the nodes of R_h^* and vanishes outside R^*. It is easy to show that

$$\|\bar{z}\|^2_{L_2(R)} \geqq \|Z\|^2 + O(h^2\|\delta Z\|^2) \tag{10}$$

and

$$\|\nabla\bar{z}\|^2_{L_2(R)} = \|\delta Z\|^2 \tag{11}$$

On the other hand, using the so-called RAYLEIGH—FABER principle (see [9], p. 231) we have

$$\|\bar{z}\|_{L_2(R)} \leqq \frac{1}{j\sqrt{\pi}}(\text{mes } R)^{\frac{1}{2}}\|\nabla\bar{z}\|_{L_2(R)}. \tag{12}$$

Substituting (10) and (11) into (12) we obtain (9).

3. THEOREM 1. *Let* $f_0 = (\text{mes } R)^{-1}\iint\limits_R f(x, y)dxdy, f^+(x, y) = \max\{f(x, y) - f_0, 0\}$. *Assume that*

$$\left[\frac{2M(2\alpha+1)}{\alpha^2} + \frac{4}{\alpha}f_0\right]c_2^2 \text{ mes } R < 1, \tag{13}$$

where

$$M = \max\{\max_{(x,y)\in\bar{R}}[d^2(x, y) + e^2(x, y)], \max_{(x,y)\in\bar{R}} f^+(x, y)\}.$$

Let $z(P) = u(P) - U(P)$. *Then for all sufficiently small* h

$$\max_{P\in\bar{R}_h^*}|z(P)| < c_3 h^{\frac{1}{2}}|\log h|^{\frac{1}{2}}, \tag{14}$$

where c_3 *is a positive constant independent of* h.

PROOF. The truncation error $z(P)$ is the solution of the problem

$$\begin{aligned} L_h z(P) &= \Phi(P), \qquad P\in R_h^*, \\ z(P) &= u(P) - u(P'), \qquad P\in C_h^*, \end{aligned} \tag{15}$$

where $\Phi = L_h u - Lu$.

It is easy to see that $z = v + w$, where v is the solution of the problem

$$\begin{aligned} L_h v(P) &= 0, \qquad P\in R_h^*, \\ v(P) &= u(P) - u(P'), \qquad P\in C_h^* \end{aligned} \tag{16}$$

and w is the solution of the problem

$$\begin{aligned} L_h w(P) &= \Phi(P), \qquad P\in R_h^*, \\ w(P) &= 0, \qquad P\in C_h^*. \end{aligned}$$

Let $v^1(P) = v(P) - \psi(P)$, where

$$\psi(P) = \begin{cases} u(P) - u(P'), & P\in C_h^*, \\ 0, & P\in R_h^*. \end{cases}$$

Then

$$L_h v^1(P) = -L_h\psi(P), \qquad P\in R_h^*,$$
$$v^1(P) = 0, \qquad P\in C_h^*.$$

We define

$$H_h(v^1)\equiv H_h(v^1, v^1).$$

Then, using the finite-difference analogue of Green's first identity (see [10]) we obtain

(17) $$H_h(v^1) = -(L_h v^1, v^1) = (L_h\psi, v^1) = -H_h(\psi, v^1).$$

It is easy to see that $f(x, y) = f^+(x, y) - f^-(x, y)$, where

$$f^+(x, y) = \max\{f(x, y) - f_0, 0\}, \quad f^-(x, y) = -f_0 + \max\{-f(x, y) + f_0, 0\}.$$

Using (2) we have

(18) $$\alpha\|\delta v^1\|^2 + h^2\sum_{P\in R_h^*} f^-(P)[v^1(P)]^2 \leqq$$

$$\leqq H_h(v^1) + 0{,}5h^2\sum_{P\in R_h^*}\{d(P)v^1(P)[v_x^1(P) + v_{\bar{x}}^1(P)] + e(P)v^1(P)[v_y^1(P) + v_{\bar{y}}^1(P)] +$$
$$+ 2f^+(P)[v^1(P)]^2\} \leqq H_h(v^1) + \varepsilon\|\delta v^1\|^2 +$$
$$+ h^2\sum_{P\in R_h^*}\left\{\left[\frac{1}{4\varepsilon}(d^2(P) + e^2(P)) + f^+(P)\right][v^1(P)]^2\right\},$$

where ε is any positive real number. But

$$h^2\sum_{P\in R_h^*}\left\{\left[\frac{1}{4\varepsilon}(d^2(P) + e(P)) + f^+(P)\right][v^1(P)]^2\right\} \leqq M\left(\frac{1}{4\varepsilon} + 1\right)\|v^1\|^2,$$

where

$$M = \max\left\{\max_{(x,y)\in\bar{R}}[d^2(x, y) + e^2(x, y)], \max_{(x,y)\in\bar{R}} f^+(x, y)\right\}$$

and, therefore, setting $\varepsilon = \frac{\alpha}{2}$ in (18) we find that

(19) $$\|\delta v^1\|^2 + \frac{2}{\alpha}h^2\sum_{P\in R_h^*} f^-(P)[v^1(P)]^2 \leqq \frac{2}{\alpha}H_h(v^1) + \frac{M(2\alpha+1)}{\alpha^2}\|v^1\|^2.$$

Substituting (17) into (19) we obtain

(20) $$\|\delta v^1\|^2 + \frac{2}{\alpha}h^2\sum_{P\in R_h^*} f^-(P)[v^1(P)]^2 \leqq -\frac{2}{\alpha}H_h(\psi, v^1) + \frac{M(2\alpha+1)}{\alpha^2}\|v^1\|^2.$$

Using Schwarz's inequality and (9) we have

(21) $$|H_h(\psi, v^1)| \leqq 0{,}5h^2\sum_{P\in R_h^*}\{|a(P)|[|\psi_x(P)|\,|v_x^1(P)| + |\psi_{\bar{x}}(P)|\,|v_{\bar{x}}^1(P)|] +$$
$$+ |b(P)|[|\psi_x(P)|\,|v_y^1(P)| + |\psi_y(P)|\,|v_x^1(P)| + |\psi_{\bar{x}}(P)|\,|v_{\bar{y}}^1(P)| + |\psi_{\bar{y}}(P)|\,|v_{\bar{x}}^1(P)|] +$$
$$+ |c(P)|[|\psi_y(P)|\,|v_y^1(P)| + |\psi_{\bar{y}}(P)|\,|v_{\bar{y}}^1(P)|] + |d(P)|\,|v^1(P)|[|\psi_x(P)| + |\psi_{\bar{x}}(P)|] +$$
$$+ |e(P)|\,|v^1(P)|[|\psi_y(P)| + |\psi_{\bar{y}}(P)|] + 2|f(P)|\,|\psi(P)|\,|v^1(P)|\} \leqq c_4\|\delta v^1\|(\|\delta\psi\| + \|\psi\|),$$

where c_4 is a positive constant depending only on the region R and the coefficients of the operator L. Inserting (21) into (20) we have

$$\|\delta v^1\|^2+\frac{2}{\alpha}h^2\sum_{P\in R_h^*}f^-(P)[v^1(P)]^2 \leqq \tag{22}$$

$$\leqq \frac{1}{2}\|\delta v^1\|^2+\frac{2}{\alpha^2}c_4^2(\|\delta\psi\|+\|\psi\|)^2+\frac{M(2\alpha+1)}{\alpha^2}\|v^1\|^2.$$

Hence it follows that

$$\|\delta v^1\|^2+\frac{4}{\alpha}h^2\sum_{P\in R_h^*}f^-(P)[v^1(P)]^2 \leqq \frac{4}{\alpha}c_4^2(\|\delta\psi\|+\|\psi\|)^2+\frac{2M(2\alpha+1)}{\alpha^2}\|v^1\|^2. \tag{23}$$

We note that $\min\limits_{(x,y)\in \bar{R}} f^-(x,y)=-f_0$. Thus, applying (9) to (23) we obtain

$$\|\delta v^1\|^2 \leqq \frac{4}{\alpha}c_4^2(\|\delta\psi\|+\|\psi\|)^2+\gamma\|\delta v^1\|^2, \tag{24}$$

where $\gamma=\left[\frac{2M(2\alpha+1)}{\alpha^2}+\frac{4}{\alpha}f_0\right]c_2^2 \operatorname{mes} R$. By (13) we have $\gamma<1$ and, therefore, it follows from (24) that

$$\|\delta v^1\| \leqq \frac{2c_4}{[\alpha(1-\gamma)]^{\frac{1}{2}}}(\|\delta\psi\|+\|\psi\|). \tag{25}$$

Our next aim is to estimate $\|\delta\psi\|$ and $\|\psi\|$.

Let A_i be a corner of C, with interior angle $\pi\alpha_i$ and let $P\in C_h^*$. Denote by $r(P, A_i)$ the distance between the points P and A_i. If r_1 is a sufficiently small positive real number and $3h<r(P,A_i)<r_1$, then using (6) and (7), respectively, we obtain

$$u(P)-u(P')=(x_P-x_{P'})u_x(P'')+(y_P-y_{P'})u_y(P'')= \tag{26}$$

$$=O\left([r(P,A_i)]^{\frac{1}{\alpha_i^*}-1-\varepsilon}h+h\right),$$

where P'' is a point in the interval PP' and ε is any positive real number. It follows from (26) that

$$\sum_{P\in C_h^*,\ 3h<r(P,A_i)<r_1}[u(P)-u(P')]^2= \tag{27}$$

$$=O\Big(\sum_{P\in C_h^*,\ 3h<r(P,A_i)<r_1}[r(P,A_i)]^{\frac{2}{\alpha_i^*}-2-2\varepsilon}h^2+h\Big)=$$

$$=O\left(h^{\frac{2}{\alpha_i^*}-2\varepsilon}\sum_{1\leqq n\leqq\frac{r_1}{h}}\frac{1}{n^{2-\frac{2}{\alpha_i^*}+2\varepsilon}}+h\right)=O(h).$$

On the other hand, using (6) and (7) we have

(28) $$\sum_{P\in C_h^*,\, r(P,A_i)\le 3h} [u(P)-u(P')]^2 = O\big(h^{\frac{2}{\alpha_i^*}-2\varepsilon}+h^2\big) = O(h).$$

Summing (27) and (28) over all corners of A_i and applying Lemmas 1 and 2 we obtain

(29) $$\sum_{P\in C_h^*} [u(P)-u(P')]^2 = O(h).$$

It follows from (29) that

(30) $$\|\delta\psi\| = O(h^{\frac{1}{2}}), \quad \|\psi\| = O(h^{\frac{3}{2}}).$$

Let

$$V(P) = \begin{cases} v(P), & P\in \bar{R}_h^*, \\ 0, & P\notin \bar{R}_h^*, \end{cases} \quad \Psi(P) = \begin{cases} \psi(P), & P\in \bar{R}_h^*, \\ 0, & P\notin \bar{R}_h^*. \end{cases}$$

From (30) we see that

(31) $$\|\delta\Psi\|_1 = O(h^{\frac{1}{2}}).$$

Let

$$V^1(P) = \begin{cases} v^1(P), & P\in \bar{R}_h^*, \\ 0, & P\notin \bar{R}_h^*. \end{cases}$$

It is easy to see that $V^1(P) = V(P)-\Psi(P)$. From (25) and (30) we obtain that

(32) $$\|\delta V^1\|_1 = \|\delta v^1\| = O(\|\delta\psi\|+\|\psi\|) = O(h^{\frac{1}{2}}).$$

It follows from (31) and (32) that

(33) $$\|\delta V\|_1 \le \|\delta V^1\|_1 + \|\delta\Psi\|_1 = O(h^{\frac{1}{2}})$$

whence, using Lemma 4, we have

(34) $$\max_{P\in R_h^*} |v(P)| = O(h^{\frac{1}{2}}|\log h|^{\frac{1}{2}}).$$

The function $w(P)$ can be estimated in the same way as in the self-adjoint case (see [11]). Thus we obtain

(35) $$\max_{P\in R_h^*} |w(P)| = O(h^{\beta^*}|\log h|^{\frac{1}{2}},$$

where

$$\beta^* = \begin{cases} \dfrac{1}{\max\limits_{i=1,\dots,n} \alpha_i^*} - \varepsilon, & \text{if } \max\limits_{i=1,\dots,n} \alpha_i^* \ge \dfrac{1}{2}, \\ 2, & \text{if } \max\limits_{i=1,\dots,n} \alpha_i^* < \dfrac{1}{2} \text{ or if there are no corners.} \end{cases}$$

It follows from (34) and (35) that

(36) $$\max_{P\in R_h^*} |z(P)| = O(h^{\frac{1}{2}}|\log h|^{\frac{1}{2}}).$$

This completes the proof of Theorem 1.

It is easy to show that the condition (13) is always satisfied if mes R is sufficiently small. On the other hand, the condition (13) is satisfied for each operator L which can be written in the form $L = L_0 - \lambda E$, where L_0 is an elliptic operator with bounded coefficients and λ is a sufficiently large positive real number.

If $b(x, y) \equiv 0$, $f(x, y) \leqq 0$ and h is sufficiently small, then the matrix corresponding to the problem (3) is diagonally dominant and of non-negative type and, consequently, the finite-difference analogue of the maximum principle is valid (see, for example, [12]). Let

$$\gamma^* = \begin{cases} \dfrac{1}{\max\limits_{i=1,\dots,n} \alpha_i^*} - \varepsilon, & \text{if} \quad \max\limits_{i=1,\dots,n} \alpha_i^* \geqq 1, \\ 1, & \text{if} \quad \max\limits_{i=1,\dots,n} \alpha_i^* < 1 \text{ or if there are no corners,} \end{cases}$$

where ε is any positive real number. (Note that $\alpha_i^* < 1$ if and only if $\alpha_i < 1$). Since by Lemma 3 $u(P) - u(P') = O(h^{\gamma^*})$ for $P \in C_h^*$, we have the sharper estimate

$$\max_{P \in \bar{R}_h^*} |v(P)| = O(h^{\gamma^*}) \tag{37}$$

instead of (34). Combining (35) and (37) we obtain

$$\max_{P \in \bar{R}_h^*} |z(P)| = O(h^{\gamma^*}). \tag{38}$$

Thus we have proved the following theorem.

THEOREM 2. *Assume that* $b(x, y) \equiv 0$, $f(x, y) \leqq 0$ *and assume that the condition* (13) *is satisfied. Then the truncation error* $z(P)$ *satisfies* (38).

REFERENCES

[1] RIVKIND, V. JA.: An approximate method of solving the Dirichlet problem and estimates of the speed of convergence of solutions of the difference equations to solutions of elliptic equations with discontinuous coefficients, *Vestnik Leningrad. Univ. Ser. Mat. Meh. Astronom.* **19** (1964) 37—52.
[2] RIVKIND, V. JA.: On an estimate of the rapidity of convergence of homogeneous difference schemes for elliptic and parabolic equations with discontinuous coefficients, *Problems Math. Anal. Boundary Value Problems Integr. Equations, Izdat. Leningrad. Univ., Leningrad,* 1966. 110—119.
[3] KONDRAT'EV, V. A.: Boundary value problems for elliptic equations in domains with conical or angular points, *Trudy Moskov. Mat. Obšč.* **16** (1967) 209—292.
[4] HANNA, M. S. and SMITH, K. S.: Some remarks on the Dirichlet problems with piecewise smooth domains, *Comm. Pure Appl. Math.* **20** (1967) 575—593.
[5] LADYŽENSKAJA, O. A. and URAL'CEVA, N. N.: *Linear and quasi-linear equations of elliptic type,* Izdat. „Nauka", Moscow, 1964.
[6] MIRANDA, C.: *Partial differential equations of elliptic type,* Sringer, Berlin—Heidelberg—New York, 1970.
[7] MORREY, C. B. and NIRENBERG, L.: On the analyticity of linear elliptic systems of partial differential equations, *Comm. Pure Appl. Math.* **10** (1957), 271—290.

[8] BRAMBLE, J. H.: A second order finite difference analog of the first biharmonic boundary value problem, *Num. Math.* **9** (1966), 236—249.
[9] PÓLYA, G. and SZEGŐ, G.: *Isoperimetric inequalities in mathematical physics,* Princeton University Press, Princeton, 1951.
[10] COURANT, R., FRIEDRICHS, K. and LEWY, H.: Über die partiellen Differenzengleichungen der mathematischen Physik. *Math. Ann.* **100** (1928), 32—74.
[11] VEIDINGER, L.: On the order of convergence of finite-difference approximations to the solution of the Dirichlet problem in a domain with corners, *Studia Sci. Math. Hungar.* **3** (1968), 337—343.
[12] BRAMBLE, J. H. and HUBBARD, B.: A theorem on error estimation for finite difference analogues of the Dirichlet problem for elliptic equations. *Contributions to Differential Equations,* 2. Wiley, New York, 1963.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received August 23, 1970)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 409—417.

ON SOME PROBLEMS CONCERNING DIRECTIONAL DIMENSION

by

J. GERLITS

In this paper we shall deal with some problems raised by E. DEÁK in [1]. For the sake of completeness we list the necessary definitions and theorems. For the proofs and for further theorems see the forthcoming book [2] or [3].

The author wishes to take this opportunity to express his indebtedness to Professor Á. CSÁSZÁR and E. DEÁK for their help in the preparation of this paper.

1. §

All results of this paragraph are due to E. DEÁK.

(1. 1) *Definition.* A *direction* on a topological space X is a family $\mathscr{R}$ of ordered pairs (G, F) of subsets of X, where G is open and F is closed, satisfying the following three conditions:[1]

(i) $(\emptyset, \emptyset)$, $(X, X) \in \mathscr{R}$

(ii) $G \subset F$ for each (G, F) of $\mathscr{R}$, and for any two (G, F) and (G', F'), either $F \subset G'$ or $F' \subset G$

(iii) let $\mathscr{G}(\mathscr{R}) = \{G; \exists F, (G, F) \in \mathscr{R}\}$ and $\mathscr{F}(\mathscr{R}) = \{F; \exists G, (G, F) \in \mathscr{R}$; then if $\mathscr{B}$ is a subfamily of $\mathscr{G}(\mathscr{R}) \cup \mathscr{F}(\mathscr{R})$, $\bigcup \{B; B \in \mathscr{B}\}$ and $\bigcap \{B; B \in \mathscr{B}\}$ are members of $\mathscr{G}(\mathscr{R}) \cup \cup \mathscr{F}(\mathscr{R})$.

(1. 2) *Definition.* If $\mathscr{R}$ is a direction on a space X, the elements of $\mathscr{G}(\mathscr{R})$ are called *lower open $\mathscr{R}$-halfspaces*, and the complements of the elements of $\mathscr{F}(\mathscr{R})$ *upper open $\mathscr{R}$-halfspaces*.

(1. 3) *Definition. A directional structure* (abbr. DS) on a topological space X is a system $\mathfrak{R} = \{\mathscr{R}_\alpha; \alpha \in A\}$ of directions on X, for which the family of all open half-spaces form a subbasis of X.

(1. 4) *Definition. The directional dimension* — denoted by Dim X — of an indiscrete space X is equal to 0. If X is a not indiscrete topological space then Dim X is the minimum of the cardinalities of its directional structures.

(1. 5) THEOREM. *Every topological space X has a directional structure and* Dim $X \leqq w(X)$, *the weight of X.*

[1] This is an equivalent modification of the original definition of E. DEÁK; see [2].

(1. 6) THEOREM. *If* $\{X_\alpha; \alpha \in A\}$ *is a family of topological spaces, then*

$$\operatorname{Dim}(\times\{X_\alpha; \alpha \in A\}) \leqq \Sigma\{\operatorname{Dim} X_\alpha; \alpha \in A\}.$$

(1. 7) THEOREM. *If* X' *is a subspace of a space* X, *then* $\operatorname{Dim} X' \leqq \operatorname{Dim} X$.

(1. 8) THEOREM. *For a discrete space* X

$$\operatorname{Dim} X \leqq 1.$$

(1. 9) THEOREM. *If* $\mathbf{R}^n$ *denotes the n-dimensional euclidean space, then* $\operatorname{Dim} \mathbf{R}^n = n$ $(n = 0, 1, 2, \ldots)$.

(1. 10) THEOREM. *For a separable metric space* X, $\operatorname{Dim} X \leqq n$ *if and only if* X *is topologically embeddable into* $\mathbf{R}^n$.

(1. 11) THEOREM. *For any separable metric space* X, *if* $\dim X < \infty$, *then*

$$\dim X \leqq \operatorname{Dim} X \leqq 2 \dim X + 1$$

and the limits are the best possible.

(1. 12) THEOREM. *If* X *is a topological space and* $\operatorname{Dim} X < \aleph_0$, *then* $\operatorname{ind} X \leqq$ $\leqq \operatorname{Dim} X$.

(1. 13) THEOREM. *Let* X *be a perfectly normal space and* $\mathfrak{n}$ *a cardinal number, then* $\operatorname{Dim} X \leqq \mathfrak{n}$ *if and only if* X *is topologically embeddable into the product of* $\mathfrak{n}$ *order topological spaces.*

(1. 14) THEOREM. *For any order topological space* X, $\operatorname{Dim} X \leqq 1$.

§ 2.

A product theorem for Dim

(2. 1) LEMMA. *Let* X *be a topological space,* $x \in X$. *Then there exist systems of open sets* $\{\mathfrak{U}_t; t \in T\}$ *such that*

a) $\overline{\overline{T}} \leqq 2 \operatorname{Dim} X$

b) *The systems* $\mathfrak{U}_t$ *are ordered by inclusion* $(t \in T)$

c) $\mathfrak{U} = \bigcup \{\mathfrak{U}_t; t \in T\}$ *is a neighbourhood-subbasis of the point* x.

PROOF. Let $\mathfrak{R} = \{\mathscr{R}_s; s \in S\}$ be a DS on X, $\overline{\overline{S}} = \operatorname{Dim} X$. Considering the systems of the lower (resp. upper) open halfspaces containing x, we obtain the required systems.

(2. 2) THEOREM. *Let* $\{X_s; s \in S\}$ *be an uncountable family of* T_0*-spaces,* $\overline{\overline{X_s}} \geqq 2$ $(s \in S)$; *then the following equality holds:*

$$\operatorname{Dim}(\times\{X_s; s \in S\}) = \sum \{\operatorname{Dim} X_s; s \in S\}.$$

PROOF. By virtue of (1. 6) the sign $\leqq$ is valid; if X denotes the space $\times\{X_s; s\in S\}$, then by (1. 7) we obtain

$$\operatorname{Dim} X \geqq \sup\{\operatorname{Dim} X_s; s\in S\}$$

and, evidently

$$\sum\{\operatorname{Dim} X_s;\ s\in S\} = \max\{\overline{\overline{S}};\ \sup\{\operatorname{Dim} X_s;\ s\in S\}\},$$

consequently it is enough to prove

$$\operatorname{Dim} X \geqq \overline{\overline{S}}. \tag{2. 3}$$

Let π_s be the projection into X_s; for every $s\in S$ we choose two points x_s, y_s of X_s such that $x_s\notin\overline{\{y_s\}}$. We denote by $\bar{x}$ the point of X, for which $\pi_s(\bar{x})=x_s$ $(s\in S)$ and by $\bar{x}^p$ $(p\in S)$ that point, for which

$$\pi_s(\bar{x}^p) = \begin{cases} y_p; & \text{if } \ s=p \\ x_s; & \text{if } \ s\in S-\{p\} \end{cases}$$

Now, let us suppose that (2. 3) is false; then by (2. 1) there exist systems $\{\mathfrak{S}_q; q\in Q\}$, $\overline{\overline{Q}}\leqq 2\operatorname{Dim} X<\overline{\overline{S}}$ (S is an infinite set!), $\mathfrak{S}_q$ are ordered by inclusion and $\mathfrak{S}=\bigcup\{\mathfrak{S}_q; q\in Q\}$ form a neighbourhood-subbasis of the point $\bar{x}$. By the definition of the product topology, for every set $U\in\mathfrak{S}$, the set

$$M_u = \{p; p\in S, \bar{x}^p\notin U\}$$

is finite. If $U, V\in\mathfrak{S}_q$ and $U\subset V$, then $M_U\supset M_V$, thus, for fixed $q\in Q$, the system $\{M_U;\ U\in\mathfrak{S}_q\}$ is ordered by inclusion, consisting of finite sets, consequently the union of that system is countable. Making use of $\overline{\overline{Q}}<\overline{\overline{S}}$ and $\overline{\overline{S}}>\aleph_0$, we obtain that there exists an element $p\in S$ such that $\bar{x}^p\in U$ for all $U\in\mathfrak{S}$.

This means that $\bar{x}\in\overline{\{\bar{x}^p\}}$, which is impossible, because $x_p\notin\overline{\{y_p\}}$. Q.e.d.

(2. 4) *Corollary.* If I denotes the interval [0, 1] and $\overline{\overline{A}}>\aleph_0$, then $\operatorname{Dim}(I^A)=\overline{\overline{A}}$.

3. §

In this paragragraph, we shall deal with some generalizations of Theorem (1. 11).

We begin with a negative result.

(3. 1) *Example.* For each natural number n, there exist a compact T_2-space C_n with

$$\dim C_n=0, \quad \operatorname{Dim} C_n=n.$$

We denote with ω_i the i-th infinite initial number, by $W(\omega_i)$ the order topological space of the ordinal numbers less than ω_i, by

$$W_i = W(\omega_i)\times\{i\} \qquad (1\leqq i\leqq 2n)$$

Let

$$\Omega\notin\bigcup\{W_i;\ i=1,2,\dots,2n\},$$

$$C_n = \bigcup\{W_i;\ i=1,2,\dots,2n\}\cup\{\Omega\}.$$

A neighbourhood basis of a point $(\alpha, i) \in W_i$, $\alpha \neq 0$ is the family of sets $(\beta, \gamma) \times \{i\}$, where

$$\beta, \gamma \in W(\omega_i), \qquad \beta < \alpha < \gamma$$

(i.e. the customary intervals in $W(\omega_i) \times \{i\}$), the points $(0, i)$ are isolated and a neighbourhood basis of the point Ω is the family of the sets

$$U(\alpha_1, \dots, \alpha_{2n}) = \bigcup_{i=1}^{2n} \{(\beta, i); \beta \in W(\omega_i), \beta > \alpha_i\} \cup \{\Omega\} \qquad (\alpha_i \in W(\omega_i),\ 1 \leqq i \leqq 2n)$$

One readily shows that the space C_n defined in this manner is the one-point compactification of the discrete topological sum of the spaces $W(\omega_i)$ $(1 \leqq i \leqq 2n)$.

Evidently C_n is a totally disconnected compact Hausdorff space, thus $\dim C_n = 0$ indeed (see e.g. [7]).

a) $\operatorname{Dim} C_n \leqq n$.

Let i $(1 \leqq i \leqq n)$ be a natural number, $\alpha \in W(\omega_i)$, $\gamma \in W(\omega_{2n+1-i})$ and let us define

$$G_i^\alpha = ([0, \alpha) \times \{i\}), \quad F_i^\alpha = ([0, \alpha] \times \{i\}),$$

$$\Gamma_i^\gamma = C_n - ((0, \gamma] \times \{2n+1-i\}), \quad \Phi_i^\gamma = C_n - ([0, \gamma) \times \{2n+1-i\}),.$$

$$G_i = W_i, \quad F_i = C_n - W_{2n+1-i},$$

$\mathscr{R}_i = \{(\emptyset, \emptyset), (C_n, C_n)\} \cup \{(G_i^\alpha, F_i^\alpha); \alpha \in W(\omega_i)\} \cup \{(G_i, F_i)\} \cup \{(\Gamma_i^\gamma, \Phi_i^\gamma); \gamma \in W(\omega_{2n+1-i})\}$. It can easily be verified that $\mathscr{R}_i$ is a direction on C_n and the system $\mathfrak{R} = \{\mathscr{R}_i;\ 1 \leqq i \leqq n\}$ is a DS on the space.

b) $\operatorname{Dim} C_n \geqq n$.

Indeed, otherwise — making use of (2. 1) for the point Ω, as x — we would find a family $\{\mathfrak{U}_i;\ 1 \leqq i \leqq k\}$, where $k < 2n$, the systems $\mathfrak{U}_i$ are ordered by inclusion and $\mathfrak{U} = \bigcup \{\mathfrak{U}_i;\ i = 1, 2, \dots, k\}$ is a neighbourhood subbasis of Ω in C_n. Let $\mathscr{V}$ be one of these systems and $H = \bigcap \{V;\ V \in \mathscr{V}\}$; then, for all but one of the indices i $(1 \leqq i \leqq k)$, the set H contains a tail of W_i since, for two distinct indices i, j $(1 \leqq i \leqq k,\ 1 \leqq j \leqq k,\ i \neq j)$ $W(\omega_i)$ is not cofinal with ω_j.

Using the fact that $k < 2n$, we obtain that $\{\Omega\} \neq \bigcap \{U;\ U \in \mathfrak{U}\}$, consequently $\mathfrak{U}$ cannot be a neighbourhood subbasis of Ω in C_n. This contradiction completes the proof.

(3. 2) *Definition.* A topological space E is said to be *strongly metrizable* if E is a regular T_1-space and has a basis which is the union of a countable family of star-finite covers of E (see e.g. [6]).

(3. 3) *Definition.* Let Ω be any set. The product of countable many copies of the discrete space on Ω is called the *Baire space* on Ω and we denote it by $N(\Omega)$.

Each strongly metrizable space is metrizable by the NAGATA—SMIRNOV metrizability theorem, but the converse is not true. It is very easy to see that the product of countable many strongly metrizable spaces is also strongly metrizable and that a subspace of a strongly metrizable space is strongly metrizable. Since a discrete space and a separable metric space is evidently strongly metrizable, we obtain that for any set Ω each subspace of the space $N(\Omega)\times I^{\omega}$ is strongly metrizable.

A. ZARELUA showed in [5] that, for a strongly metrizable space X, $\operatorname{ind} X = \operatorname{Ind} X = \dim X$. Employing (1. 11), we obtain that, for strongly metrizable spaces, $\dim X \leqq \operatorname{Dim} X$ if $\operatorname{Dim} X < \aleph_0$. On the other hand, a theorem of J. NAGATA ([6]) asserts that, if X is strongly metrizable and $\dim X \leqq n$, then X is topologically embeddable into the space $N(\Omega)\times I_{2n+1}$ for a suitable Ω $(0 \leqq n < \aleph_0)$.

(3. 4) LEMMA. *For an arbitrary set* Ω

$$\operatorname{Dim} N(\Omega) \leqq 1.$$

PROOF. If $\Omega_1 \subset \Omega_2$, then $N(\Omega_1)$ is topologically embeddable into $N(\Omega_2)$, and so by (1. 7) it is sufficient to show the statement for infinite Ω. By (1. 14) is sufficient to prove that the topology of $N(\Omega)$ is induced by a suitable order. We choose an order $<$ on Ω, for which Ω has neither a first nor a last element, and denote by $\prec$ the lexicographic order on $N(\Omega)$. We will show that $\prec$ induces the original topology of $N(\Omega)$.

a) The order topology is coarser than the original topology.

If $\mathbf{x} = (x_k)_1^{\infty} \in N(\Omega)$ and M is a natural number, we denote by $U_M(\mathbf{x})$ the set $\{\mathbf{y} = (y_k) \in N(\Omega),\ x_l = y_l\ (l \leqq M)\}$.

If $\mathbf{a}, \mathbf{b}, \mathbf{x} \in N(\Omega)$,

$$\mathbf{a} = (a_n)_1^{\infty}, \quad \mathbf{b} = (b_n)_1^{\infty}, \quad \mathbf{x} = (x_n)_1^{\infty},$$

$\mathbf{a} \prec \mathbf{x} \prec \mathbf{b}$, then, for suitable natural numbers n, m, we have $a_n < x_n$, $x_m < b_m$, but $a_k = x_k$ for $k < n$, and $x_l = b_l$ for $l < m$.

Put $N = \max(n, m)$. The set $U_N(\mathbf{x})$ is a neighbourhood of $\mathbf{x}$ with respect to the product topology, and if

$$\mathbf{y} \in U_N(\mathbf{x}), \quad \text{then} \quad \mathbf{a} \prec \mathbf{y} \prec \mathbf{b}.$$

b) The order topology is finer that the product topology.

Let, for $\mathbf{x} = (x_n) \in N(\Omega)$, U be a neighbourhood of $\mathbf{x}$ in the product topology. For a suitable natural number M, $U_M(\mathbf{x}) \subset U$; we select elements $a, b \in \Omega$ for which $a < x_{M+1} < b$ holds. If we denote by $\mathbf{a}$ and $\mathbf{b}$ the points of $N(\Omega)$ for which

$$\pi_n(\mathbf{a}) = \begin{cases} x_n & (n \neq M+1) \\ a & (n = M+1) \end{cases} \qquad \pi_n(\mathbf{b}) = \begin{cases} x_n & (n \neq M+1) \\ b & (n = M+1) \end{cases}$$

(π_n denotes the n-th projection on Ω), then $\mathbf{a} \prec \mathbf{x} \prec \mathbf{b}$, and $\mathbf{a} \prec \mathbf{y} \prec \mathbf{b}$, $\mathbf{y} \in N(\Omega)$ implies $\mathbf{y} \in U_M(\mathbf{x}) \subset U$. Q.e.d.

(3. 5) THEOREM. *If X is a metrizable space and* $\dim X = 0$, *then* $\operatorname{Dim} X \leqq 1$.

PROOF. All such spaces are topologically embeddable into a space $N(\Omega)$ for a suitable Ω (see e.g. [7], Th. 7. 3. 9), and so by (3. 4) and (1. 7) the theorem is true.

(3. 6) THEOREM. *If X is a strongly metrizable space and one of the numbers* $\dim X$ *and* $\operatorname{Dim} X$ *is finite, then so is also the other and*

$$\dim X \leqq \operatorname{Dim} X \leqq 2 \dim X + 2.$$

PROOF. The assertion is a direct consequence of Lemma (3. 4), of the Theorems (1. 6) and (1. 7) and of the above-mentioned theorems of A. ZARELUA and J. NAGATA.

It is a very natural question whether the upper bound is reducible to $2\dim X+1$. We will answer this question in the negative; for each natural number n $(n\geqq 1)$ we find a strongly metrizable space X_n such that $\dim X_n=n$, $\operatorname{Dim} X_n = 2n+2$. We denote by $\mathcal{M}^n_{2n+1}$ the set of points of I_{2n+1} at most n of whose coordinates are rational. By a well known theorem ([4] Theorem V. 5) $\mathcal{M}^n_{2n+1}$ is an universal space for the at most n-dimensional separable metric spaces, thus (see [4], Ex. V. 3) $\mathcal{M}^n_{2n+1}$ is not embeddable into $\mathbf{R}^{2n}$. Making use of (1. 10), we obtain

$$\text{(3. 7)} \qquad \dim \mathcal{M}^n_{2n+1} = n, \quad \operatorname{Dim} \mathcal{M}^n_{2n+1} = 2n+1$$

Let Ω be any uncountable set; we assert that the space

$$\text{(3. 8)} \qquad X_n = N(\Omega)\times \mathcal{M}^n_{2n+1}$$

satisfies the above conditions. Now, X_n is topologically embeddable in $N(\Omega)\times I^{\omega}$; thus it is strongly metrizable; from the product theorem for dim[2] (e.g. [7] Th. 7. 3. 10 and 7. 3. 11) we obtain the relation $\dim X_n=n$. We shall now prove the equality $\operatorname{Dim} X_n = 2n+2$.

(3. 8) LEMMA. *Let X and Y be metric spaces such that*
a) *X is not locally separable,*
b) *Y is arcwise connected,*
c) *if $\emptyset\neq G\subset Y$, G is open in Y, then* $\operatorname{Dim} G=k<\aleph_0$.

Then $\operatorname{Dim} Z>k$ for $Z = X\times Y$.

PROOF. Let us suppose that the statement of the lemma is false; then by (1. 13) there exist order topological spaces $R_1, \ldots, R_k$ such that

$$Z\subset R = \mathop{\times}_{i=1}^{k} R_i$$

We shall denote by π_i the projection onto R_i and let x_0 be a point of X which has no separable neighbourhood. Since the space Y is connected, the sets $\pi_i(\{x_0\}\times Y)$ are connected sets J_i in R_i $(1\leqq i\leqq k)$. We assert that there exists a point $y_0\in Y$ such that $\pi_i((x_0, y_0))$ is an interior point of J_i $(1\leqq i\leqq k)$. Indeed, if $p\in R_i$, then the directional dimension of the space $\pi_i^{-1}(\{p\})$ is at most $(k-1)$ by (1. 6) hence — making use of the property c) of Y — the set $A_p = \pi_i^{-1}(\{p\})\cap(\{x_0\}\times Y)$ is nowhere dense closed set in $\{x_0\}\times Y$. Let p run over the end points of J_i for $i=1, \ldots, k$, and select

$$(x_0, y_0)\in(\{x_0\}\times Y)-\bigcup A_p$$

(This set is not empty because $\bigcup A_p$ is nowhere dense in $(\{x_0\}\times Y)$). Denote by $p_i\in R_i$ the point $\pi_i((x_0, y_0))$ $(1\leqq i\leqq k)$. We select for each i two points r_i, q_i of

[2] "For every countable family $\{X_i\}_{i=1}^{\infty}$ of metrizable spaces the conditions $\dim(\times_{i=1}^{\infty} X_i)=0$ and $\dim X_i=0$ $(i=1,2, \ldots)$ are equivalent."
"For every two metrizable spaces X and Y we have: $\dim X\times Y\leqq \dim X+\dim Y$.

$\pi_i(\{x_0\}\times Y)$ with $r_i<p_i<q_i$, and let u_i and v_i be points of $(\{x_0\}\times Y)$ for which

$$r_i=\pi_i(u_i), \quad q_i=\pi_i(v_i)$$

Now, by the arcwise connectedness of Y, there exists a Peano continuum I_i in $\{x_0\}\times Y$ joining u_i, (x_0, y_0) and v_i. The subspace $E_i=\pi_i(I_i)$ of R_i is a Peano continuum, hence it is a separable metrizable space and p_i is an interior point of E_i in R_i $(i=1, 2, \ldots, k)$. Now, if we denote by E the set $\bigtimes_1^k E_i$, then E is a separable metrizable neighbourhood of (x_0, y_0) in R; so the set $E\cap(X\times\{y_0\})$ is a separable neighbourhood of (x_0, y_0) in $X\times\{y_0\}$, which contradicts the definition of x_0. Q.e.d.

(3. 9) COROLLARY. Dim $X_n = 2n+2$ $(n\geqq 1)$.

PROOF. Indeed, if we denote by X the space $N(\Omega)$ $(\overline{\overline{\Omega}}>\aleph_0)$, by Y the space $\mathcal{M}^n_{2n+1}$ and by k the number $2n+1$, then the requirements of the lemma are fulfilled. The validity of the property a) is obvious. If G is a non empty open set in $\mathcal{M}^n_{2n+1}$, then G contains a topological image of $\mathcal{M}^n_{2n+1}$ and so by (3. 7) and (1. 7) Dim $G = 2n+1$ indeed.

Finally we must show that $\mathcal{M}^n_{2n+1}$ is arcwise connected. Let $\mathbf{x}=(x_k)$ denote a point in $\mathcal{M}^n_{2n+1}$ having only irrational coordinates. We will show that for each point $\mathbf{y}=(y_k)\in\mathcal{M}^n_{2n+1}$ there exists an arc joining $\mathbf{x}$ and $\mathbf{y}$ in $\mathcal{M}^n_{2n+1}$. We can suppose that only the first l coordinates of $\mathbf{y}$ are rational $(l\leqq n)$. If we denote by J_j $(j = 1, 2, \ldots, 2n+1)$ the set of all $\mathbf{z}=(z_k)\in I_{2n+1}$ where

$$z_k = \begin{cases} y_k, & \text{if } k<j \\ x_k, & \text{if } k>j \end{cases}$$

and z_j run over the interval with end points x_j and y_j, then J_j is an arc in $\mathcal{M}^n_{2n+1}$ and $\bigcup_{j=1}^{2n+1} J_j$ is a Peano continua joining x and y in $\mathcal{M}^n_{2n+1}$.

4. §

We see that for strongly metrizable spaces the directional dimension and the classical dimensions are "approximately equal" in the sense of (3. 6); but this is not true for general metric spaces.

E. DEÁK proved that if $S(A)$ denotes the star-space ("hedgehog" in the terminology of [7]) over the set A and $\overline{\overline{A}}>\aleph^+$, where $\aleph^+$ denotes the cardinal number following the continuum, then Dim $S(A)=\overline{\overline{A}}$, although for each set A ind $S(A)=$ $=$dim $S(A)=$Ind $S(A)=1$. One of his problems in [1] is the determination of the directional dimension of the star-spaces of any weight. It is very easy to see that if $\overline{\overline{A}}\leqq\aleph_0$, then Dim $S(A)\leqq 2$. More generally we shall deal with the directional dimension of arcwise connected metric spaces. By a well known theorem (see e.g. [7] ch. 6 Problem N) the continuous image of an arcwise connected space is also arcwise connected if it is Hausdorff, so by (1. 13) we need to characterize the arcwise connected order topological spaces.

(4. 1) *Definition.* Let W be the set of countable ordinal numbers and $L = W \times [0, 1)$; if (α, x), $(\beta, y) \in L$, then let $(\alpha, x) \prec (\beta, y)$ if and only if $\alpha < \beta$ or $\alpha = \beta$ and $x < y$.

The order topological space L obtained in this manner is called the *long halfline.*

By joining an inversely ordered long halfline to a long halfline, we obtain a *long line.*

It is very easy to see that a closed interval of a long line is homeomorphic with the interval $[0, 1]$ of the real line.

(4. 2) LEMMA. *If R is an arcwise connected order topological space, then R is topologically embeddable into the long line.*

PROOF. It is clear that an order topological space is arcwise connected if and only if each closed interval is an arc. Let now $x \in R$ be any point, obviously it is enough to prove that the subspace $P = \{y; y \in R, y \geqq x\}$ is topologically embeddable into L, the long halfline. But this fact was proved in [8].

(4. 3) LEMMA. *If W denotes the set of countable ordinals, then* $\operatorname{Dim} S(W) \leqq \aleph_0$.

PROOF. Let L denote the long halfline; by (1. 6) and (1. 7) it is sufficient to show that $S(W)$ is embeddable into $L^\omega \times I^\omega$.

We denote by I_α the interval

$$\{(\beta, x);\ (\beta, x) \in L,\ (\beta, x) \leqq (\alpha+1, 0)\} \quad \text{in } L \quad (\alpha \in W).$$

Now there exists for each natural number $n > 1$ and ordinal $\alpha \in W$ a homeomorphism $\varphi_n^\alpha: [0, 1] \to I_{\alpha+1}$ such that

$$\varphi_\alpha^n(0) = (0, 0), \quad \varphi_\alpha^n\left(\frac{1}{2n}\right) = (1, 0), \quad \varphi_\alpha^n\left(\frac{1}{n}\right) = (\alpha+1, 0), \quad \varphi_\alpha^n(1) = (\alpha+2, 0).$$

We denote by $f_n: S(W) \to (L \times I)$ $(n = 2, 3, \dots)$ the following mapping:

$$f_n\big((x, \alpha)\big) = \big(\varphi_\alpha^n(x), x\big) \qquad (\alpha \in W)$$

Let now f be the diagonal function

$$f = \mathop{\Delta}_{n=2}^{\infty} f_n : S(W) \to (L \times I)^{\aleph_0}$$

Making use of the Diagonal Lemma in [7], it is very easy to see that f is an embedding indeed.

(4. 4) COROLLARY. *Let X be a metric space. If $w(X) \leqq \aleph_1$; then* $\operatorname{Dim} X \leqq \aleph_0$.

PROOF. Indeed, all such spaces are topologically embeddable into $(S(W))^{\aleph_0}$ (see e.g. [7]).

(4. 5) THEOREM. *Let X be an arcwise connected metric space. Then*

$$\operatorname{Dim} X = \begin{cases} w(X) & \text{if} \quad w(X) > \aleph_1 \\ \aleph_0 & \text{if} \quad w(X) = \aleph_1 \end{cases}$$

PROOF. Let $\{R_\alpha; \alpha \in A\}$ be a family of order topological spaces, $\overline{\overline{A}} = \mathrm{Dim}\, X$, such that X is embeddable into $R = \times\{R_\alpha; \alpha \in A\}$. We denote by $\pi_\alpha: R \to R_\alpha$ the projection into R_α. We can suppose without loss of generality that $R_\alpha = \pi_\alpha(X)$, because is very easy to see, that a connected subspace of an orderable space is orderable, too. Now, each space R_α, as a continuous image of an arcwise connected space, is arcwise connected. Hence by the Lemma (4. 2) it is embeddable into the long line. Making use of the fact that the weight of the long line is $\aleph_1$ we obtain, that

$$w(X) \leqq w(R) \leqq \aleph_1 \cdot \overline{\overline{A}}$$

Therefore if $w(X) > \aleph_1$, then $\mathrm{Dim}\, X = \overline{\overline{A}} \geqq w(X)$, but $\mathrm{Dim}\, X \leqq w(X)$ so $\mathrm{Dim}\, X = w(X)$ indeed.

Let us now assume that $w(X)$ is equal to $\aleph_1$; by the Corollary (4. 4) we need only to show that Dim X is infinite. But otherwise X would be topologically embeddable into the product of a finite number of copies of the long line so it would be locally separable. By a well known theorem (see e.g. [7] ch. 4 Problem C) a locally separable connected metric space is separable, and this contradiction completes the proof of the theorem.

REFERENCES

[1] DEÁK, E.: Dimenzió és konvexitás I—IV. (Dimension and Convexity, in Hungarian) *MTA III. Osztály Közleményei* **17** (1967), **18** (1968).

[2] DEÁK, E.: *Dimension und Konvexität*. Akadémiai Kiadó, Budapest (to appear).

[3] DEÁK, E.: Eine vollständige Charakterisierung der Teilräume eines euklidischen Raumes mittels der Richtungsdimension *Publ. Math. Inst. Hung. Acad. Sci.* **9**. Series A (1964), 437—465.

[4] HUREWICZ, W.—WALLMAN, H.: *Dimension Theory*. Princeton Univ. Press. 1948.

[5] ZARELUA, A.: О теореме Гуревича, *Доклабы АН СССР* **141** (1961).

[6] NAGATA, J.: Note on dimension theory for metric spaces, *Fund. Math.* **45** (1958) 143—181.

[7] ENGELKING, R.: *Outline of general topology*. Warszawa 1968.

[8] THOMAS, J. P.: The long line as a subset of $P(R)$, *Amer. Math. Monthly* **76** (1969), 675—677.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received July 24, 1970)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 419—426

ABSOLUTE SUMMABILITY AND CONVERGENCE OF FOURIER SERIES

by

P. CHANDRA

Dedicated to the memory of late Professor Alfréd Rényi

1. *Definitions and Notations.* Let $L=L(w)$ be a continuous, differentiable and monotonic increasing function of w, and let it tend to infinity with w. Suppose that $\sum_{n=1}^{\infty} a_n$ is given infinite series, then

$$\sum_{n=1}^{\infty} a_n \quad \text{is summable} \quad |R, L, r| \qquad (r>0)$$

$$\left(\sum_{n=1}^{\infty} a_n \in |R, L, r|\right),$$

if

$$\int_A^{\infty} L'(w) L^{-r-1}(w) \Big| \sum_{n \leqq w} \{L(w)-L(n)\}^{r-1} L(n) a_n \Big| \, dw < \infty,$$

where A is a finite number.

Let $f(t)$ be a periodic function with period 2π and integrable in the sense of Lebesgue over $(-\pi, \pi)$. We can, without any loss of generality, write the Foureir series of $f(t)$ as

$$\sum_{n=1}^{\infty} (a_n \cos nt + b_n \sin nt) = \sum_{n=1}^{\infty} A_n(t),$$

assuming that the constant term is zero.

Throughout this paper we shall use the following notations; c is non-negative.

$$\Phi(t) = \frac{1}{2}\{f(x+t)+f(x-t)\}. \tag{1.1}$$

$$h(w, t) = \sum_{n \leqq w} n^{-1} L(n) \left(\log(n+1)\right)^{c-1} \sin nt. \tag{1.2}$$

2. *Introduction.* IZUMI [3] proved the following:

THEOREM A. *If, for* $c=0, 1,$

$$\int_0^{\pi} \left(\log \frac{2\pi}{t}\right)^c |d\Phi(t)| < \infty, \tag{2.1}$$

and

$$\{n^d \Delta(nA_n(x))\} \in BV, \quad \textit{for some} \quad d > 0, \text{ [1]} \tag{2.2}$$

then $\sum_{n=1}^{\infty} |A_n(x)| (\log(n+1))^{c-1} < \infty$.

Replacing the condition (2. 2) by the lighter condition (2. 4), MAZHAR [4] recently established the following theorem for the absolute convergence of Fourier series.

THEOREM B. *Let* $c=0, 1$ *and* $k \geqq \pi e^2$. *If*

$$\int_0^{\pi} \left(\log \frac{k}{t}\right)^c |d\Phi(t)| < \infty; \tag{2.3}$$

and

$$\left\{\exp(-n^a) \sum_{m=1}^{n} \exp(m^a) (\log(m+1))^{c-1} A_m(x)\right\} \in BV, \tag{2.4}$$

then $\sum_{n=1}^{\infty} |A_n(x)| (\log(n+1))^{c-1} < \infty$, *where* $0 < a < 1$.

In 1950, MOHANTY [5] gave the following criterion for the absolute convergence of Fourier series.

THEOREM C. Let $0 < \delta < 1$ and $k \geqq \pi e^2$. *If* (i) $\Phi(t) \log \frac{k}{t} \in BV(0, \pi)$ [2] *and* (ii) $\{n^{\delta} A_n(x)\} \in BV$, then $\sum_{n=1}^{\infty} |A_n(x)| < \infty$.

IZUMI [3] has shown that the conditions (2. 2) and (ii) of Theorem C are mutually exclusive and MAZHAR [4] has shown that condition (2. 2) implies (2. 4). Therefore naturally the question arises as to whether it is possible to establish any relationship between (ii) of Theorem C and (2. 4) for $c=1$. In this paper, Lemma 3, we have shown that the condition

$$\{n^{1-a} A_n(x) (\log n)^{-d}\} \in BV,$$

for finite d and $0 < a < 1$, implies

$$\left\{\exp(-n^a) \sum_{m=1}^{n} \exp(m^a) (\log m)^{-d} A_m(x)\right\} \in BV. \text{ [3]}$$

The technique used by MAZHAR [4] was to obtain the following result concerning the absolute Riesz summablity of Fourier series at a point and to deduce Theore B by means of a Tauberian theorem established by BHATT [1].

---

[1] $\{t_n\} \in BV$ we mean $\sum_n |\Delta t_n| < \infty$, where $\Delta t_n = t_n - t_{n+1}$.

[2] $f(x) \in BV(a, b)$ we mean $\int_a^b |df(x)| < \infty$.

[3] Also we can follow from Lemma 6 of the present paper.

THEOREM D. *Let* (2. 3) *for* $c=1$ *hold. If* $0<a<1$, *then* $\sum_{n=1}^{\infty} A_n(x) \in |R, \exp(n^a), 1|$.

Taking (2. 3) for $c=0$, the corresponding theorem has been established by MOHANTY [6] else where.

The purposes of this paper are the following:

1. First to bridge the gap $0<c<1$; indeed, we deduce it from a more general result established in Theorem 1.

2. We findout suitable absolute summability processes and absolute convergence factors (monotonic nature) on taking more general condition ((2. 3) for non-negative c) than MAZHAR [4] imposed upon the generating function of the Fourier series in Theorems D and B respectively.

3. We study the absolute summability and absolute convergence factor problems stated below and to provide suitable answers to these.

MAZHAR, in Theorem B, took the conditions (2. 3) for $c=0$ and 1. Corresponding to the condition (2. 3) for $c=0$, MAZHAR [4] obtained a suitable absolute convergence factor $\{\log(n+1)\}^{-1}$ under the condition (2. 4) for $c=0$. Now naturally the question arises as to whether it is possible to replace $\Phi(t) \in BV(0, \pi)$ by a still more general condition and to obtain suitable absolute summability and absolute convergence factors.

Theorems 1 and 2 of this paper provide an answer to this question. Indeed, the author has replace $\Phi(t) \in BV(0, \pi)$ by (2. 3) for non-negative c and has obtained a suitable absolute summability and absolute convergence factor sequence $\{\log(n+1)\}^{c-1}$.

We establish the following theorems.

THEOREM 1. *Let* $0<a<1$ *and* c *is non-negative. If the type of Riesz means* $L(w)$ *satisfies the following conditions:*

(2. 5) $\{L(w)/w(\log w)^{1-c}\}$ *is monotonic increasing with* $w \geqq w_0$; [4]

$$w^{1-a}L'(w) = O\{L(w)\}, \quad \text{as} \quad w \to \infty, \tag{2. 6}$$

and

$$\int_0^{\pi} \left(\log \frac{k}{t}\right)^c |d\Phi(t)| < \infty, \tag{2. 7}$$

then $\sum_{n=1}^{\infty} A_n(x)\left(\log(n+1)\right)^{c-1} \in |R, L(w), 1|$.

THEOREM 2. *If, for* non-negative c *and* $0<a<1$,

(2. 7) *holds and*

$$\{n^{1-a}A_n(x)\left(\log(n+1)\right)^{c-1}\} \in |R, \exp(n^a), 1|, \; ^5 \tag{2. 8}$$

[4] In the case $\{L(w)|w(\log w)^{1-c}\}$ is monotonic decreasing with $w \geqq w_0$, the result follows by using the second theorem of consistency for absolute Riesz summability.

[5] In view of Lemma 5 of this paper, this can be replaced by $\left\{\exp(-n^a) \sum_{m=1}^{n} \exp(m^a) A_m(x) \left(\log(m+1)\right)^{c-1}\right\} \in BV$.

then $\sum_{n=1}^{\infty} |A_n(x)| (\log(n+1))^{c-1} < \infty.$

3. We shall use the following lemmas.

LEMMA 1. *If, for non-negative c, L(n) satisfies* (2. 5), *then*

$$\sum_{n \leqq w} L(n) \sin nt/n(\log(n+1))^{1-c} = O\{t^{-1} w^{-1} L(w) (\log w)^{c-1}\},$$

uniformly in $0<t<\pi$.

PROOF. By using the Abel's lemma and the condition (2. 5), the result follows.

LEMMA 2 [5]. *If* $\{s_n\} \in BV$, *then*

$$\left\{(P_n)^{-1} \sum_{m=1}^{n} p_m s_m\right\} \in BV,$$

where $P_n = \sum_{m=1}^{n} p_m \to \infty$, as $n \to \infty$ *and* $p_m > 0$.

LEMMA 3. *If* $\{n^{1-a} A_n(x) (\log n)^{-d}\} \in BV$, *then*

$$\{T_n\} = \left\{\exp(-n^a) \sum_{m=2}^{n} \exp(m^a) (\log m)^{-d} A_m(x)\right\} \in BV,$$

where $0<a<1$ *and d is finite.*

PROOF. We have, for $t_m = m^{1-a} A_m(x) (\log m)^{-d}$,

$$T_n = \exp(-n^a) \sum_{m=2}^{n} m^{a-1} \exp(m^a) t_m =$$

$$= \frac{\sum_{m=2}^{n} \left\{\frac{m^{a-1} \exp(m^a)}{\exp(m^a) - \exp\{(m-1)^a\}}\right\} t_m \{\exp(m^a) - \exp\{(m-1)^a\}\}}{\sum_{m=2}^{n} \{\exp(m^a) - \exp\{(m-1)^a\}\}} \cdot \{1 - e \cdot \exp(-n^a)\} =$$

$$= R_n \cdot Q_n.$$

Since $\{t_m\} \in BV$, $\{m^{a-1} \exp(m^a)/(\exp(m^a) - \exp((m-1))^a\} \in BV$ and $\sum_{m=2}^{n} \{\exp(m^a) - \exp((m-1)^a)\} \to \infty$, as $n \to \infty$, the sequence $\{R_n\} \in BV$, by Lemma 2. Also

$$\{Q_n\} = \{1 - e \cdot \exp(-n^a)\} \in BV.$$

Therefore $\{T_n\} \in BV$, completes the proof of the lemma.

LEMMA 4 (DAS [2], Lemma 17.) [6]. *For any sequence* $\{L_n\}$, $\{b_n\} \in |R, L_n, 1|$ *and* $\{d_n\} \in BV$ *imply* $\{b_n d_n\} \in |R, L_n, 1|$.

[6] This is proved for L_{n-1} but, by similar arguments, Lemma also holds good for L_n.

LEMMA 5.[7] *If* $\sum_{n=1}^{\infty} a_n \in |R, L_n, 1|$, *then a necessary and sufficient condition for the absolute convergence of the series* $\sum_{n=1}^{\infty} a_n$ *is*

$$\left\{L_{n+1}^{-1} \sum_{m=1}^{n} L_m a_m\right\} \in BV.$$

PROOF. It has been observed by BOSANQUET (see MOHANTY [6]), that the methods, $|R, L_n, 1|$ and $|R', L_n, 1|$ equivalent, where $(R', L_n, 1)$ mean of $\sum_{n=1}^{\infty} a_n$ is

$$t_n = L_{n+1}^{-1} \sum_{m=1}^{n} l_{m+1} s_m, \qquad l_{m+1} = L_{m+1} - L_m.$$

By Abel's transformation, we have

$$t_n = L_{n+1}^{-1} \sum_{m=1}^{n-1} \Delta s_m \sum_{k=1}^{m} l_{k+1} + s_n L_{n+1}^{-1} \sum_{m=1}^{n} l_{m+1}$$

$$= -L_{n+1}^{-1} \sum_{m=1}^{n-1} L_{m+1} a_{m+1} + (s_n - s_1) l_1 L_{n+1}^{-1} + s_n (1 - l_1 L_{n+1}^{-1})$$

$$= s_n - L_{n+1}^{-1} \sum_{m=1}^{n-1} L_{m+1} a_{m+1} - a_1 L_1 / L_{n+1} = s_n - L_{n+1}^{-1} \sum_{m=1}^{n} L_m a_m.$$

Now if $\sum_{n=1}^{\infty} a_n \in |R, L_n, 1|$, then $\sum_{n=1}^{\infty} |a_n| < \infty$, iff

$$\left\{(L_{n+1})^{-1} \sum_{m=1}^{n} L_m a_m\right\} \in BV.$$

LEMMA 6. *If* $\{n^{1-a} A_n(x) (\log (n+1))^{c-1}\} \in |R, \exp (n^a), 1|$ $(0<a<1)$, *then*

$$\left\{\exp(-(n+1)^a) \sum_{m=1}^{n} \exp(m^a) A_m(x) (\log(m+1))^{c-1}\right\} \in BV,$$

where c is non-negative.

PROOF. Since, for $L_n = \exp(n^a)$ $(0<a<1)$, $\{n^{a-1} L_n / (L_{n+1} - L_n)\} \in BV$ and $\{n^{1-a} A_n(x) (\log (n+1))^{c-1}\} \in |R, \exp (n^a), 1|$, we have by Lemma 4

$$\{L_n A_n(x) (\log(n+1))^{c-1} l_{n+1}^{-1}\} \in |R, \exp(n^a), 1|,$$

which completes the proof of the lemma.

---

[7] Compare with BHATT [1].

4. Proof of Theorem 1. We have

$$A_n(x) = \frac{2}{\pi}\int_0^\pi \Phi(t)\cos nt\,dt = -\frac{2}{\pi}\int_0^\pi (\sin nt/n)\,d\Phi(t),$$

integrating by parts.

The series $\sum_{n=1}^{\infty} A_n(x)\left(\log(n+1)\right)^{c-1} \in |R, L(w), 1|$, if

$$I = \frac{2}{\pi}\int_1^\infty \frac{L'(w)}{\{L(w)\}^2}\left|\sum_{n\leqq w}\frac{L(n)}{\{\log(n+1)\}^{1-c}}\int_0^\pi \frac{\sin nt}{n}\,d\Phi(t)\right| dw < \infty.$$

Now,

$$I \leqq \frac{2}{\pi}\int_0^\pi |d\Phi(t)| \int_1^\infty L'(w)\left(L(w)\right)^{-2}|h(w,t)|\,dw.$$

For the proof of the theorem, it is sufficient to show that

$$I' = \int_1^\infty L'(w)\left(L(w)\right)^{-2}|h(w,t)|\,dw = O\left\{\left(\log(k/t)\right)^c\right\},$$

uniformly in $0<t<\pi$.

For $T_1 = k/t$ and $T_2 = (k/t)^{-1/(1-a)}$, we have

$$I' = \left(\int_1^{T_1} + \int_{T_1}^{T_2} + \int_{T_2}^{\infty}\right)\left(L'(w)\left(L(w)\right)^{-2}|h(w,t)|\,dw\right) = I_1+I_2+I_3.$$

Now, by using the fact that $|\sin nt| \leqq nt$, we have

$$I_1 = O\left\{t\int_1^{T_1} L'(w)\left(L(w)\right)^{-2}\left|\sum_{n\leqq w} L(n)\left(\log(n+1)\right)^{c-1}\right| dw\right\}.$$

Therefore, by (2. 5) and (2. 6), we have

$$I_1 = O\left\{t\int_1^{T_1} L'(w)\left(L(w)\right)^{-2} dw \int_1^w L(x)\left(\log(x+1)\right)^{c-1} dx\right\}$$

$$+O\left\{t\int_1^{T_1} L'(w)\left(L(w)\right)^{-1}\left(\log(w+1)\right)^{c-1} dw\right\} + O(t)$$

$$= O\left\{t\int_1^{T_1} L(x)\left(\log(x+1)\right)^{c-1} dx \int_x^{T_1} L'(w)\left(L(w)\right)^{-2} dw\right\}$$

$$+O\left\{t\int_1^{T_1} w^{a-1}\left(\log(w+1)\right)^{c-1} dw\right\} + O(t)$$

$$= O\left\{t\int_1^{T_1}\left(\log(x+1)\right)^{c-1} dx\right\} + O\left\{t\int_1^{T_1} w^{a-1}\left(\log(w+1)\right)^{c-1} dw\right\} + O(t)$$

$$= O\left\{\left(\log(k/t)\right)^c\right\},$$

uniformly in $0<t<\pi$. And, since $\sin nt = O(1)$,

$$I_2 = O\Big\{\int_{T_1}^{T_2} L'(w)\,(L(w))^{-2}\Big|\sum_{n\leqq w} L(n)n^{-1}(\log(n+1))^{c-1}\Big|\,dw\Big\}$$

$$= O\Big\{\int_{T_1}^{T_2} L'(w)\,(L(w))^{-2}\,dw\int_1^w L(x)x^{-1}(\log(x+1))^{c-1}\,dx\Big\}$$

$$+O\Big\{\int_{T_1}^{T_2} L'(w)\,(wL(w))^{-1}(\log(w+1))^{c-1}\,dw\Big\}+O\{(\log(k/t))^c\}$$

$$= O\{I_{2,1}\}+O\{I_{2,2}\}+O\{(\log(k/t))^c\}.$$

Now, by (2. 6), we have

$$I_{2,2} = O\Big\{\int_{T_1}^{T_2} w^{a-2}(\log(w+1))^{c-1}\,dw\Big\} = O\{(\log(k/t))^c\},$$

and by the change of order of integration, we have

$$I_{2,1} = O\Big\{\int_{T_1}^{T_2} L(x)x^{-1}(\log(x+1))^{c-1}\,dx\int_x^{T_2} L'(w)\,(L(w))^{-2}\,dw\Big\}$$

$$+O\Big\{\int_{T_1}^{T_2} L'(w)\,(L(w))^{-2}\,dw\int_1^{T_1} L(x)x^{-1}(\log(x+1))^{c-1}\,dx\Big\}$$

$$= O\Big\{\int_{T_1}^{T_2} x^{-1}(\log(x+1))^{c-1}\,dx\Big\}+O\{(\log(k/t))^c\} = O\{(\log(k/t))^c\},$$

uniformly in $0<t<\pi$. Hence combining $I_{2,1}$ and $I_{2,2}$, we have

$$I_2 = O\{(\log(k/t))^c\},$$

uniformly in $0<t<\pi$. And finally, by Lemma 1 and by (2. 6), we have

$$I_3 = O\Big\{t^{-1}\int_{T_2}^{\infty} L'(w)\,(L(w))^{-1}w^{-1}(\log w)^{c-1}\,dw\Big\} = O\Big\{t^{-1}\int_{T_2}^{\infty} w^{a-2}(\log w)^{c-1}\,dw\Big\}$$

$$= O\{t^{-1}T_2^{a-1}|\log T_2|^{c-1}\} = O\{(\log(k/t))^c\},$$

uniformly in $0<t<\pi$.

This terminates the proof of Theorem 1.

4. 1. Corollary of Theorem 1. *If, for non-negative c and $0<a<1$, (2. 7) holds. Then the series*

$$\sum_{n=1}^{\infty} A_n(x)\,(\log(n+1))^{c-1}\in|R,\,\exp(w^a),\,1|.$$

5. PROOF of Theorem 2. Since, by using the standard definition of Riesz means, the series

$$\sum_{n=1}^{\infty} A_n(x)\left(\log(n+1)\right)^{c-1} \in |R, \exp(n^a), 1| \qquad (0<a<1)$$

whenever (2. 7) holds, by the corollary of Theorem 1, the proof of the theorem follows by using Lemma 4 and Lemma 3 of the present paper.

REFERENCES

[1] BHATT, S. N.: A Tauberian theorem for absolute Riesz summability, *Indian Jour. Math.* **1** (1958) 29—32.

[2] DAS, G.: Tauberian theorems for absolute Nörlund summability, *Proc. London Math. Soc.* (3), XIX (1969) 357—385.

[3] IZUMI, S.: Absolute convergence of some trigonometric series II, *Jour. Math. Analysis and Applications* **1** (1960) 184—194.

[4] MAZHAR, S. M.: On the absolute convergence of Fourier series, *Proc. Japan Acad.* **44** (1968) 756—761.

[5] MOHANTY, R.: A criterian for absolute convergence of Fourier series, *Proc. London Math. Soc.* **51** (1950) 181—196.

[6] MOHANTY, R.: On the absolute Riesz summability of Fourier series and allied series, *Proc. London Math. Soc.* **52** (1951) 295—320.

402, West-Ghamapur, Jabalpur, M. P., India

(Received September 15, 1970)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 427—429.

ÜBER EINE LOKALE VARIANTE DES PRIWALOWSCHEN SATZES

von

M. SALLAY

Es sei $f(x)$ eine 2π periodische Funktion die in dem Intervall $[0, 2\pi]$ eine Lipschitz-Bedingung β-ter Ordnung erfüllt $(0<\beta<1)$. Bezeichnen wir mit $\tilde{f}(x)$ die konjugierte Funktion von $f(x)$. Nach dem Satz von I. I. PRIWALOW [3] erfüllt $\tilde{f}(x)$ in $[0, 2\pi]$ ebenfalls eine Lipschitz-Bedingung derselben Ordnung β.

Setzen wir nun voraus, daß $f(x)$ in einem einzigen Punkte $\xi\in[0, 2\pi]$ für ein $\alpha>\beta$ auch ein lokale Lipschitz-Bedingung erfüllt, d.h. es gibt eine positive Zahl $\delta>0$ derart, daß für die Punkte $\xi+h$ mit $|h|\leqq\delta$ die Relation

$$(1)\qquad |f(\xi+h)-f(\xi)|\leqq M_1|h|^\alpha \qquad (M_1=\text{konst.})$$

gilt. Es stellt sich die Frage, was man unter der Bedingung (1) über den lokalen Lipschitz-Exponent von $\tilde{f}(\xi)$ aussagen kann?

G. FREUD hat in seiner Arbeit [2] die folgende Erweiterung des Priwalowschen Satzes bewiesen: Es sei $f(x)\in C_{2\pi}$, $f\in\text{Lip}\,\beta$; ferner sei $m_0(f)$ die Menge der Punkte x für welche gleichmäßig in x die Relation $f(x+h)-f(x)=o(|h|^\beta)$ $(h\to 0)$ gilt. Es sei $\tilde{f}(x)$ die harmonische Konjugierte von $f(x)$. Die Mengen $m_0(f)$ und $m_0(\tilde{f})$ sind äquivalent, d.h. $|[m_0(f)\setminus m_0(\tilde{f})]\cup[m_0(\tilde{f})\setminus m_0(f)]|=0.$[1] In verbindung damit hat A. A. GONTSCHAR die Obenstehende Frage gestellt. Auf das Problem hat meine Aufmerksamkeit G. FREUD gelenkt.

Wir beweisen den folgenden:

SATZ: *Es sei in dem Intervall* $[0, 2\pi]$ $f(x)\in\text{Lip}\,\beta$ $(0<\beta<1)$ *und befriedige in dem Punkte* ξ *auch die Bedingung* (1). *Dann gibt es eine Umgebung* $\delta_1>0$ *des Punktes* ξ, *so dass für* $|h|\leqq\delta_1$ *die Beziehung*

$$(2)\qquad |\tilde{f}(\xi+h)-\tilde{f}(\xi)|\leqq M_2|h|^\gamma$$

gilt mit

$$(3)\qquad \gamma=\beta+\beta\frac{\alpha-\beta}{1+\beta}\qquad (\beta<\gamma<\alpha).$$

Beweis: Bezeichnen wir mit $\sigma_n(f; x)$ das n-te Fejérsche Mittel der Fourierschen Reihe von $f(x)$ und mit $\sigma_n(\tilde{f}; x)$ das n-te Fejérsche Mittel der konjugierten Funktion $\tilde{f}(x)$.

Nun besteht für beliebige n $(n=1, 2, \dots)$

$$(4)\qquad \tilde{f}(\xi+h)-\tilde{f}(\xi)=[\tilde{f}(\xi+h)-\sigma_n(\tilde{f};\xi+h)]-[\tilde{f}(\xi)-\sigma_n(\tilde{f};\xi)]+\int_\xi^{\xi+h}\sigma_n'(\tilde{f},t)\,dt.$$

[1] G. FREUD hat den Satz für eine allgemeinere Funktionklasse bewiesen.

Da die Funktion $\tilde{f}(x)$ im Intervall $[0, 2\pi]$ eine Lipschitz-Bedingung β-ter Ordnung erfüllt, besteht in $[0, 2\pi]$ gleichmäßig (S., S. N. BERNSTEIN [1])

$$|\tilde{f}(x) - \sigma_n(\tilde{f}; x)| \leqq k_1 n^{-\beta}. \tag{5}$$

(Mit k_ν $(\nu = 1, 2, \ldots)$ bezeichnen wir von n und x unabhängige Konstanten.)

Es bezeichne $\sigma_n^{(2)}(f; x)$ das n-te $(C, 2)$-Mittel der Fourierreiche von $f(x)$, dann gilt die Formel (Vgl. z. B. A. ZYGMUND [4] S. 269).

$$\begin{aligned}\sigma_n'(\tilde{f}; x) &= (n+2)[\sigma_n(f; x) - \sigma_n^{(2)}(f; x)] = \\ &= (n+2)\sigma_n(f; x) - \frac{2}{n+1}\sum_{k=0}^{n}(k+1)\sigma_k = n\sigma_n(f; x) - \frac{2}{n+1}\sum_{k=0}^{n-1}(k+1)\sigma_k = \\ &= n[\sigma_n(f; x) - f(x)] - \frac{2}{n+1}\sum_{k=0}^{n-1}(k+1)[\sigma_k(f; x) - f(x)].\end{aligned} \tag{6}$$

Es sei nun $x = \xi + h$ mit $|h| \leqq \frac{\delta}{2}$. Da

$$\sigma_n(f; \xi+h) - f(\xi+h) = \frac{2}{\pi}\int_0^\pi [f(\xi+h+t) + f(\xi+h-t) - 2f(\xi+h)]K_n(t)\,dt \tag{7}$$

ist und

$$0 \leqq K_n(t) \leqq k_3 \min\left(n, \frac{1}{nt^2}\right) \tag{8}$$

gilt, besteht wegen (1), (5), (7) und (8) die folgende Abschätzung:

$$\begin{aligned}&|\sigma_n(f; \xi+h) - f(\xi+h)| \leqq \\ &\leqq k_4\Big[\int_0^{1/n}(|h|+t)^\alpha K_n(t)\,dt + \int_{1/n}^{\delta/2}(|h|+t)^\alpha K_n(t)\,dt + \int_{\delta/2}^{\pi} t^\beta K_n(t)\,dt\Big] \leqq \\ &\leqq k_5\Big[n\int_0^{1/n}(|h|+t)^\alpha\,dt + \int_{1/n}^{\delta/2}(|h|+t)^\alpha \frac{1}{nt^2}\,dt + \frac{1}{n}\int_{\delta/2}^{\pi} t^{\beta-2}\,dt\Big] \leqq \\ &\leqq k_6[n^{-\alpha} + |h|^\alpha + n^{-1}\delta^{\beta-1}].\end{aligned} \tag{9}$$

Aus den Beziehungen (5), (6), (4) und (9) folgt somit für $|h| \leqq \frac{\delta}{2}$

$$|\tilde{f}(\xi+h) - \tilde{f}(\xi)| \leqq k_7[n^{-\beta} + n|h|^{1+\alpha} + n|h|^{-\alpha}n^{-\alpha} + |h|\delta^{\beta-1}]. \tag{10}$$

Es sei nun

$$n = \big[|h|^{-1-\frac{\alpha-\beta}{1+\beta}}\big] + 1.$$

Dann ist für $|h| \leqq \delta_1 < \left(\frac{\delta}{2}\right)^{\frac{1-\beta^2}{1-\alpha\beta}}$ (2) mit (3) begfriedigt, w. z. b. w.

LITTERATURVERZEICHNIS

[1] Бернщейн С. Н.: О наилучшем приближении непрерывных функций посредством многочленов данной степени. *Сообщ. Харьк. Матем. об.—ва. сер.* 2 T. **13** (1912) 49—194.

[2] Freud, G.: An approximation theoretical study of the structure of real functions. *Studia Sci. Math. Hungar.* **5** (1970) 141—150.

[3] Priwalow, I. I.: Sur les fonctions conjugées. *Bull. de la Soc. Math. de France* **44** (1916).

[4] Zygmund, A.: *Trigonometric Series 2.* Aufl. University Press Cambridge 1959.

Mathematisches Institut der Ungarischen Akademie der Wissenschaften, Budapest

(Eingegangen: 16. September 1970.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 431—438.

BEWEGUNGSSTABILE PACKUNGEN KONSTANTER NACHBARNZAHL

von

G. WEGNER

1. Einleitung

Unter einer Scheibe in der euklidischen Ebene wollen wir eine kompakte, konvexe Punktmenge verstehen. Eine Familie von Scheiben, die paarweise keine inneren Punkte gemeinsam haben, heiße Packung. Zwei Scheiben einer Packung mit mindestens einem gemeinsamen Randpunkt nennen wir benachbart und wir sprechen von einer n-Nachbarnpackung, wenn jedes Mitglied der Familie dieselbe Anzahl n von Nachbarn hat. Eine Packung kann translationsstabil bzw. bewegungsstabil sein; jede Scheibe der Packung wird dann von ihren Nachbarn gegenüber Translationen bzw. Bewegungen fixiert. L. FEJES TÓTH hat in [1] nachgewiesen, daß es für jedes $n \geqq 3$ translationsstabile n-Nachbarnpackungen kongruenter Scheiben gibt und (neben einigen anderen Problemen) die Frage aufgeworfen, ob es auch bewegungsstabile n-Nachbarnpackungen kongruenter Scheiben für jedes $n \geqq 3$ gibt. Wir werden zeigen, daß dies tatsächlich zutrifft.

2. Bewegungsstabile Packungen

SATZ: *Zu jeder natürlichen Zahl $n \geqq 3$ existiert in der euklidischen Ebene eine bewegungsstabile n-Nachbarnpackung kongruenter Scheiben.*

BEWEIS: Fig. 1 zeigt eine Packung der verlangten Art für $n=11$. Man erhält sie aus der bekannten 16-Nachbarnpflasterung (Fig. 8 in [1]), indem man einen Teil der Dreiecke wegläßt und von den verbleibenden die rechten Winkel abschneidet. Da außerdem für $n=3$ bis 10 (und für einige Werte $n \geqq 12$) schon in [1] bewegungsstabile n-Nachbarnpackungen kongruenter Scheiben angegeben wurden, können wir uns im folgenden auf $n \geqq 12$ beschränken.

Wir beginnen mit der Konstruktion einer speziellen bewegungsstabilen 12-Nachbarnpackung. Wir gehen hierzu wieder aus von der eben erwähnten 16-Nachbarnpflasterung. Wir fassen eines der Dreiecke $\langle abc \rangle$ [1] ins Auge (Fig. 2), schneiden hiervon das Dreieck $\langle acd \rangle$ ab, wobei $\sphericalangle\, bad$ ein gegebener Winkel α zwischen 0° und 30° ist, ferner das Dreieck $\langle bef \rangle$, wobei e die Strecke $\langle ab \rangle$ im Verhältnis $(\sqrt{3}-1):(2-\sqrt{3})$ teilt und $\sphericalangle\, bef = 2\alpha$ ist. Damit erhalten wir ein Viereck $V := \langle adfe \rangle$ und die analoge Reduzierung aller Dreiecke der 16-Nachbarnpflasterung führt zu einer 12-Nachbarnpackung (in Fig. 2 stark gezeichnet), welche unabhängig vom Winkel α bewegungsstabil ist. Um dies einzusehen, vermerken wir zunächst,

[1] Mit $\langle x_1 \dots x_n \rangle$ wird das von den Ecken $x_1, \dots x_n$ aufgespannte n-Eck bezeichnet; insbesondere ist $\langle xy \rangle$ die Strecke mit den Endpunkten x und y.

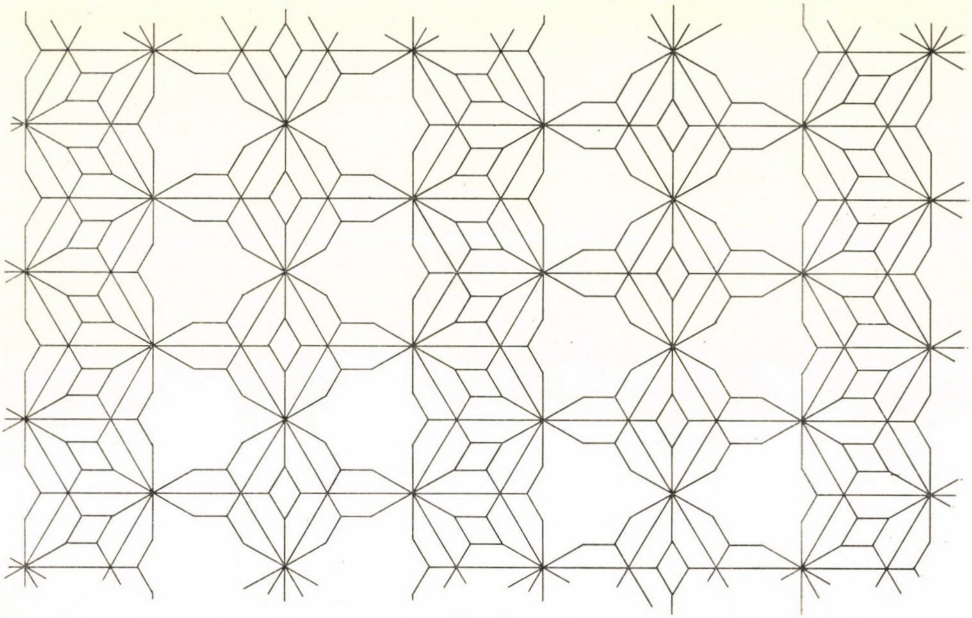


Fig. 1

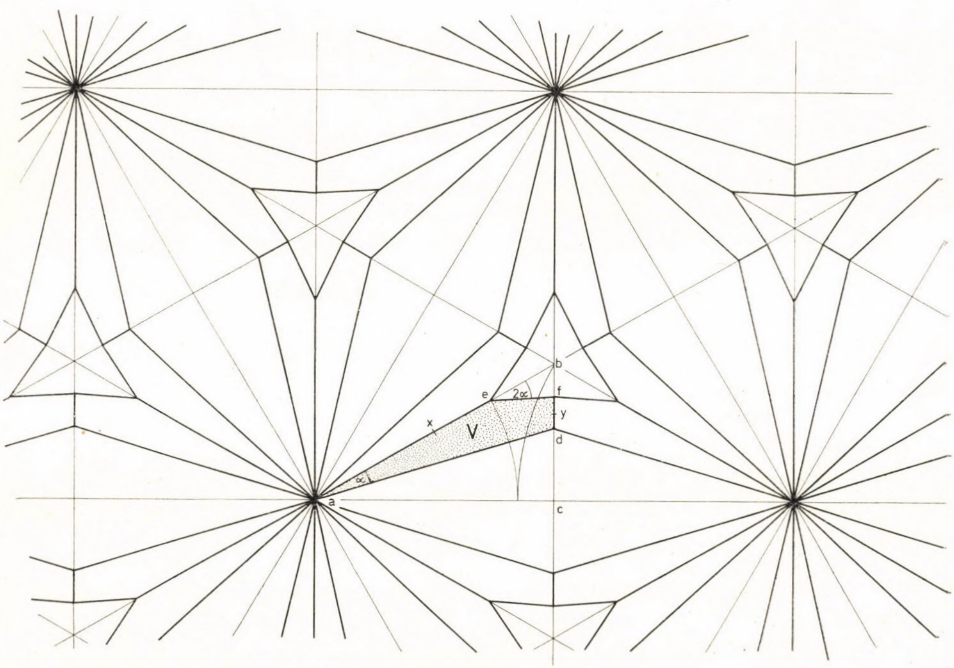


Fig. 2

daß die Vierecke der Packung untereinander äquivalent sind; es genügt also, die Stabilität von V nachzuweisen. Man betrachte hierzu drei Punkte von V, nämlich die Ecke a, den Punkt $x := \frac{a+b}{2}$ und einen Punkt y im Innern der Strecke $\langle df \rangle$. Bei einer Bewegung von V sind diese Punkte Funktionen der Zeit: $a(t)$, $x(t)$ und $y(t)$, und wenn $t=0$ der durch die Packung gegebenen Lage entspricht, so ist für ein geeignetes Koordinatensystem

$$a(0) = (0, 0), \quad x(0) = \frac{1}{2}(\sqrt{3}, 1), \quad y(0) = (\sqrt{3}, \tau)$$

mit geeignetem τ zwischen 0 und 1. Setzen wir ferner $u := (\sqrt{3}, 1)$ und $v := (\sqrt{3}, -1)$, so führen die Nachbarn von V in der Packung zu folgenden Beschränkungen für die Tangentialvektoren der Bewegungskurven der Punkte a, x und $y = (y_1, y_2)$ [2]; wobei also $\dot{\varphi}(0)$ die einseitige Ableitung $\lim\limits_{t \searrow 0} t^{-1}(\varphi(t) - \varphi(0))$ bezeichnet:

$$\begin{aligned} \dot{a}(0) &= \varrho u + \sigma v \quad \text{mit} \quad \varrho \geqq 0, \quad \sigma \geqq 0 \\ \dot{x}(0) &= \varrho' u + \sigma' v \quad \text{mit} \quad \sigma' \geqq 0 \\ \dot{y}_1(0) &\leqq 0 \end{aligned}$$

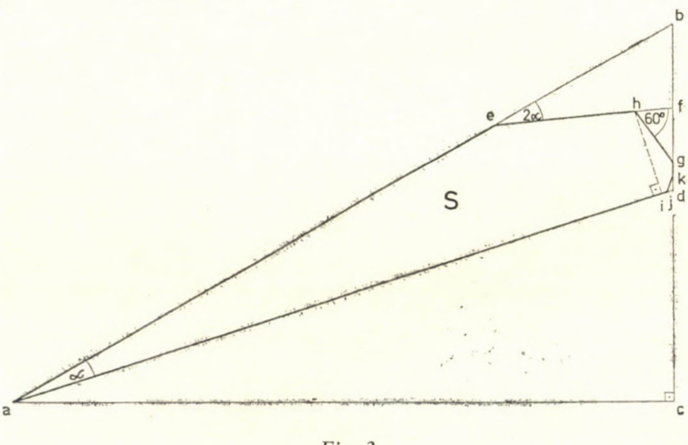


Fig. 3

Eine einfache Rechnung zeigt, daß diese Bedingungen nur mit $\dot{a}(0) = \dot{x}(0) = \dot{y}(0) = 0$ zu erfüllen sind, was die Bewegungsstabilität der Packung beweist.

Es sei vermerkt, daß man die Vierecke noch weiter verkleinern kann, ohne die Stabilität der Packung zu beeinträchtigen, solange V nur die Punkte a und x und mehr als einen Punkt von $\langle df \rangle$ mit den jeweiligen Nachbarn gemeinsam hat. Dies nützen wir aus, um bewegungsstabile n-Nachbarnpackungen für $n > 12$ zu konstruieren. Wir schneiden von V das Dreieck $\langle fgh \rangle ab$ (Fig. 3), wobei $g = \frac{1}{3}(f + 2d)$

[2] Tatsächlich sind die Beschränkungen sogar noch stärker (allerdings in Abhängigkeit von α), als dies in den Ungleichungen zum Ausdruck kommt.

und $\sphericalangle fhg = 60°$. Der Fußpunkt i des Lotes von h auf die Gerade $\overline{ad}$ liegt dann im Innern von $\langle ad \rangle$ und wir können noch mit einer Geraden, die i und g streng von d separiert, die Ecke d abschneiden. Legen wir die durch diesen Schnitt entstehenden Ecken j und k auf $j = \frac{d+i}{2}$ und $k = \frac{d+g}{2}$ fest, so erhalten wir zu jedem Winkel α mit $0° < \alpha < 30°$ ein eindeutig bestimmtes Sechseck $S := S_\alpha := \langle ajkghe \rangle$, welches, wie wir oben vermerkt haben, analog zu Fig. 2 Anlaß zu einer bewegungsstabilen 12-Nachbarnpackung gibt. Wir betrachten in dieser Packung den von zwei benachbarten Rosetten eingegrenzten Raum (unter einer Rosette in einer doppeltperiodischen Packung wollen wir eine maximale Teilfamilie von Mitgliedern verstehen, die einen Punkt gemeinsam haben). Zu gegebenem $n > 12$ existiert ein eindeutig bestimmter Winkel α derart, daß sich in diesen Zwischenraum $2n-24$ weitere Exemplare von S so einlagern lassen, daß (vgl. Fig. 4 bis 6 für $n = 13$ bis 15)

a) jeder der beiden Rosetten $n-12$ neue Mitglieder zugeordnet werden,

b) jedes der neuen Sechsecke die der längsten Seite $\langle aj \rangle$ von S entsprechende Seite mit einem anderen Mitglied gemeinsam hat,

c) jedes Mitglied einen Punkt mit (genau) einem Mitglied der anderen Rosette gemeinsam hat.

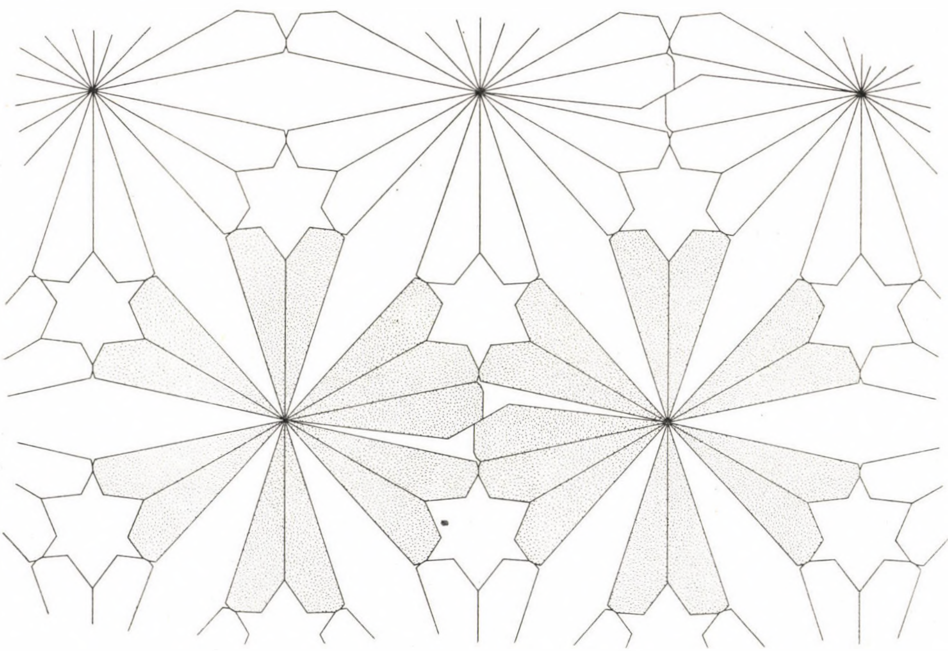


Fig. 4

Für die neuen Sechsecke, ausgenommen die beiden mittleren Sechsecke bei ungeradem n, ist dieser Berührpunkt eindeuting bestimmt, und zwar trifft das eine Mitglied mit der h entsprechenden Ecke das andere Mitglied im Innern der $\langle eh \rangle$ entsprechenden Seite. Die neu hinzukommenden Sechsecke werden also wechselseitig paarweise

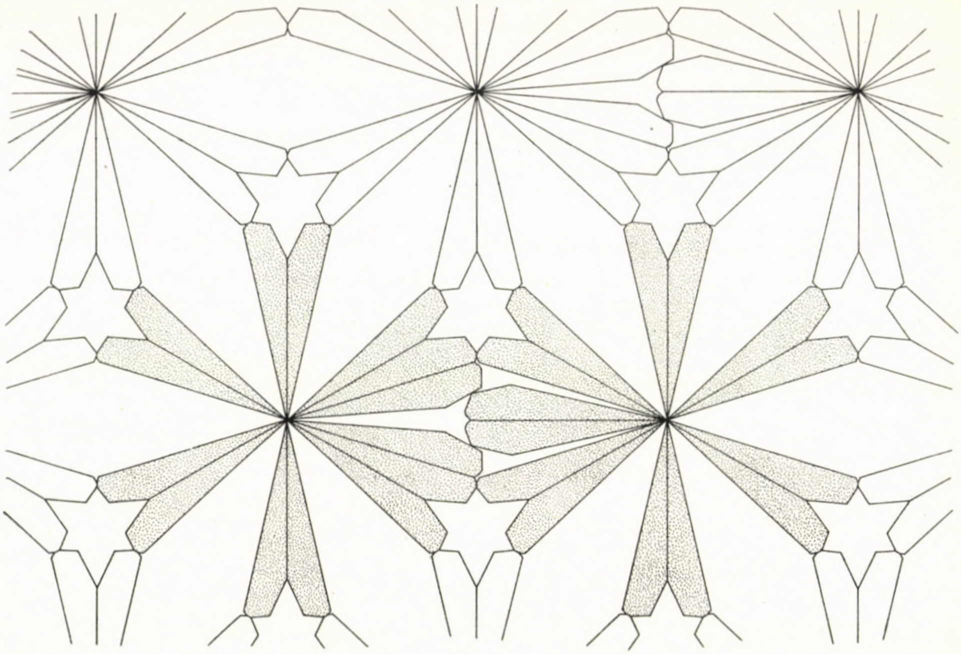


Fig. 5

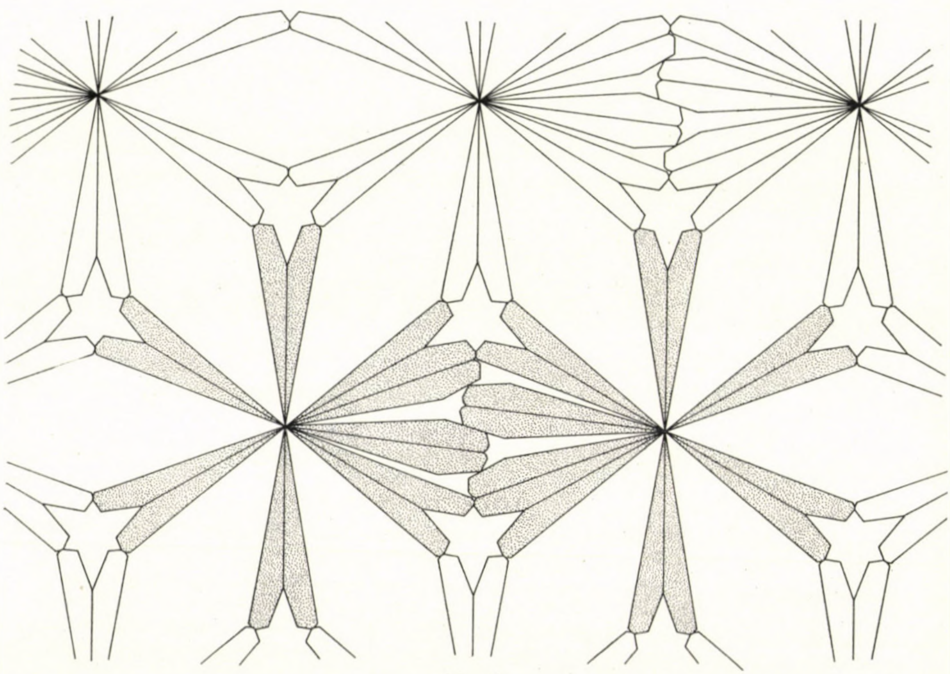


Fig. 6

den beiden Rosetten zugeordnet mit Ausnahme der beiden „äußeren' Sechsecke, welche ihre längsten Seiten mit Sechsecken der ursprünglichen 12-Nachbarnpackung gemeinsam haben und bei geradem n derselben, bei ungeradem n verschiedenen Rosetten angehören. Das derart erweiterte Rosettenpaar (in Fig. 4 bis 6 schattiert) ist Translationseinheit einer doppeltperiodischen n-Nachbarnpackung, denn jedes Mitglied hat $n-1$ Nachbarn in der Rosette, der es selbst angehört, und einen Nachbarn in einer anderen Rosette. Diese Packung ist bewegungsstabil: Nun sind die neu eingelagerten Mitglieder — nur deren Bewegungsstabilität ist noch nachzuweisen — zwar nicht untereinander äquivalent, jedoch läßt sich der Beweis einheitlich führen. Eines dieser Sechsecke S' denken wir uns wieder durch drei starr miteinander verbundene Punkte x, y, z repräsentiert, wobei x der Ecke a entspreche, y einem Berührpunkt dieses Sechseckes mit dem Mitglied der Nachbarrosette (also einem Punkt der Seite $\langle eh \rangle$) und z dem Fußpunkt des Lotes von y auf die der Seite $\langle aj \rangle$ entsprechende Seite von S'. Wieder fassen wir diese Punkte als Funktionen der Zeit auf und $t=0$ entspreche der durch die Packung definierten Lage. Für ein geeignetes Koordinatensystem ist

$$x(0)=(0,0), \quad y(0)=(1,0), \quad z(0)=(1,\lambda)$$

mit $0<\lambda<\frac{1}{\sqrt{3}}$. Wir setzen noch $u=(1,0)$, $v=(\sqrt{3},1)$ und $w=(-\sqrt{3},1)$; für jedes $n>12$ und jede Wahl von S' sind dann die durch die Nachbarn von S' gegebenen Bewegungsbeschränkungen der Punkte x und y stärker als die durch die folgenden Ungleichungen gegebenen Beschränkungen:

$$\begin{aligned} \dot{x}(0) &= \varrho u+\sigma v \quad \text{mit} \quad \varrho\geqq 0 \quad \sigma\geqq 0 \\ \dot{y}(0) &= \varrho' u+\sigma' v \quad \text{mit} \quad \sigma'\geqq 0 \end{aligned}$$

$\dot{z}(0)$ wird durch $\dot{x}(0)$ und $\dot{y}(0)$ bereits eindeutig bestimmt und es ist leicht nachzurechnen, daß $\dot{z}(0)$ dann in der positiven Hülle von u und w liegt:

$$\dot{z}(0) = \varrho'' u+\tau w \quad \text{mit} \quad \varrho''\geqq 0 \quad \text{und} \quad \tau\geqq 0,$$

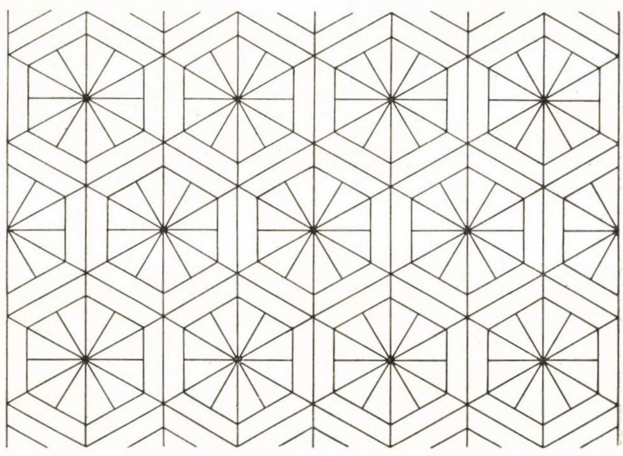


Fig. 7

wobei zudem $\varrho''=0$ nur dann eintreten kann, wenn $\varrho=\sigma=\varrho'=\sigma'=0$. Da nun S' im Punkte z derart in Berührung tritt mit einem Mitglied der Nachbarrosette, daß in keinem Falle $\dot{z}(0)$ eine Positivkombination von u und w sein kann mit $\varrho''>0$, folgt die Bewegungsstabilität der gesamten Packung.

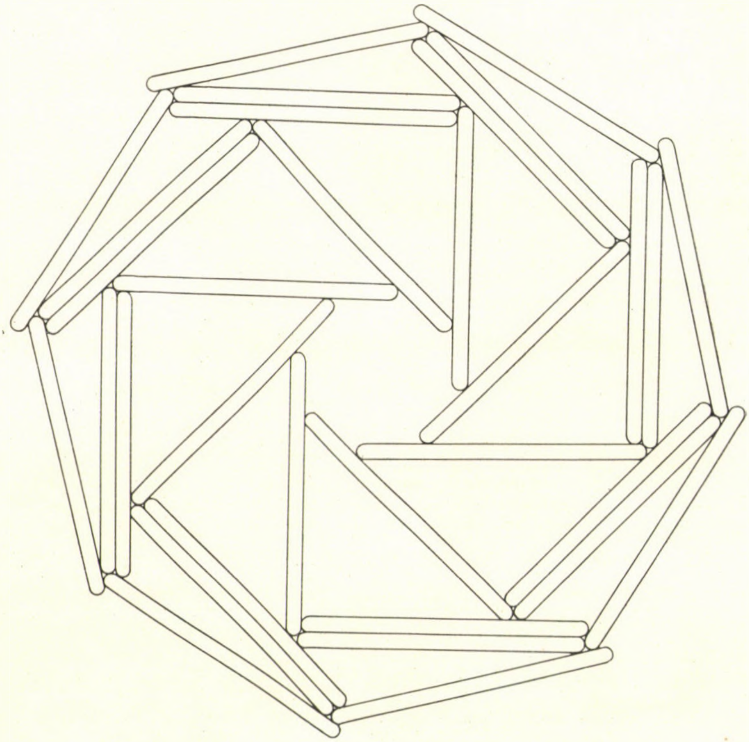


Fig. 8

3. Bemerkungen

3. 1 Aus einer $(12+k)$-Nachbarnpackung der eben konstruierten Art erhält man natürlich eine $(12+6k)$-Nachbarnpackung (im Falle der Figuren 4 bis 6 also eine 18-, 24- bzw. 30-Nachbarnpackung), wenn man die Lücken zwischen allen Rosettenpaaren in entsprechender Weise auffüllt. Passende teilweise Auffüllungen liefern $(12+mk)$-Nachbarnpackungen mit $m=2, 3, 4$ und 5. Man erhält so aus den Figuren 4 bis 6 schon n-Nachbarnpackungen für alle Zahlen von 13 bis 30.

3. 2 Wir haben eine Packung $\mathcal{G}$ bewegungsstabil genannt, wenn jedes Mitglied von $\mathcal{G}$ von seinen Nachbarn gegenüber Bewegungen fixiert wird. Es ist offenbar eine stärkere Forderung, wenn wir verlangen, daß für jede endliche Teilfamilie $\mathcal{G}'$ von $\mathcal{G}$ die Vereinigungsmenge $\bigcup_{S\in\mathcal{G}} S$ von der Restfamilie $\mathcal{G}\setminus\mathcal{G}'$ gegenüber Bewegungen fixiert wird. Mit Ausnahme der in Figur 1 dargestellten 11-Nachbarn-

packung haben alle hier wie auch in [1] betrachteten bewegungsstabilen n-Nachbarnpackungen diese stärkere Stabilitätseigenschaft. Dem Verfasser ist dagegen keine 11-Nachbarnpackung mit dieser Eigenschaft bekannt.

3. 3 In einem Aufsatz [2] fragt G. FEJES TÓTH nach der kleinsten ganzen Zahl $K=K(n)$ mit der Eigenschaft, daß sich aus kongruenten Exemplaren von K konvexen Polygonen eine n-Nachbarnparkettierung konstruieren lässt. Er weist darauf hin, daß $K(6)=K(7)=K(8)=K(9)=K(10)=K(12)=K(14)=K(16)=K(21)=1$ gilt und zeigt an einem Beispiel, daß $K(11)\leqq 2$ ist. Die in Fig. 7 dargestellte Parkettierung zeigt, daß auch $K(13)\leqq 2$ ausfällt.

3. 4 In [1] wird eine 4-Nachbarnpackung aus 12 kongruenten glatten Scheiben konstruiert und die Vermutung ausgesprochen, daß eine 5-Nachbarnpackung aus endlich vielen kongruenten glatten Scheiben nicht existiert. Fig. 8 zeigt eine 5-Nachbarnpackung aus 32 kongruenten "Stangen", die die obige Vermutung widerlegt.

LITERATURVERZEICHNIS

[1] FEJES TÓTH, L.: Scheibenpackungen konstanter Nachbarnzahl. *Acta Math. Acad. Sci. Hung.* **20** (1969), 375—381.
[2] FEJES TÓTH, G.: Über Parkettierungen konstanter Nachbarnzahl. *Studia Sci. Math. Hung.* **6** (1971) 133—135.

Math. Institut der Univ. Dortmund

(Eingegangen: 11. September 1970)

PUNKTVERTEILUNGEN IN EINEM QUADRAT

Von

L. FEJES TÓTH

In einem rechtwinkligen Koordinatensystem xy definieren wir den „*Abstand*" von zwei Punkten $P_1=(x_1, y_1)$ und $P_2=(x_2, y_2)$ durch die Größe

$$a(P_1, P_2) = |x_1-x_2|+|y_1-y_2|.$$

Dies ist der natürlichste Abstandsbegriff in einer Stadt, in der die Straßen zu den Koordinatenachsen parallel verlaufen. Ein wenig anders und allgemeiner formuliert legen wir der Abstandsmeßung zwei oder mehr Scharen paralleler Geraden zugrunde. Wir nennen diese Geraden „Straßen" und definieren den *Abstand* von zwei Punkten als die Länge eines kürzesten Weges, der die Punkte entlang irgendwelcher Straßen verbindet.

Das Hauptproblem, mit dem wir uns beschäftigen wollen, lautet folgendermaßen: Wie soll man in einem Quadrat Q, in dem die Straßen zu den Seiten parallel sind, n Punkte so verteilen, daß der Mindestabstand zwischen je zwei Punkten maximal wird?

Für $n \leqq 5$ sind die Lösungen trivial bzw. einfach (Abb. 1). Für $n=2$ haben wir zwei diametrale Ecken von Q. Bei $n=3$ Punkten liegt ein Punkt in einer Ecke von Q, während die beiden übrigen Punkte auf dem Rand von Q so verteilt sind, daß die drei Punkte den Rand von Q in drei gleich lange Teile zerlegen. Die Verteilung von $n=4$ Punkten ist nicht eindeutig. Eine Lösung liefern uns 4 Punkte auf dem Rand von Q, die den Rand in 4 gleich lange Bogen zerlegen. Wir können aber auch 3 Punkte in je eine Ecke von Q und den vierten Punkt auf diejenige Strecke legen, die den Mittelpunkt von Q mit der vierten Ecke verbindet. Die beste Verteilung von $n=5$ Punkten ist durch die Ecken und den Mittelpunkt von Q gegeben.

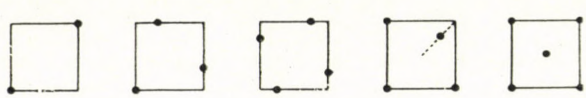


Abb. 1

Im folgenden lösen wir das Problem in demjenigen Fall, daß n die Summe von zwei aufeinanderfolgenden Quadratzahlen ist: $n = k^2+(k+1)^2$, $k=1, 2, \ldots$. Die extremale Verteilung von $k^2+(k+1)^2$ Punkten im Quadrat $0 \leqq x$, $y \leqq 2k$ ist durch die Punkte mit ganzzahligen Koordinaten mit einer geraden Summe gegeben. Abb. 2 stellt den Fall $k=3$ dar. Für beliebige Werte von n scheint das Problem recht schwerig zu sein.

Unser Ergebnis ist enthalten im folgenden

SATZ. *Von n Punkten $P_1, \ldots, P_n$ eines Einheitsquadrates, in dem die Straßen zu den Seiten parallel liegen, lassen sich stets zwei Punkte P_i, P_j $(i \neq j)$ mit einem Abstand*

$$a(P_i, P_j) \leqq \frac{1+\sqrt{2n-1}}{n-1}$$

herausgreifen. Das Gleichheitszeichen wird nur im Falle einer Punktzahl der Form $n = k^2+(k+1)^2$ $(k=1, 2, \ldots)$ beansprucht.

BEWEIS. Das Einheitsquadrat Q, in dem die Punkte $P_1, \ldots, P_n$ liegen, sei das Quadrat $0 \leqq x, y \leqq 1$. Ferner sei a der Mindestabstand zwischen je zwei Punkten, und q_i die Menge derjenigen Punkte, deren Abstand von P_i kleiner ist als $a/2$. q_i ist ein offenes Quadrat mit dem Mittelpunkt P_i und mit Diagonalen der Länge a, die zu den Koordinatenachsen parallel sind. Da die Mittelpunkte dieser Quadrate voneinander einen Abstand $\geqq a$ haben, haben zwei verschiedene Quadrate keine gemeinsamen Punkte. Deshalb ist die Inhaltssumme $n \frac{a^2}{2}$ der Quadrate $q_1, \ldots, q_n$ höchstens der Inhaltssumme von Q und der von Q herausragenden Teile von $q_1, \ldots, q_n$ gleich. Um den Gesamtinhalt der von Q herausragenden Teile abzuschätzen zerlegen wir diejenigen Quadrate, die eine Ecke von Q enthalten, durch die Diagonalen in je vier Dreiecke, und heben diejenigen Dreiecke, die eine Ecke von Q enthalten, heraus. Der Gesamtinhalt dieser Dreiecke ist höchstens $a^2/2$. Nun betrachten wir etwa die im Halbstreifen $1 \leqq x$, $0 \leqq y \leqq 1$ liegenden Teile der Quadrate $q_1, \ldots, q_n$, bzw. der gebliebenen Quadratteile. Diese bestehen aus nicht überlappenden Dreiecken, deren Basen auf der Seite $x=1$, $0 \leqq y \leqq 1$ von Q liegen, und deren Höhen höchstens $a/2$ sind. Deshalb beträgt die Inhaltssumme dieser Teile höchstens $a/4$. Folglich ist der Gesamtinhalt der ausserhalb von Q

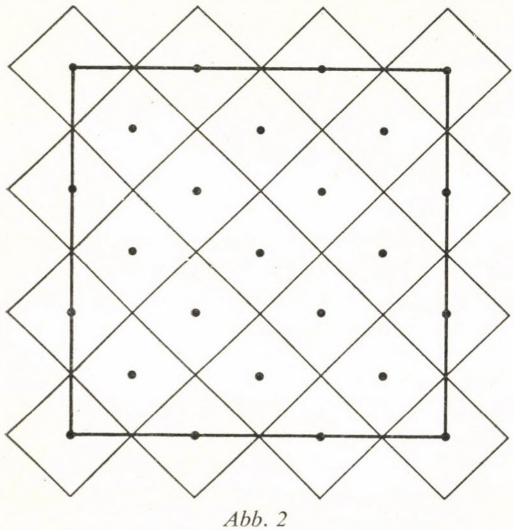


Abb. 2

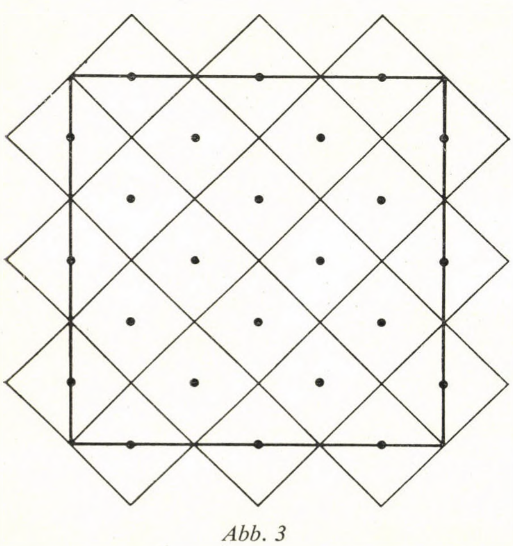


Abb. 3

liegenden Teile von $q_1, \ldots, q_n$ höchstens $a+\frac{a^2}{2}$. Wir haben also

$$n\frac{a^2}{2} \leqq 1+a+\frac{a^2}{2},$$

was mit der zu beweisenden Ungleichung äquivalent ist. Der Fall der Gleichheit leuchtet ein.

Wir geben noch einige günstige Punktanordnungen an. Die $2k^2$ Punkte mit ganzzahligen Koordinaten mit einer geraden Summe, die im Quadrat $0 \leqq x$, $y \leqq 2k-1$ liegen ($k=1, 2, \ldots$), scheinen in diesem Quadrat ebenfalls extremal verteilt zu sein. Bezeichnen wir mit a_n den maximalen Mindestabstand zwischen n Punkten eines Einheitsquadrats, so zeigt diese Punktanordnung zusammen mit dem obigen Satz, daß

$$\frac{2}{2k-1} \leqq a_{2k^2} < \frac{1+\sqrt{4k^2-1}}{2k^2-1}.$$

Z. B. gilt $0{,}4 \leqq a_{18} < 0{,}406$. Weiterhin ist die Verteilung von $2k(k+1)$ Punkten, die aus der extremalen Anordnung von $k^2+(k+1)^2$ Punkten durch Fortlaßung eines Punktes entsteht, vermutlich noch immer extremal. Es gibt aber auch eine ebenso gute Verteilung von anders angeordneten $2k(k+1)$ Punkten, nämlich die Punkte mit ganzzahligen Koordinaten mit einer ungeraden Summe, die im Quadrat $0 \leqq x, y \leqq 2k$ liegen (Abb. 3).

Zum Schluss erwähnen wir noch einige weitere Probleme.

Wir wollen im Quadrat Q n Punkte, die wir „Schulen" nennen, so verteilen, daß der Maximalabstand zwischen einem Punkt von Q und der nächsten Schule minimal wird. Es handelt sich hier um die sparsamste Überdeckung von Q durch n kleinere, um 45° verdrehte Quadrate. Vermutlich ist die zuletzt betrachtete Verteilung von $2k(k+1)$ Punkten im Quadrat $0 \leqq x$, $y \leqq 2k$ auch bezüglich dieses Problems extremal. Die Schwierigkeit im Beweis dieser Vermutung wird durch die in Abb. 4 dargestellte, sehr günstige Verteilung von 17 Schulen illustriert, wo von den kleinen Quadraten ein verhältnismäßig noch kleinerer Teil von Q herausragt, als bei der betrachteten Verteilung von $2k(k+1)$ Punkten. Zerlegt man ein Quadrat in vier Teilquadrate und konstruiert in jedem Teilquadrat die betrachtete Anordnung von 17 Schulen, so fallen 16 Schulen paarweise zusammen, so daß in dem großen

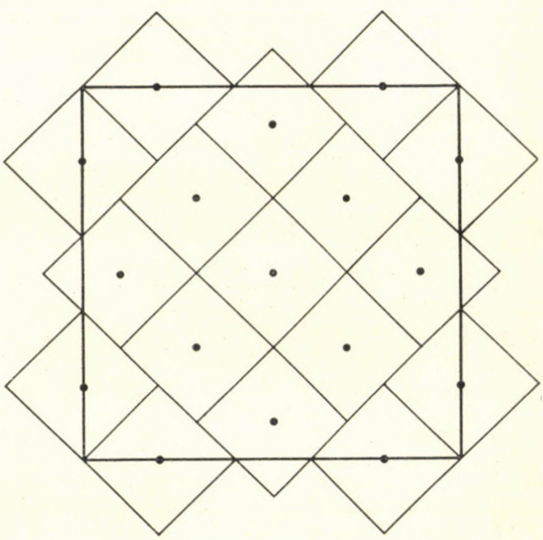


Abb. 4

Quadrat eine Verteilung von $4 \cdot 17 - 8 = 60$ Schulen entsteht. Diese Verteilung ist genau so gut, wie die obige Verteilung von $2k(k+1)$ Schulen im Falle von $k=5$.

Die obigen Probleme lassen sich dem Bereich, in dem die Punkte liegen, und den Straßenrichtungen entsprechend in mannigfaltiger Weise variieren. Betrachten wir z. B. ein gleichseitiges Dreieck, in dem die Straßen in den Höhenrichtungen verlaufen. Der Abstand zwischen zwei Ecken sei 1. Dann lassen sich von n Punkten des Dreiecks stets zwei Punkte mit einem Abstand

$$a \leqq \frac{3+\sqrt{8n+1}}{4n-4}$$

herausgreifen, wobei das Gleichheitszeichen nur im Falle $n = \frac{1}{2}k(k+1)$ beansprucht wird ($k=2, 3, \ldots$). Dies läßt sich in ähnlicher Weise zeigen wie der obige Satz. Abb. 5 zeigt die extremale Verteilung von $\frac{1}{2}k(k+1)$ Punkten im Falle $k=4$.

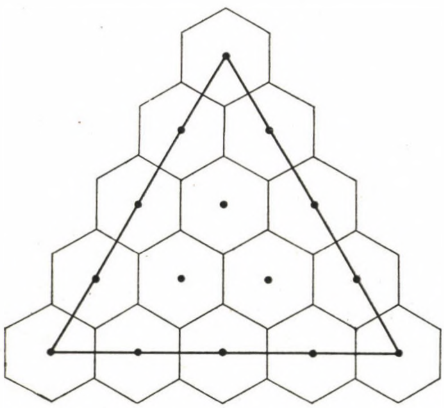


Abb. 5

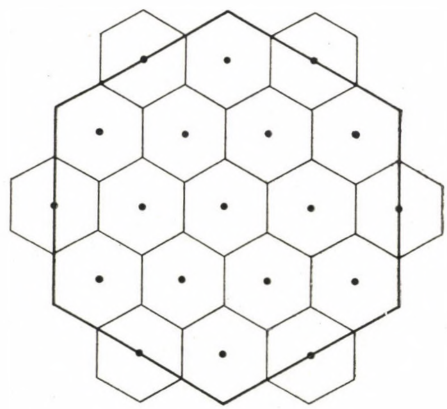


Abb. 6

Abb. 6 stellt eine sehr günstige Verteilung von 19 Schulen in einem regulären Sechseck dar, in dem die Straßen zu den Seiten parallel verlaufen. Ähnliche günstige Anordnungen existieren auch, wenn die Zahl der Schulen $1+9k(k+1)$ ist.

Mathematisches Institut der Ungarischen Akademie der Wissenschaften, Budapest

(Eingegangen: 16. September 1970.)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 443—452.

ON LIMITING DISTRIBUTIONS FOR THE SUMS OF RANDOM NUMBER OF RANDOM VARIABLES CONCERNING THE RAREFACTION OF RECURRENT PROCESS

by

T. SZÁNTAI

1. Introduction

In [8] A. RÉNYI has solved the following problem: let us consider a recurrent process for which $a = \int_0^{+\infty} x\, dF(x) < +\infty$, where $F(x)$ is the distribution function of the time interval between consecutive renewal points. We rarefy the process such that, independently from one another, every event of it will be maintained with probability q and cancelled with probability $1-q$, where $0<q<1$. In the so obtained process let us make a coordinate transformation such that the expectation of the time interval between two consecutive renewal points be again a. In this way we get a recurrent process with the same intensity as that of the original one. Moreover, the Laplace—Stieltjes transform of the distribution function of the time interval between consecutive renewal points is

$$\psi(s) = \frac{q\varphi(qs)}{1-(1-q)\varphi(qs)}, \tag{1.1}$$

where $\varphi(s) = \int_0^{+\infty} e^{-sx}\, dF(x)$. The question is what kind of process we shall get if we use the above defined operation succesively several times for a recurrent process. A. RÉNYI ([8]) has proved that the limit process is a homogeneous Poisson process possessing the same intensity as that of the starting process.

The same problem was investigated by I. N. KOVALENKO ([6]) slightly more generally. He did not leave out the infinite expectation recurrent process in his investigations. If ε denotes the probability of maintaining an event of the original recurrent process and if a coordinate transformation, $t' = \delta \cdot t$, is made, then the Laplace—Stieltjes transform of the distribution function of the time interval be- (ween two consecutive renewal points of the new process will be as follows:

$$\varphi_\varepsilon(s) = \frac{\varepsilon\varphi(\delta s)}{1-(1-\varepsilon)\varphi(\delta s)}. \tag{1.2}$$

Let $\varepsilon = \varepsilon(\delta) \to 0$ as $\delta \to 0$ and assume that

$$\varphi_0(s) = \lim_{\delta\to 0} \frac{\varepsilon(\delta)\varphi(\delta s)}{1-(1-\varepsilon(\delta))\varphi(\delta s)} \tag{1.3}$$

exists.

In [6] I. N. KOVALENKO has proved that $\varphi_0(s)$ could be written only in two forms:

$$\varphi_0(s) = \frac{1}{1+cs^\beta}, \qquad \operatorname{Re} s \geqq 0, \ c>0, \ 0<\beta\leqq 1, \tag{1.4}$$

or

$$\varphi_0(s) \equiv 1, \tag{1.5}$$

where β is defined by the relation

$$\lim_{\delta\to 0} \frac{\varepsilon(\delta z)}{\varepsilon(\delta)} = z^\beta.$$

In order that $\varphi_0(s)$ is of the form (1. 4) I. N. KOVALENKO gave a necessary and sufficient condition, concerning $\varphi(s)$. In [3] B. V. GNEDENKO has corrected a statement of KOVALENKO; more precisely, he has shown that there are recurrent processes with infinite mean-value such that in the limiting case we obtain a homogeneous Poisson process. In addition B. V. GNEDENKO has given the exact domain of attraction of the possible limiting distributions (1. 4).

The above results can be illustrated as limiting distribution theorems for the sums of a random number of independent random variables. Indeed, denoting the time intervals between consecutive renewal points of the starting process by $\xi_1, \xi_2, \dots, \xi_n, \dots$, then they are independent and commonly distributed random variables. After the first rarefaction the time distance between the consecutive renewal points of the new process will be

$$\zeta_{\nu_1} = \xi_1+\xi_2+\dots+\xi_{\nu_1}, \tag{1.6}$$

where ν_1-1 is the number of the cancelled events. The random variable ν_1 is independent of every ξ_n and its distribution is

$$\mathsf{P}(\nu_1=k) = q(1-q)^{k-1}, \qquad k=1,2,\dots$$

The time distance between consecutive events after the n-th rarefaction will be

$$\zeta_{\nu_n} = \xi_1+\xi_2+\dots+\xi_{\nu_n},$$

where ν_n is also independent of the summands and its distribution is

$$\mathsf{P}(\nu_n=k) = q^n(1-q^n)^{k-1}, \qquad k=1,2,\dots$$

Now RÉNYI's limiting distribution theorem can be formulated as follows: when the common mean-value of the random variables is a finite number a then the distribution function of the suitable normed sums ζ_{ν_n} converges to the exponential distribution, or more exactly

$$\mathsf{P}(q^n\zeta_{\nu_n}<x) \to 1-e^{-\frac{x}{a}}, \qquad n\to\infty. \tag{1.7}$$

The results due to I. N. KOVALENKO [6] and B. V. GNEDENKO [3] give, if $0<q<1$, all possible limiting distributions for the type (1. 6), and the domains of attraction of limiting distributions in any cases.

It is easy to see that for arbitrary real, independent and commonly distributed random variables with positive mean-value the preceding results are valid and it can be seen, by a simple transformation, that if the mean-value is negative we obtain also a limiting distribution. If the expectation is zero then the situation is different from the previous ones; in section 2. this problem will be solved.

On the basis of KONDRATAI TI GENE's university dissertation (see in [5])* it is known that if ξ_i, $i=1, 2, \ldots$ have finite variance then the rate of convergence of (1. 7) is of order $cq^n \ln \frac{1}{q^n}$, where the constant c is independent of n. In section 3. an estimation of the rate of convergence in the limit distribution theorem from section 2. will be given.

Finally, in section 4. the same questions will be investigated for the rarefaction method introduced by J. MOGYORÓDI in [7].

2. The limit distribution theorem

THEOREM 1. *Let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ be a sequence of independent, commonly distributed random variables, for which $\int_{-\infty}^{+\infty} x\,dF(x)=0$ and $\int_{-\infty}^{+\infty} x^2\,dF(x) = \frac{2}{\lambda^2} < +\infty$ are satisfied, where $\mathsf{P}(\xi_n<x)=F(x)$ is the common distribution function.*

Moreover let ν_n be a sequence of geometrically distributed random variables with parameter $q^n (0<q<1,\ n=1, 2, \ldots)$, and suppose that ν_n is independent of the sequence $\xi_1, \xi_2, \ldots, \xi_n, \ldots$.

If $\zeta_{\nu_n} = \xi_1+\xi_2+\cdots+\xi_{\nu_n}$ then

$$\lim_{n\to\infty} \mathsf{P}\left(\sqrt{q^n}\,\zeta_{\nu_n}<x\right) = \int_{-\infty}^{x} \frac{\lambda}{2} e^{-\lambda|y|}\,dy, \qquad \lambda>0. \tag{2. 1}$$

PROOF. It is easy to see for the characteristic function of the distribution of $\sqrt{q^n}\zeta_{\nu_n}$:

$$\sum_{k=1}^{\infty} q^n(1-q^n)^{k-1}\varphi^k\left(\sqrt{q^n}\,t\right) = \frac{q^n\varphi\left(\sqrt{q^n}\,t\right)}{1-(1-q^n)\varphi\left(\sqrt{q^n}\,t\right)} = \frac{\varphi\left(\sqrt{q^n}\,t\right)}{\dfrac{1-\varphi\left(\sqrt{q^n}\,t\right)}{q^n}+\varphi\left(\sqrt{q^n}\,t\right)},$$

where $\varphi(t) = \int_{-\infty}^{+\infty} e^{itx}\,dF(x)$.

Using the rule of L'Hospital

$$\lim_{n\to\infty}\frac{1-\varphi\left(\sqrt{q^n}\,t\right)}{q^n} = \lim_{n\to\infty}\frac{1-\varphi\left(\sqrt{q^n}\,t\right)}{q^n t^2}t^2 = \lim_{n\to\infty}\frac{-\varphi'\left(\sqrt{q^n}\,t\right)}{2\sqrt{q^n}\,t}t^2 =$$

$$= -\frac{t^2}{2}\lim_{n\to\infty}\frac{\varphi'\left(\sqrt{q^n}\,t\right)-\varphi'(0)}{\sqrt{q^n}\,t-0} = -\frac{t^2}{2}\varphi''(0) = \left(-\frac{t^2}{2}\right)\left(-\frac{2}{\lambda^2}\right) = \frac{t^2}{\lambda^2}.$$

* The result can be also found in the author's university doctoral dissertation, Eötvös Loránd University, Budapest.

Since $\lim_{t\to 0}\varphi(t)=1$, we obtain for the characteristic function of the distribution of $\sqrt{q^n}\zeta_{\nu_n}$ as $n\to+\infty$

$$\lim_{n\to\infty}\frac{\varphi(\sqrt{q^n}t)}{\frac{1-\varphi(\sqrt{q^n}t)}{q^n}+\varphi(\sqrt{q^n}t)}=\frac{1}{\frac{t^2}{\lambda^2}+1}=\frac{\lambda^2}{\lambda^2+t^2}.$$

This result proves the statement of the theorem by the well-known continuity theorem for characteristic functions; indeed, $\frac{\lambda^2}{\lambda^2+t^2}$ is exactly the characteristic function of socalled two sided exponential distribution for which the distribution function is $\int_{-\infty}^{x}\frac{\lambda}{2}e^{-\lambda|y|}\,dy$, $\lambda>0$. The proof is completed.

Following the methods of proof of DOBRUSHINE's lemma ([1]) a new proof can be given for the theorem. It is worth showing this because the limit distribution theorem from section 3. can be proved only by this method.

The second proof of theorem.

First of all we remark that by the simplest form of the central limit theorem

$$\lim_{m\to\infty}\mathsf{P}\left(\frac{\zeta_m}{\sqrt{m}\frac{\sqrt{2}}{\lambda}}<x\right)=\Phi(x), \tag{2.2}$$

where $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{u^2}{2}}\,du$.

On the other hand from the relation

$$\mathsf{P}(\nu_n q^n<x)=\sum_{k=1}^{\left[\frac{x}{q^n}\right]}q^n(1-q^n)^{k-1}=1-(1-q^n)^{\left[\frac{x}{q^n}\right]}$$

it follows immediately that

$$\lim_{n\to\infty}\mathsf{P}(\nu_n q^n<x)=E(x), \tag{2.3}$$

where $E(x)=1-e^{-x}$, if $x>0$ and $E(x)=0$ if $x\leq 0$.

The proof is based on the simple fact that if a sequence of distribution functions $K_n(x)$, $n=1,2,\dots$, is weakly convergent to $K(x)$ then a sequence χ_n of random variables can be constructed such that $\mathsf{P}(\chi_n<x)=K_n(x)$ and χ_n converges in probability to a random variable χ for which $\mathsf{P}(\chi<x)=K(x)$. To prove this it is sufficient to consider the interval [0, 1] as the space of the elementary events and to put $\chi_n(y)=K_n^{-1}(y)$ where $K_n^{-1}(y)$ is the inverse of the function $K_n(x)$.

Now let the space of the elementary events be the unit square of the plane (u, v). In the interval [0, 1] of axis u let a sequence φ_m of random variables $(m=1,2,\dots)$ be constructed for the elements of which the followings are satisfied:

a) their distributions are equal to the distributions of $\dfrac{\zeta_m}{\sqrt{m}\dfrac{\sqrt{2}}{\lambda}}$, $m=1, 2, \ldots$;

b) they converge in probability to a random variable $\bar{\zeta}$ with distribution function $\Phi(x)$ as $m \to +\infty$.

Such a construction is possible because of (2. 2) and the preceding remark. Let further ψ_n be a sequence of random variables, $n=1, 2, \ldots$, constructed on the interval $[0, 1]$ of the axis v for the elements of which the followings are satisfied:

a) their distributions are equal to the distributions of $v_n q^n$, $n=1, 2, \ldots$;

b) they converge in probability, as $n \to +\infty$, to a random variable $\bar{v}$ having the distribution function $E(x)$, and which is independent of $\bar{\zeta}$.

Let these random variables be extended to the whole unit square by the following definition: $\varphi_m(u, v)=\varphi_m(u, 0)$, $\psi_n(u, v)=\psi_n(0, v)$, $\bar{\zeta}(u, v)=\bar{\zeta}(u, 0)$ and $\bar{v}(u, v)=\bar{v}(0, v)$.

Now it is easy to see that the random variable

$$\tilde{\zeta}_m = \sqrt{m}\,\frac{\sqrt{2}}{\lambda}\,\varphi_m \tag{2. 4}$$

is of the same distribution as that of ζ_m, $m=1, 2, \ldots$, and the random variable

$$\tilde{v}_n = \frac{1}{q^n}\,\psi_n \tag{2. 5}$$

has a distribution which is the same as that of v_n, $n=1, 2, \ldots$. Moreover the random variables $\tilde{\zeta}_m$, $m=1, 2, \ldots$ and $\tilde{v}_n$, $n=1, 2, \ldots$ are mutually independent. Hence the random variables $\tilde{\zeta}_{\tilde{v}_n}$ and ζ_{v_n} have the same distributions, $n=1, 2, \ldots$, and thus the investigation of ζ_{v_n}, $n=1, 2, \ldots$ can be replaced by the investigation of $\tilde{\zeta}_{\tilde{v}_n}$, $n=1, 2, \ldots$.

The definitions (2. 4) and (2. 5) of random variables $\tilde{\zeta}_m$, $m=1, 2, \ldots$ and $\tilde{v}_n$, $n=1, 2, \ldots$ immediately imply

$$\tilde{\zeta}_m = \sqrt{m}\,\frac{\sqrt{2}}{\lambda}\,\bar{\zeta}+o(\sqrt{m}) \quad \text{as} \quad m \to +\infty \tag{2. 6}$$

and

$$\tilde{v}_n = \frac{1}{q^n}\,\bar{v}+o\left(\frac{1}{q^n}\right) \quad \text{as} \quad n \to +\infty \tag{2. 7}$$

where $o(\cdot)$ is meant in the sense of convergence in probability. Hence it follows that

$$\tilde{\zeta}_{\tilde{v}_n} = \sqrt{\frac{1}{q^n}\,\bar{v}+o\left(\frac{1}{q^n}\right)}\,\frac{\sqrt{2}}{\lambda}\,\bar{\zeta}+o\left(\sqrt{\frac{1}{q^n}\,\bar{v}+o\left(\frac{1}{q^n}\right)}\right) \quad \text{as} \quad n \to +\infty,$$

i.e.

$$\sqrt{q^n}\,\tilde{\zeta}_{\tilde{v}_n} = \frac{\sqrt{2}}{\lambda}\sqrt{\bar{v}+o(1)}\,\bar{\zeta}+o\left(\sqrt{\bar{v}+o(1)}\right) \quad \text{as} \quad n \to +\infty. \tag{2. 8}$$

So for $n\to+\infty$ the sequence of random variables $\sqrt{q^n}\tilde{\zeta}_{\tilde{v}_n}$, $n=1,2,\dots$ converges in probability to the random variable $\frac{\sqrt{2}}{\lambda}\sqrt{\tilde{v}}\,\tilde{\zeta}$. Since, the convergence in probability implies the convergence of the distribution functions, it follows that

$$\lim_{n\to\infty}\mathsf{P}\left(\sqrt{q^n}\,\tilde{\zeta}_{\tilde{v}_n}<x\right)=\mathsf{P}\left(\frac{\sqrt{2}}{\lambda}\sqrt{\tilde{v}}\,\tilde{\zeta}<x\right)$$

and thus the same is true for the random variable ζ_{v_n}:

$$\lim_{n\to\infty}\mathsf{P}\left(\sqrt{q^n}\,\xi_{v_n}<x\right)=\mathsf{P}\left(\frac{\sqrt{2}}{\lambda}\sqrt{\bar{v}}\,\zeta<x\right)= \tag{2.9}$$

$$=\iint\limits_{\frac{\sqrt{2}}{\lambda}\sqrt{z}\,y<x} d\Phi(y)\,dE(z)=\int_0^{+\infty}\int_{-\infty}^{\frac{\lambda x}{\sqrt{2z}}} d\Phi(y)\,dE(z)=\int_0^{+\infty}\Phi\left(\frac{\lambda x}{\sqrt{2z}}\right)e^{-z}\,dz.$$

Finally, we prove the equality

$$\int_0^{+\infty}\Phi\left(\frac{\lambda x}{\sqrt{2z}}\right)e^{-z}\,dz=\int_{-\infty}^{x}\frac{\lambda}{2}\,{}^{-\lambda|y|}\,dy. \tag{2.10}$$

It is enough to see instead of the equality (2. 10) the equality of the corresponding density functions. But this can be easily seen by the relations

$$\int_0^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{\lambda^2x^2}{4z}}\frac{\lambda}{\sqrt{2z}}e^{-z}\,dz=\frac{\lambda}{\sqrt{\pi}}\int_0^{+\infty}e^{-\frac{\lambda^2x^2}{4u^2}}e^{-u^2}\,du=\frac{\lambda}{\sqrt{\pi}}\frac{\sqrt{\pi}}{2}e^{-2\frac{\lambda|x|}{2}}=\frac{\lambda}{2}e^{-\lambda|x|}.$$

For the calculation of the integral $\int_0^{+\infty}e^{-\frac{\lambda^2x^2}{4u^2}}e^{-u^2}\,du$ the formula

$$\int_0^{+\infty}e^{-\frac{b^2}{u^2}}e^{-a^2u^2}\,du=\frac{\sqrt{\pi}}{2a}e^{-2ab},\qquad a>0,\ b>0$$

was used (see in [2] 860. 25.).

It is similarly simple to show the equality of their characteristic functions:

$$\int_{-\infty}^{+\infty}e^{itx}\,d_x\left\{\int_0^{+\infty}\Phi\left(\frac{\lambda x}{\sqrt{2z}}\right)dE(z)\right\}=\int_0^{+\infty}\int_{-\infty}^{+\infty}e^{itx}\,d_x\Phi\left(\frac{\lambda x}{\sqrt{2z}}\right)dE(z)=$$

$$=\int_0^{+\infty}e^{-\frac{t^2\,2z}{2\lambda^2}}\,dE(z)=\int_0^{+\infty}e^{-\frac{t^2z}{\lambda^2}}e^{-z}\,dz=\int_0^{+\infty}e^{-z\left(\frac{t^2}{\lambda^2}+1\right)}\,dz=\frac{\lambda^2}{\lambda^2+t^2}.$$

The proof is completed.

3. The rate of convergence

THEOREM 2. *If* $\int_{-\infty}^{+\infty} |x|^3 dF(x) < +\infty$ *is satisfied beside the conditions of Theorem 1. then there exists a constant c which is independent of n such that*

$$\left| \mathsf{P}(\sqrt{q^n}\,\zeta_{\nu_n} < x) - \int_{-\infty}^{x} \frac{\lambda}{2} e^{-\lambda|y|}\,dy \right| \leqq cM\left(\frac{1}{\sqrt{\nu_n}}\right) \tag{3.1}$$

holds if n is suitable large.

PROOF. According to the theorem of total probability

$$\mathsf{P}(\sqrt{q^n}\,\zeta_{\nu_n} < x) = \sum_{k=1}^{\infty} \mathsf{P}\left(\frac{\zeta_k}{\sqrt{k}\,\frac{\sqrt{2}}{\lambda}} < \frac{x}{\sqrt{k}\,\frac{\sqrt{2}}{\lambda}\sqrt{q^n}} \right) \mathsf{P}(\nu_n = k). \tag{3.2}$$

By the condition $\int_{-\infty}^{+\infty} |x|^3 dF(x) < +\infty$ it follows from the BERRY—ESSEEN theorem (see for example in [4]) that

$$\left| \mathsf{P}\left(\frac{\zeta_k}{\sqrt{k}\,\frac{\sqrt{2}}{\lambda}} < \frac{x}{\sqrt{k}\,\frac{\sqrt{2}}{\lambda}\sqrt{q^n}} \right) - \Phi\left(\frac{x}{\sqrt{k}\,\frac{\sqrt{2}}{\lambda}\sqrt{q^n}} \right) \right| < c_1 \frac{1}{\sqrt{k}}, \tag{3.3}$$

where c_1 is a constant being independent of k.

By (3. 2) and (3. 3) for the probability $\mathsf{P}(\sqrt{q^n}\cdot\zeta_{\nu_n} < x)$ the inequalities

$$\sum_{k=1}^{\infty} \Phi\left(\frac{x}{\sqrt{k}\,\frac{\sqrt{2}}{\lambda}\sqrt{q^n}} \right) \mathsf{P}(\nu_n = k) - c_1 \sum_{k=1}^{\infty} \frac{1}{\sqrt{k}} \mathsf{P}(\nu_n = k) \leqq \tag{3.4}$$

$$\leqq \mathsf{P}(\sqrt{q^n}\,\zeta_{\nu_n} < x) \leqq$$

$$\leqq \sum_{k=1}^{\infty} \Phi\left(\frac{x}{\sqrt{k}\,\frac{\sqrt{2}}{\lambda}\sqrt{q^n}} \right) \mathsf{P}(\nu_n = k) + c_1 \sum_{k=1}^{\infty} \frac{1}{\sqrt{k}} \mathsf{P}(\nu_n = k)$$

hold.

Let $E_n(x) = \mathsf{P}(\nu_n q^n < x)$. Then using this notation

$$\sum_{k=1}^{\infty} \Phi\left(\frac{x}{\sqrt{k}\,\frac{\sqrt{2}}{\lambda}\sqrt{q^n}} \right) \mathsf{P}(\nu_n = k) = \int_{0}^{+\infty} \Phi\left(\frac{x}{\frac{\sqrt{2}}{\lambda}\sqrt{z}} \right) dE_n(z). \tag{3.5}$$

Moreover, if $E(x) = 1-e^{-x}$ as $x>0$ and $E(x)=0$ as $x\leqq 0$ then by (2. 10) the equation

$$\int_0^{+\infty} \Phi\left(\frac{\lambda x}{\sqrt{2z}}\right) dE(z) = \int_{-\infty}^{x} \frac{\lambda}{2} e^{-\lambda|y|}\, dy \tag{3. 6}$$

is true.

From relations (3. 5) and (3. 6):

$$\left|\sum_{k=1}^{\infty} \Phi\left(\frac{x}{\sqrt{k}\frac{\sqrt{2}}{\lambda}\sqrt{q^n}}\right) \mathsf{P}(v_n=k) - \int_{-\infty}^{x} \frac{\lambda}{2} e^{-\lambda|y|}\, dy\right| = \tag{3. 7}$$

$$= \left|\int_0^{+\infty} \Phi\left(\frac{x}{\frac{\sqrt{2}}{\lambda}\sqrt{z}}\right) dE_n(z) - \int_0^{+\infty} \Phi\left(\frac{\lambda x}{\sqrt{2z}}\right) dE(z)\right| =$$

$$= \left|\left[\Phi\left(\frac{\lambda x}{\sqrt{2z}}\right) E_n(z)\right]_{z=0}^{\infty} - \int_0^{+\infty} E_n(z)\, d_z \Phi\left(\frac{\lambda x}{\sqrt{2z}}\right) - \left[\Phi\left(\frac{\lambda x}{\sqrt{2z}}\right) E(z)\right]_{z=0}^{\infty} +\right.$$

$$\left. + \int_0^{+\infty} E(z)\, d_z \Phi\left(\frac{\lambda x}{\sqrt{2z}}\right)\right| \leqq \int_0^{+\infty} |E(z)-E_n(z)|\, d_z \Phi\left(\frac{\lambda x}{\sqrt{2z}}\right).$$

Now $\lim\limits_{n\to\infty} E_n(z) = E(z)$, and concerning the rate of convergence the following is true: there exists a constant c_2 which is independent of n and for which

$$|E(z)-E_n(z)| \leqq c_2 q^n \ln \frac{1}{q^n} = c_3 M\left(\frac{1}{v_n}\right)$$

as n is suitable large.

By this and from (3. 7):

$$\left|\sum_{k=1}^{\infty} \Phi\left(\frac{x}{\sqrt{k}\frac{\sqrt{2}}{\lambda}\sqrt{q^n}}\right) \mathsf{P}(v_n=k) - \int_{-\infty}^{x} \frac{\lambda}{2} e^{-\lambda|y|}\, dy\right| \leqq \frac{1}{2} c_3 M\left(\frac{1}{v_n}\right).$$

This and the estimation (3. 4) give together

$$\int_{-\infty}^{x} \frac{\lambda}{2} e^{-\lambda|y|}\, dy - \frac{1}{2} c_3 M\left(\frac{1}{v_n}\right) - c_1 M\left(\frac{1}{\sqrt{v_n}}\right) \leqq \tag{3. 8}$$

$$\leqq \mathsf{P}\left(\sqrt{q^n}\, \zeta_{v_n} < x\right) \leqq$$

$$\leqq \int_{-\infty}^{x} \frac{\lambda}{2} e^{-\lambda|y|}\, dy + \frac{1}{2} c_3 M\left(\frac{1}{v_n}\right) + c_1 M\left(\frac{1}{\sqrt{v_n}}\right).$$

This proves the statement (3. 1) of Theorem 2.

4. Case of a more general rarefaction and further developings

J. MOGYORÓDI ([7]) has investigated a more general rarefaction when ν_1 is an arbitrary positive integer-valued random variable — see the definition of ν_1 in section 1. — for which $1 < \mathsf{M} = \mathsf{M}(\nu_1) < +\infty$ and $\mathsf{D}^2(\nu_1) < +\infty$. If $f(z)$, for $|z| \leq 1$, is the generating function of the random variable ν_1 then it is easy to see that the generating function of random variable ν_n is the n-th iteration of $f(z)$, i.e. this is $f_n(z)$. J. MOGYORÓDI ([7]) has proved by the theory of Galton—Watson processes that

$$\text{(4. 1)} \qquad \lim_{n\to\infty} \mathsf{P}\left(\frac{\nu_n}{\mathsf{M}^n} < x\right) = G(x),$$

where $G(x)$ is a distribution function whose expectation is 1 and its variance is $\dfrac{\mathsf{D}^2(\nu_1)}{\mathsf{M}^2 - \mathsf{M}}$.

Now, however, the MOGYORÓDI's limiting distribution theorem has the following meaning for the sum

$$\zeta_{\nu_n} = \xi_1 + \xi_2 + \cdots + \xi_{\nu_n}$$

of a random number of non-negative, independent and commonly distributed random variables: when the common expectation of the summands ξ_k, $k = 1, 2, \ldots$ is a finite number a then the distribution function of the suitable normed sum of ζ_{ν_n} converges, as $n \to +\infty$, to the distribution function $G(x)$, or more precisely:

$$\text{(4. 2)} \qquad \mathsf{P}\left(\frac{\zeta_{\nu_n}}{\mathsf{M}^n} < x\right) \to G\left(\frac{x}{a}\right), \qquad n \to +\infty, \ a \neq 0.$$

The following limit distribution theorem can be proved for arbitrary real-valued, independent and commonly distributed summands with mean-value 0.

THEOREM 3. *Let $\xi_1, \xi_2, \ldots, \xi_k, \ldots$ be a sequence of independent, commonly distributed random variables for which $\int_{-\infty}^{+\infty} x\,dF(x) = 0$ and $\int_{-\infty}^{+\infty} x^2 dF(x) = 2\sigma^2 < +\infty$ are satisfied, where $\mathsf{P}(\xi_k < x) = F(x)$ is the common distribution function.*

Moreover let ν_n be independent of the sequence $\xi_1, \xi_2, \ldots, \xi_k, \ldots$ positive integer-valued random variables, with generating functions $f_n(z)$, $n = 1, 2, \ldots$, and let $1 < f'(1) = \mathsf{M} < +\infty$ and $f''(1) < +\infty$.

Denoting by $\zeta_{\nu_n} = \xi_1 + \xi_2 + \cdots + \xi_{\nu_n}$, we have

$$\lim_{n\to\infty} \mathsf{P}\left(\frac{\zeta_{\nu_n}}{\sqrt{\mathsf{M}^n}} < x\right) = \int_0^{+\infty} \Phi\left(\frac{x}{\sigma\sqrt{2z}}\right) dG(z), \qquad \sigma > 0,$$

where $\Phi(x)$ is the standard normal distribution function.

PROOF. Starting from limiting relations (2. 2) and (4. 1), instead of (2. 2) and (2. 3), the proof can be completed by the thoughts of the second proof of Theorem 2., hence it will not be detailed.

Remarks.

1. It is difficult to estimate the rate of convergence of limit distribution theorems concerning the more general rarefaction due to J. MOGYORÓDI ([7]). These theorems are based on formula (4. 1) which can be proved by martingale convergence theorem; it causes all difficulties. However J. MOGYORÓDI has proved some results in a manuscript, which also would be interesting at the theory of Galton—Watson processes.

2. It can be given the practicable limiting distributions for sums of random number of random variables related to the rarefaction due to A. RÉNYI ([8]) and their domains of attraction in that case also when the expectation of addable sums are zero and their variances are finite or infinite. It is possible by the method of I. N. KOVALENKO ([6]) and B. V. GNEDENKO ([3]). This will not be detailed here. In his manuscript J. MOGYORÓDI investigated similar questions for the sum of random number of random variables concerning his more general rarefaction method, too.

3. Introducing the notion of the rarefaction of one dimensional random wandering it can be investigated the analogies of invariance problems defined for recurrent processes (see in [7] and [8]). In the case of zero expectation the following result is true: only the one dimensional random wandering constructed by finite variance distributions can be invariant for the once rarefaction and the subsequent suitable coordinate transformation, and, among them, exactly those are invariant, which can be constructed by the limit distributions being at section 2. and 4.

Finally, I express my sincere thanks to J. MOGYORÓDI for his valuable remarks.

REFERENCES

[1] Добрушин, Р. Л.: Лемма о пределе сложной случайной функции, *Успехи Математических Наук*, **10** (1955) 157—159.

[2] DWIGHT, H. B.: *Tables of integrals and other mathematial data,* New York, Macmillan Company, 1961.

[3] Гнеденко, Б. В. и Фрайер,: Несколько замечаний к одной работе И. Н. Коваленко, *Лит. Мат. Сбор.*, IX. (1969), № 3., 463—470.

[4] GNEDENKO, B. V. and KOLMOGOROV, A. N.: *Független valószínűségi változók összegeinek határeloszlásai,* Akadémiai Kiadó Budapest, 1951.

[5] KONDRATAI TI GENE: *Retejanciu sriantu ribines teoremos patikslinimas,* Lithuanian university dissertation, Vilnius, 1968.

[6] Коваленко, И. Н.: О классе предельных распределений для редеющих потоков однородных событий, *Лит. Мат. Сбор.*, V. (1965), № 4., 569—573.

[7] MOGYORÓDI, J.: A rekurrens folyamat ritkításáról, *MTA III. Oszt. Közl.*, **19** (1969) 25—31.

[8] RÉNYI, A.: A Poisson folyamat egy jellemzése, *MTA Mat. Kut. Int. Közl.*, **1** (1956) 519—527.

Technical University of Budapest

(Received September 24, 1970)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 453—456.

ON AN INVARIANCE PROBLEM RELATED TO DIFFERENT RAREFACTIONS OF RECURRENT PROCESSES

By

T. SZÁNTAI

1. Preliminaries and fundamental notations

Let us consider a recurrent process and make a rarefaction of it as follows: let us maintain, independently from one another, every event of the process with probability q and cancel them with probability $1-q$, where $0<q<1$. Moreover let us make a coordinate transformation in the new process such that the expectation of the time interval between consecutive events be equal to the original one. A. RÉNYI raised the question, (see in [5]), what kind of recurrent processes are invariant for the above introduced rarefaction and coordinate transformation. Under invariance we mean that the distribution function of the time interval between two consecutive events is the same as before the rarefaction and coordinate transformation.

We remark that it is easy to see from RÉNYI's limiting distribution theorem (see in [5]) that among the recurrent processes having finite expectation only the homogeneous Poisson process is invariant for the once rarefaction and the subsequent coordinate transformation. Indeed, if we apply infinitely the rarefaction and the coordinate transformation to the finite expectation recurrent process which is invariant in all steps then it should be equal to the limiting recurrent process, i.e.: to a homogeneous Poisson process possessing the same intensity. It is also evident that the homogeneous Poisson process is in fact invariant for the rarefaction and the suitable normalization operation. This is reasonable by proving the existence of the three characteristic properties (see them in FELLER's fundamental book [1]) of the homogeneous Poisson process for the process obtained from the original homogeneous Poisson process. It is also easy to get this result by the method of the characteristic functions (see in [5]). A third proof based on a theorem due to A. PRÉKOPA (see in [4]) concerning the secondary processes generated by a random point distribution of Poisson type, is as follows. Let us consider the original homogeneous Poisson process on the axis x of the plane (x, y) and let us derive a secondary process in the plane (x, y) with the aid of the independent and commonly distributed random variables ξ_i, $i=0, 1, 2, \dots$, $\mathsf{P}(\xi_i=1)=q$, $\mathsf{P}(\xi_i = 2) = 1-q$. A. PRÉKOPA has shown that it is a Poisson process in the plane. Since the events of the secondary process being on the line $y=1$ are exactly the events of the process obtained by the rarefaction, we obtain on the basis of the properties of two-dimensional Poisson processes that the rarefied process should be a homogeneous Poisson process, too. Thus a finite expectation recurrent process is invariant under the rarefaction and the coordinate transformation if and only if it is a homogeneous Poisson process. In addition A. RÉNYI has proved by the method of characteristic functions, that an infinite expectation recurrent process is never invariant for the rarefaction and the suitable normalization operation.

If the number of events cancelled from the original process is denoted by $\nu-1$, then it is easy to see that, in the case of the above defined rarefaction, ν is geometrically distributed random variable with parameter q. J. MOGYORÓDI investigated a more general rarefaction procedure (see in [3]), when ν is an arbitrary positive integer valued random variable for which $1 < \mathsf{M} = \mathsf{M}(\nu) < +\infty$ and $\mathsf{D}^2(\nu) < +\infty$*. In this case it is also evident that only the limit recurrent process of recurrent processes with finite expectation maybe invariant for the rarefaction and the subsequent suitable coordinate transformation. In [3] J. MOGYORÓDI has proved by the aid of the Galton—Watson processes that the limit recurrent process is really invariant.

The purpose of this note is to show that a recurrent process with infinite expectation is never invariant for the above defined more general rarefaction procedure and the subsequent suitable normalization. The following theorem will be proved:

THEOREM. *If a recurrent process is invariant under the above defined general rarefaction and normalization procedure then the expectation of time interval between consecutive events is necessarily finite.*

2. The PROOF of the theorem. It is easy to see (see in [3]) that for the Laplace—Stieltjes transform $\varphi(s)$ of distribution function $F(x)$ of the invariant recurrent process the functional equation

$$\varphi(s) = f\left(\varphi\left(\frac{s}{\mathsf{M}}\right)\right), \quad \operatorname{Re} \quad s \geqq 0 \tag{2.1}$$

holds, where $f(z)$, for $|z| \leqq 1$, is the generating function of the random variable ν (see the definition of ν in 1.) and $\mathsf{M} = \mathsf{M}(\nu)$. By applying n-times 2. 1. and considering it only for $s \geqq 0$ we have:

$$\varphi(s) = f_n\left(\varphi\left(\frac{s}{\mathsf{M}^n}\right)\right), \qquad s \geqq 0, \ n = 1, 2, \ldots \tag{2.2}$$

where $f_n(z)$ is the n-th iteration of $f(z)$. Thus it is enough to consider $f_n(z)$ for only real values, $0 \leqq x \leqq 1$. But for real x in [0, 1] there exists the inverse function of $f(x)$, denoted by $u(x)$, and by induction it can be proved that the inverse of $f_n(x)$ also exists and it is equal to the n-th iteration of $u(x)$, i.e. this is $u_n(x)$.

Taking the inverse of the functional equation 2. 2. we get

$$u_n(\varphi(s)) = \varphi\left(\frac{s}{\mathsf{M}^n}\right), \qquad s \geqq 0, \ n = 1, 2, \ldots$$

wherehence for positive s

$$\frac{u_n(\varphi(s)) - 1}{\dfrac{s}{\mathsf{M}^n}} = \frac{\varphi\left(\dfrac{s}{\mathsf{M}^n}\right) - 1}{\dfrac{s}{\mathsf{M}^n}}, \qquad s > 0, \ n = 1, 2, \ldots \tag{2.3}$$

is satisfied.

* Throughout the paper M stands for the expectation of a random variable and D^2 denotes its variance.

Let us introduce the function $v(x) = u(x+1)-1$, then $v_n(x) = u_n(x+1)-1$ and $u_n(x) = v_n(x-1)+1$, and for (2. 3) we obtain

$$\frac{v_n(\varphi(s)-1)}{\frac{s}{M^n}} = \frac{\varphi\left(\frac{s}{M^n}\right)-1}{\frac{s}{M^n}}, \qquad s>0,\ n=1,2,\ldots \tag{2. 4}$$

Let us consider the Schröder functional equation

$$\chi(v(x)) = \frac{1}{M}\chi(x), \qquad x\in I,\ I=(-1,0]. \tag{2. 5}$$

The function $v(x)$ satisfies all conditions of Theorem 6. 1. from [2].

Indeed, $v'(0) = u'(1) = \frac{1}{f'(1)} = \frac{1}{M}$, $0<v'(0)<1$, $0\in I$, and since $D^2(v)<+\infty$, $v(x)$ is twice continuously differentiable function on the interval $I=(-1,0]$, moreover $u(x)<1$ for $x\in(0,1)$ hence $[v(x)-0]\cdot[0-x] = [u(x+1)-1]\cdot[-x] < 0$ if $x\in(-1,0)$ and from $f'(1)=M>1$ $f(x)<x$ follows whenever $x\in(0,1)$ that is $u(x)>x$ for $x\in(0,1)$ and as a consequence of the preceding

$$[v(x)-x]\cdot[0-x] = [u(x+1)-1-x]\cdot[-x] > 0$$

when $x\in(-1,0)$.

On the basis of the Theorem 6. 1. of [2] there exists only one solution $\chi(x)$ of equation (2. 5) being twice continuously differentiable on I and for which $\chi'(0)=1$ This solution can be written in the form as follows

$$\chi(x) = \lim_{n\to\infty}\frac{v_n(x)}{\frac{1}{M^n}}, \qquad x\in(-1\ 0], \tag{2. 6}$$

where the limit exists in the sense of uniform convergence.

If we consider (2. 4) as $n\to+\infty$ and regard (2. 6), we obtain:

$$\lim_{n\to\infty}\frac{\varphi\left(\frac{s}{M^n}\right)-1}{\frac{s}{M^n}} = \lim_{n\to\infty}\frac{v_n(\varphi(s)-1)}{\frac{s}{M^n}} = \frac{\chi(\varphi(s)-1)}{s}, \quad \text{if} \quad s>0. \tag{2. 7}$$

Let us choose a positive number s_0 such that $\frac{\chi(\varphi(s_0)-1)}{s_0} > -\infty$. The continuity of $\chi(x)$ and $\varphi(s)$, further the facts that $\varphi(0)=1$ and $\chi(0)=0$, always imply the existence of the number s_0.

Taking (2. 7) at the point s_0 in integral form we get that

$$\lim_{n\to\infty}\int_0^{+\infty}\frac{1-e^{-\frac{s_0}{M^n}y}}{\frac{s_0}{M^n}}\,dF(y) = -\frac{\chi(\varphi(s_0)-1)}{s_0} < +\infty.$$

We see that the sequence of the integrals of the sequence $f_n(y) = \dfrac{1-e^{-\frac{s_0}{M^n}y}}{\frac{s_0}{M^n}} \geqq 0$ is bounded and since $f_n(y) \to f(y) = y$, as $n \to +\infty$, we obtain by Fatou's lemma:

$$\int_0^{+\infty} y\,dF(y) \leqq \varliminf_{n\to\infty} \int_0^{+\infty} \frac{1-e^{-\frac{s_0}{M^n}y}}{\frac{s_0}{M^n}}\,dF(y) =$$

$$= \lim_{n\to\infty} \int_0^{+\infty} \frac{1-e^{-\frac{s_0}{M^n}y}}{\frac{s_0}{M^n}}\,dF(y) = -\frac{\chi(\varphi(s_0)-1)}{s_0} < +\infty.$$

The proof of the theorem is completed.

Finally, I express my sincere gratitude to J. MOGYORÓDI for his valueable remarks.

REFERENCES

[1] FELLER, W.: *An Introduction to Probability Theory and its Applications,* John Wiley, New York, 1966.
[2] KUCZMA, M.: *Functional equations in a single variable,* Varso, 1968, Polish Scientific Publishens.
[3] MOGYORÓDI, J.: A rekurrens folyamat ritkításáról, *MTA III. Oszt. Közl.,* **19** (1969) 25—31.
[4] PRÉKOPA, A.: On secondary processes generated by a random point distribution of Poisson type, *Ann. Univ. Sc. Budapest de Rol. Eötvös nom. Sectio Math.* **1** (1958) 153—170.
[5] RÉNYI, A.: A Poisson folyamat egy jellemzése, *MTA Mat. Kut. Int. Közl.,* **1** (1956) 519—527.

Technical University of Budapest

(Received September 24, 1970)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 457—465.

REGULARIZATION OF CERTAIN OPERATOR EQUATIONS BY FILTERS

by

K. MARTON and L. VARGA

1. Introduction

Consider the integral equation

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} k(y) f(x-y)\, dy = g(x) \tag{1.1}$$

where the functions $k \in \mathscr{L}_1$, $g \in \mathscr{L}_2$ are given*. We wish to determine the function $f \in \mathscr{L}_2$.

We suppose that the function g is available only in the sense that we know an approximation to it in the space $\mathscr{L}_2$, i.e. we are given a function $\tilde{g} \in \mathscr{L}_2$ such that $\|g-\tilde{g}\|_2$ is small**. We wish to find, by making use of $\tilde{g}$, an approximation (in the sense of the norm $\|\cdot\|_2$) to the exact solution of (1. 1).

In this paper we shall deal with a slightly more general equation than (1. 1). (It will be formulated in section 2.) We use equation (1. 1) as an example to illustrate some difficulties which may arise in solving equations of the form (2. 5) below.

Equation (1. 1) may be written in the form

$$Kf = g \tag{1.2}$$

where the operator K is defined by the formula

$$Kf(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} k(y) f(x-y)\, dy. \tag{1.3}$$

(As it is well known, K is a bounded linear mapping from $\mathscr{L}_2$ into $\mathscr{L}_2$; c.f. e.g. [1], p. 397.)

Equation (1. 2) (or, what is the same thing, equation (1. 1)) is incorrect, i.e. the operator K on the left of (1. 2) has no single-valued continuous inverse, defined on the whole space $\mathscr{L}_2$. To prove this statement, let $\hat{k}$ be the Fourier transform of k,

$$\hat{k}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} k(x) e^{ixt}\, dx \qquad (t \text{ real}) \tag{1.4}$$

and let F denote the Fourier—Plancherel operator. As is well known, F is an isometry of $\mathscr{L}_2$ onto itself ([1], p. 411.).

* By $\mathscr{L}_1$ ($\mathscr{L}_2$) we denote the space of all complex valued integrable (square integrable) functions of a real variable.

** $\|.\|_2$ denotes the usual norm on $\mathscr{L}_2$.

According to the properties of the Fourier and Fourier—Plancherel operators, equation (1. 1) is then equivalent to

$$\hat{k}F(f) = F(g). \tag{1. 5}$$

On the other hand from the Riemann—Lebesgue lemma we have

$$\lim_{t \to \pm\infty} \hat{k}(t) = 0. \tag{1. 6}$$

Hence, for $g \in \mathcal{L}_2$, the quotient $\dfrac{F(g)}{\hat{k}}$ may not belong to $\mathcal{L}_2$ (in fact, it may not exist). (1. 6) also shows that for $f, f_n \in \mathcal{L}_2$ $(n = 1, 2, \ldots)$ the relation

$$\lim_{n \to \infty} \|Kf_n - Kf\|_2 = 0$$

does not imply

$$\lim_{n \to \infty} \|f_n - f\|_2 = 0.$$

Therefore, though for $g \in K\mathcal{L}_2$ equation (1. 1) does have a solution $f \in \mathcal{L}_2$, the function f may depend on g discontinuously.

Since equation (1. 1) is incorrect, it seems natural to look for its solution by making use of TIHONOV's regularization method. (C.f. [2], [3]. We recall the definition in section 3.)

In this paper we propose a method for solving equations of the more general form (2. 5). The solution we propose is a special case of TIHONOV's regularization, and, at the same time, a common generalization of some special methods for solving equations of the form (2. 5). We shall see that in this way these methods become comparable. The method we propose will be referred to as the *method of regularization by filters.*

In section 2 we formulate the equation we deal with. Section 3 contains the definition of TIHONOV's regularization. In section 4 we formulate the method of regularization by filters, and prove, that it is a special case of TIHONOV's regularization. In section 5 two methods for solving equations of the form (2. 5) (an iterative method and one related to that formulated by D. PHILLIPS in [4]) are proved to be special cases of regularization by filters.

2. A generalization of equation (1.1)

Let $\mathcal{X}$ denote the following linear subset of $\mathcal{L}_2$:

$$\mathcal{X} = \{f \in \mathcal{L}_2 : \hat{l}F(f) \in \mathcal{L}_2 \text{ for } \hat{l} \in \Lambda\} \tag{2. 1}$$

where

$$\Lambda = \{1, \hat{k}, \hat{l}_1, \hat{l}_2, \ldots, \hat{l}_m\}.$$

Here $\hat{k}, \hat{l}_1, \hat{l}_2, \ldots, \hat{l}_m$ are complex valued measurable functions of a real variable, and 1 denotes the function identically equal to 1 on the real line. (For $m = 0$ $\Lambda = \{1, \hat{k}\}$.)

Since $\mathscr{X}$ lies in $\mathscr{L}_2$ (i.e. in the domain of F), we may (and do) define the set $\hat{\mathscr{X}}$ as the image of $\mathscr{X}$ under F:

$$\hat{\mathscr{X}} = F(\mathscr{X}) = \{\hat{f}\in\mathscr{L}_2 : \hat{l}\cdot\hat{f}\in\mathscr{L}_2 \text{ for } \hat{l}\in\Lambda\}.$$

Let us define a norm on $\hat{\mathscr{X}}$ as follows:

$$\|\hat{f}\|_{\hat{\mathscr{X}}} = \max_{\hat{l}\in\Lambda} \|\hat{l}\cdot\hat{f}\|_2 \qquad (\hat{f}\in\hat{\mathscr{X}}). \tag{2. 2}$$

It is easy to see that $\|\,.\,\|_{\hat{\mathscr{X}}}$ is a norm, indeed, which makes $\hat{\mathscr{X}}$ a Banach space.

Finally, let us define a norm on $\mathscr{X}$ by setting

$$\|f\|_{\mathscr{X}} = \|F(f)\|_{\hat{\mathscr{X}}} \qquad (f\in\mathscr{X}). \tag{2. 3}$$

Using the fact that F is an isometry of $\mathscr{L}_2$ onto itself, it is easy to see that (2. 3) defines a norm on $\mathscr{X}$, making the space $\mathscr{X}$ a Banach space, too.

Define a linear operator K on $\mathscr{X}$ by the formula

$$F(Kf) = \hat{k}F(f). \tag{2. 4}$$

According to (2. 1), for $f\in\mathscr{X}$ we have $\hat{k}F(f)\in\mathscr{L}_2$, so K is a mapping from $\mathscr{X}$ into $\mathscr{L}_2$. (Note that K is bounded, in fact, we have $\|K\|\leqq 1$. As a matter of fact, for every $f\in\mathscr{X}$ the definitions (2. 1)—(2. 4) imply

$$\|Kf\|_2 = \|F(Kf)\|_2 = \|\hat{k}F(f)\|_2 \leqq \|F(f)\|_{\hat{\mathscr{X}}} = \|f\|_{\mathscr{X}}.)$$

We are going to deal with the equation

$$Kf = g \qquad (f\in\mathscr{X},\ g\in\mathscr{L}_2). \tag{2. 5}$$

We suppose the function g to be known in the sense, that we are given an approximation to it in $\mathscr{L}_2$, i.e. we know a function $\tilde{g}\in\mathscr{L}_2$ such that $\|g-\tilde{g}\|_2$ is small. We wish to find, by making use of $\tilde{g}$, an approximation (in the sense of the norm $\|\,.\,\|_{\mathscr{X}}$) to the exact solution of (2. 5).

Equation (1. 1) is a special case of equation (2. 5), corresponding to the set $\Lambda=\{1, \hat{k}\}$, where $\hat{k}$ has the form (1. 4) with $k\in\mathscr{L}_1$. As a matter of fact, for this Λ we have $\mathscr{X}=\mathscr{L}_2$ and the norms $\|\,.\,\|_{\mathscr{X}}$ and $\|\,.\,\|_2$ are equivalent ([5], p. 102.).

3. Tihonov's regularization method

Definition (c.f. [2], [3])

Let $(\mathscr{Y}, \|\,.\,\|_{\mathscr{Y}})$ and $(\mathscr{Z}, \|\,.\,\|_{\mathscr{Z}})$ be Banach spaces, and let R be a linear operator from $\mathscr{Y}$ into $\mathscr{Z}$. Let A denote a set of real numbers, the closure of A containing the number 0. For each $\alpha\in A$ let S_α be a linear operator from $\mathscr{Z}$ into $\mathscr{Y}$.

The set of operators $\{S_\alpha : \alpha\in A\}$ is said to regularize the equation

$$Ry = z \qquad (y\in\mathscr{Y},\ z\in\mathscr{Z}), \tag{3. 1}$$

if the following conditions hold:

a) for each $\alpha\in A$ S_α is bounded,

b) for each $y \in \mathscr{Y}$ we have $\lim_{\alpha \to 0} \|S_\alpha Ry - y\|_{\mathscr{Y}} = 0$.

If R has a single-valued continuous inverse, defined on the whole space $\mathscr{X}$, then, setting $A = \{0\}$, $S_0 = R^{-1}$ yields a set of operators (consisting of a single element) which regularizes equation (3. 1). If this is not the case then the inverse of R may be replaced in some sense with an arbitrary set of operators $\{S_\alpha : \alpha \in A\}$, regularizing equation (3. 1). As a matter of fact, for each $y \in \mathscr{Y}$ the above definition implies

$$\lim_{\|Ry - \tilde{z}\|_{\mathscr{X}} \to 0} \inf_{\alpha \in A} \|y - S_\alpha \tilde{z}\|_{\mathscr{Y}} = 0,$$

that is, $S_\alpha \tilde{z}$ approaches the exact solution of (3. 1), if $\tilde{z}$ approaches z and if α is properly chosen (depending both on y and the distance $\|z - \tilde{z}\|_{\mathscr{X}}$).

4. Regulalization by filters

We have seen in section 1, that equation (2. 5) may be incorrect, so it seems natural to try to solve it by making use of TIHONOV's regularization method.

We shall need the following

Definition. Let A denote a set of real numbers, the closure of A containing the number 0. For each $\alpha \in A$ let D_α be a complex valued measurable function of a real variable.

The set of functions $\{D_\alpha : \alpha \in A\}$ is said to be a *filter*, regularizing equation (2. 5), if the following conditions hold:

(4. 1) a) $\sup_\alpha \text{ess.}\sup_t |D_\alpha(t)| < \infty$, *

(4. 2) b) $\text{ess.}\sup_t \left| \dfrac{D_\alpha(t)}{\hat{k}(t)} \right| < \infty$ for each $\alpha \in A$,

(4. 3) c) $\text{ess.}\sup_t \left| \dfrac{\hat{l}_i(t) D_\alpha(t)}{\hat{k}(t)} \right| < \infty$ for each $\alpha \in A$ and $i = 1, 2, \ldots, m$,

(4. 4) d) $\lim_{\alpha \to 0} D_\alpha(t) = 1$ for almost every t.

The above definition is justified by the following

THEOREM 1. *Let* $\{D_\alpha : \alpha \in A\}$ *be a filter, regularizing equation* (2. 5), *and for each* $\alpha \in A$ *let* S_α *be a linear operator on* $\mathscr{L}_2$, *defined by the formula*

$$F(S_\alpha g) = \frac{F(g)}{\hat{k}} D_\alpha \qquad (g \in \mathscr{L}_2). \tag{4. 5}$$

Then the set of operators $\{S_\alpha : \alpha \in A\}$ *regularizes equation* (2. 5).

* If h is a real valued measurable function of a real variable then by $\text{ess.}\sup_t h$ we denote the smallest number γ, satisfying the inequality $h(t) \leq \gamma$ for almost every t.

PROOF. Conditions (4. 1)—(4. 3) show that for each $\alpha \in A$ S_α is a bounded linear operator from $\mathcal{L}_2$ into $\mathcal{X}$. As a matter of fact, for each $\alpha \in A$, $g \in \mathcal{L}_2$ and $\hat{l} \in \Lambda$ we have

$$\hat{l} \cdot F(S_\alpha g) = \hat{l} \frac{D_\alpha}{\hat{k}} F(g) \in \mathcal{L}_2,$$

i.e. for each $\alpha \in A$ and $g \in \mathcal{L}_2$

$$S_\alpha g \in \mathcal{X}.$$

Moreover, according to the definitions (2. 3) and (2. 2),

$$\|S_\alpha g\|_{\mathcal{X}} = \|F(S_\alpha g)\|_{\hat{\mathcal{X}}} = \max_{\hat{l} \in \Lambda} \left\| \hat{l} \frac{D_\alpha}{\hat{k}} F(g) \right\|_2 \leqq \max_{\hat{l} \in \Lambda} \operatorname{ess.sup}_t \left| \hat{l}(t) \frac{D_\alpha(t)}{\hat{k}(t)} \right| \cdot \|g\|_2 .$$

Now let $f \in \mathcal{X}$ be fixed. We are going to prove that

$$\lim_{\alpha \to 0} \|S_\alpha K f - f\|_{\mathcal{X}} = 0. \tag{4. 6}$$

Using (2. 3), (4. 5), (2. 4) and (2. 2), we obtain

$$\|S_\alpha K f - f\|_{\mathcal{X}}^2 = \|F(S_\alpha K f) - F(f)\|_{\hat{\mathcal{X}}}^2 = \|F(f)(1 - D_\alpha)\|_{\hat{\mathcal{X}}}^2 = \tag{4. 7}$$

$$= \max_{\hat{l} \in \Lambda} \int_{-\infty}^{\infty} |\hat{l}(t) F(f)(t)(1 - D_\alpha(t))|^2 \, dt$$

for each $\alpha \in A$.

Condition (4. 4) shows that the integrand in (4. 7) tends to zero almost everywhere as $\alpha \to 0$. Moreover, from the fact that $\hat{l} F(f) \in \mathcal{L}_2$ for each $\hat{l} \in \Lambda$, in conjunction with (4. 1), it follows that the dominated convergence theorem applies to the right-hand side of (4. 7). Thus the right-hand side of (4. 7) tends to zero as $\alpha \to 0$, and the proof is complete.

Note that a filter $\{D_\alpha : \alpha \in A\}$ is of practical use in solving equations of the form (2. 5) only if (4. 6) holds uniformly on some well-treatable subset of $\mathcal{X}$. This is the case, e.g., if (4. 4) holds uniformly on every finite interval. As a matter of fact, define

$$\mathcal{X}_{C,N} = \left\{ f \in \mathcal{X} : \|f\|_{\mathcal{X}} \leqq C, \ \max_{\hat{l} \in \Lambda} \int_{|t| \geqq T} |\hat{l}(t) F(f)(t)|^2 \, dt \leqq N^2(T) \text{ for all } T > 0 \right\},$$

where C is a positive number, and N denotes a positive function, defined on the set of all positive numbers and satisfying the condition

$$\lim_{T \to \infty} N(T) = 0.$$

From the relation

$$\|S_\alpha K f - f\|_{\mathcal{X}} \leqq C \sup_{|t| \leqq T} |1 - D_\alpha(t)| + \sup_\alpha \operatorname{ess.sup}_t |1 - D_\alpha(t)| \, N(T),$$

valid for every $f \in \mathcal{X}_{C,N}$ and $T > 0$, we infer that (4. 6) holds uniformly on $\mathcal{X}_{C,N}$.

5. Examples

1. Let A be the set of all positive numbers, and for each $\alpha>0$ let D_α be the characteristic function of the interval $\left[-\frac{1}{\alpha}, \frac{1}{\alpha}\right]$. The set $\{D_\alpha : \alpha>0\}$ is then a filter, regularizing equation (1. 1), provided that

$$\hat{k}(t) \neq 0 \tag{5. 1}$$

for each t. More generally, the set $\{D_\alpha : \alpha>0\}$ is a filter, regularizing equation (2. 5), provided that

$$\operatorname*{ess.\,sup}_{|t| \leqq T} \left| \frac{\hat{l}(t)}{\hat{k}(t)} \right| < \infty$$

for each $\hat{l} \in \Lambda$ and $T>0$. — A similar filter was considered in [6].

The following two examples are intended to show that some special methods for solving equations of the form (2. 5) are special cases of regularization by filters. By showing this, we show as well, that these methods are special cases of TIHONOV's regularization. In the case of example 3 this is not quite obvious.

2. Consider the equation (1. 1), and suppose that, instead of (5. 1), the condition

$$|1-\hat{k}(t)| < 1 \tag{5. 2}$$

holds for almost every t. The solution of equation (1. 1) can be then found by iteration according to the formulae

$$f_1 = g, \; f_n = f_{n-1} - Kf_{n-1} + g \qquad (n=2, 3, \ldots). \tag{5. 3}$$

It is easy to verify that (5. 3) is equivalent to the formulae

$$F(f_n) = \sum_{j=0}^{n-1} (1-\hat{k})^j F(g) = \frac{F(g)}{\hat{k}}(1-(1-\hat{k})^n) \qquad (n=1, 2, \ldots). \tag{5. 4}$$

Let $A = \left\{1, \frac{1}{2}, \frac{1}{3}, \ldots\right\}$ and for each $\alpha \in A$ define

$$D_\alpha = 1-(1-\hat{k})^{1/\alpha}.$$

From (5. 2) we infer that the set $\{D_\alpha : \alpha \in A\}$ is a filter, regularizing equation (1. 1). As a matter of fact, conditions a) and d) of section 4 hold obviously, while condition b) follows from the relations

$$\operatorname*{ess.\,sup}_{t} \left| \frac{D_\alpha(t)}{\hat{k}(t)} \right| = \operatorname*{ess.\,sup}_{t} \left| \sum_{j=0}^{1/\alpha-1} (1-\hat{k}(t))^j \right| \leqq \frac{1}{\alpha} \qquad (\alpha \in A).$$

On the other hand, (5. 4) may be written in the form

$$F(f_n) = \frac{F(g)}{\hat{k}} D_\alpha,$$

i.e.

$$f_n = S_\alpha g$$

where $\alpha = \frac{1}{n}$ and S_α is the operator, associated with D_α, as in section 4.

3. Let $\Lambda = \{1, \hat{k}, \hat{l}\}$ where the function $\hat{k}$ is of the form (1. 4) with $k \in \mathscr{L}_1$. Suppose that the functions $\hat{k}$ and $\hat{l}$ satisfy the following two conditions:

$$(5.6) \qquad |\hat{k}(t)|^2 + |\hat{l}(t)|^2 > 0$$

for almost every t, and

$$(5.7) \qquad \operatorname{ess.\,sup}_t \frac{|\hat{k}(t)|}{|\hat{k}(t)|^2 + |\hat{l}(t)|^2} < \infty.$$

Let $(\mathscr{X}, \|\cdot\|_{\mathscr{X}})$ be the Banach space, associated with the set Λ, as in section 2.

For each $g \in \mathscr{L}_2$ and $\alpha > 0$ define a functional $\Omega_{g,\alpha}$ on $\mathscr{X}$ by setting

$$\Omega_{g,\alpha}(f) = \|Kf - g\|_2^2 + \alpha \|\hat{l}F(f)\|_2^2 \qquad (f \in \mathscr{X}).$$

Here K is the operator, associated with $\hat{k}$, as in section 2 (or, equivalently, the restriction of the operator, defined by (1. 3), to the space $\mathscr{X}$). Finally, for each $\alpha > 0$ define an operator S_α on $\mathscr{L}_2$ by setting

$$(5.8) \qquad S_\alpha g = f_{g,\alpha} \qquad (g \in \mathscr{L}_2)$$

where $f_{g,\alpha} \in \mathscr{X}$ is determined by the condition

$$(5.9) \qquad \Omega_{g,\alpha}(f_{g,\alpha}) = \inf_{f \in \mathscr{X}} \Omega_{g,\alpha}(f).$$

THEOREM 2. *For each $\alpha > 0$ the operator S_α is well defined on the whole space $\mathscr{L}_2$ and is associated with the function*

$$(5.10) \qquad D_\alpha = \frac{|\hat{k}|^2}{|\hat{k}|^2 + \alpha |\hat{l}|^2},$$

as in section 4.

Remark 1. Conditions (5. 6) and (5. 7) imply that we have

$$(5.7') \qquad \operatorname{ess.\,sup}_t \frac{|\hat{k}(t)|}{|\hat{k}(t)|^2 + \alpha |\hat{l}(t)|^2} < \infty$$

and

$$(5.11) \qquad \operatorname{ess.\,sup}_t \frac{|\hat{k}(t)|\,|\hat{l}(t)|}{|\hat{k}(t)|^2 + \alpha |\hat{l}(t)|^2} \leqq \frac{1}{2\sqrt{\alpha}} < \infty$$

for each $\alpha > 0$. The comparison of (5. 7') and (5. 11) with (5. 10) shows that the set $\{D_\alpha : \alpha > 0\}$ is a filter, regularizing equation (2. 5).

PROOF of Theorem 2. It is easy to verify that $\Omega_{g,\alpha}$ is a convex differentiable ([5], p. 143) functional for each $g \in \mathscr{L}_2$ and $\alpha > 0$. Denote by $\Omega'_{g,\alpha}(f)$ the derivative of $\Omega_{g,\alpha}$ at a fixed $f \in \mathscr{X}$. An easy computation shows that to an arbitrary $h \in \mathscr{X}$ the linear functional $\Omega'_{g,\alpha}(f)$ makes correspond the number

$$(5.12) \quad (\Omega'_{g,\alpha}(f), h) = 2 \operatorname{Re} \int_{-\infty}^{\infty} \{[\hat{k}(t)\hat{f}(t) - \hat{g}(t)]\overline{\hat{k}(t)} + \alpha |\hat{l}(t)|^2 \hat{f}(t)\} \overline{\hat{h}(t)}\, dt \qquad (h \in \mathscr{X})$$

where we used the notations $\hat{f}=F(f)$, $\hat{g}=F(g)$, $\hat{h}=F(h)$, while by $\bar{\gamma}$ we denoted the complex conjugate of the number γ.

Expression (5. 12), combined with the convexity of $\Omega_{g,\alpha}$, shows, that condition (5. 9) holds if and only if

$$[\hat{k}(t)\hat{f}_{g,\alpha}(t)-\hat{g}(t)]\overline{\hat{k}(t)}+\alpha|\hat{l}(t)|^2\hat{f}_{g,\alpha}(t)=0$$

for almost every t, i.e. if and only if

$$\hat{f}_{g,\alpha}=\frac{\hat{g}}{\hat{k}}\frac{|\hat{k}|^2}{|\hat{k}|^2+\alpha|\hat{l}|^2}=\frac{F(g)}{\hat{k}}D_\alpha. \tag{5. 13}$$

Now it follows from (5. 7′) and (5. 11) that the function $\hat{f}_{g,\alpha}$, defined by (5. 13), belongs to $\hat{\mathscr{X}}$, i.e. $f_{g,\alpha}\in\mathscr{X}$, so that the proof is complete.

Corollary. The set of operators $\{S_\alpha:\alpha>0\}$, defined by formulae (5. 8)—(5. 9), regularizes equation (2. 5) in TIHONOV's sense.

Remark 2. Suppose that instead of (5. 6) the stronger condition

$$\hat{l}(t)\neq 0 \tag{5. 14}$$

holds for almost every t. Let $0\neq g\in\mathscr{L}_2$ and $0<\varepsilon<\|g\|_2$ be fixed.

It can be shown that there exists a unique function $f_g^{(\varepsilon)}\in\mathscr{X}$, which satisfies both of the following conditions:

$$\|Kf_g^{(\varepsilon)}-g\|_2\leqq\varepsilon \tag{5. 15}$$

and

$$\|\hat{l}F(f_g^{(\varepsilon)})\|_2=\inf\{\|\hat{l}F(f)\|_2:f\in\mathscr{X},\ \|Kf-g\|_2\leqq\varepsilon\}; \tag{5. 16}$$

in fact, we have

$$f_g^{(\varepsilon)}=f_{g,\alpha}$$

for some $\alpha>0$ (depending both on g and ε). Here we only show the existence of such a function $f_g^{(\varepsilon)}$, the unicity can be proved by a slight modification of Lagrange's principle of indefinite multipliers.

To prove the existence of $f_g^{(\varepsilon)}$ note first that, by virtue if (5. 10) and (5. 14), we have

$$\lim_{\alpha\to 0}D_\alpha(t)=1$$

and

$$\lim_{\alpha\to\infty}D_\alpha(t)=0$$

for almost every t. (5. 10) and (5. 14) show as well that the functions D_α depend on α continuously and are strictly decreasing with α.

Let us now define

$$\varepsilon_{g,\alpha}=\|Kf_{g,\alpha}-g\|_2$$

for all $\alpha>0$. We infer from the definition (2. 4) and Theorem 2 that $\varepsilon_{g,\alpha}$ may be written in the form

$$\varepsilon_{g,\alpha}=\|Kf_{g,\alpha}-g\|_2=\|F(Kf_{g,\alpha})-F(g)\|_2=\|\hat{k}F(f_{g,\alpha})-F(g)\|_2=\|F(g)(1-D_\alpha)\|_2.$$

Thus we see from what has been established on $D_\alpha(t)$, as a function of the parameter α, that the following four statements are true:

$$\lim_{\alpha\to 0} \varepsilon_{g,\alpha} = 0,$$

$$\lim_{\alpha\to\infty} \varepsilon_{g,\alpha} = \|g\|_2,$$

$\varepsilon_{g,\alpha}$ depends on α continuously and is strictly increasing with α. (Here we used the dominated convergence theorem.)

Consequently, there exists a (unique) number $\alpha > 0$ such that

$$\varepsilon_{g,\alpha} = \varepsilon,$$

and it is clear that relations (5. 15)—(5. 16) become true with $f_g^{(\varepsilon)}$ replaced by $f_{g,\alpha}$.

This interpretation shows that, if we set

$$\hat{l}(t) = t^2,$$

then the regularization of equation (2. 5) by the filter, described in this example, becomes very similar to that proposed in [4].

REFERENCES

[1] HEWITT, E. — STROMBERG, K.: *Real and Abstract Analysis*, Springer Verlag, 1965.

[2] Тихонов, А. Н.: О решении некорректно поставленных задач и методе регуляризции. *D.A.H. CCCP*, **151,** № 3, (1963) 501—504.

[3] Тихонов, А. Н.: О регуляризации некорректно поставленных задач. *D.A.H. CCCP*, **153,** № 1, (1963) 49—52.

[4] PHILLIPS, D. L.: A Technique for the Numerical Solution of Certain Integral Equations of the First Kind. *J.A.C.M.* **9,** № 1. (1962) 84—97.

[5] DIEUDONNÉ, J.: *Foundations of Modern Analysis*. Academic Press, New York and London, 1960.

[6] Страхов, В. Н.: О численном решении некорректных задач, представляемых интегральными уравнениями типа свёртки. *D.A.H. CCCP*, **178,** № 2, (1968) 299—302.

Central Research Institute of Physics, Budapest

(Received November 2, 1970)

Studia Scientiarum Mathematicarum Hungarica 6 (1971) 467—480.

REMARKS ON THE THREE BODY PROBLEM IN QUANTUM MECHANICS

by

E. MAKAI

1. Let x_i, y_i, z_i $(i=1, 2, 3)$ be the rectilinear coordinates of three points P_i in three dimensional space. If $\Delta_i = \partial^2/\partial x_i^2 + \partial^2/\partial y_i^2 + \partial^2/\partial z_i^2$, then the Schrödinger equation of the quantum mechanical three body problem is of the form

$$\left\{\sum_{i=1}^{3} \mu_i \Delta_i + U + E\right\}\psi = 0 \tag{1. 1}$$

where the μ_i's are non-negative constants, E is a real constant and U is some function of the variables $x_1, y_1, \ldots, z_3$. Let us write

$$r_1 = \overline{P_2 P_3}, \quad r_2 = \overline{P_3 P_1}, \quad r_3 = \overline{P_1 P_2}, ^* \tag{1. 2}$$

where $\overline{P_i P_k}$ denotes the distance between the points P_i and P_k and let us suppose that U depends only on the quantities r_1, r_2, r_3.

Then it is known that (1. 1) has solutions depending only on the quantities r_1, r_2, r_3 and some of these solutions — for some E's — are quadratically integrable in the whole nine dimensional space $x_1, y_1, \ldots, z_3$.

In some simple cases, e.g. if U is a constant or of the form $\Sigma c_i r_i^2$, i.e. if the system is a three-body harmonic oscillator, the solutions of (1. 1) may be found by group-theoretical methods, namely by calculating the mean of a certain simple solution over the group of rotations: the solutions are then given by way of multiple integrals.

In this paper we shall find some of these solutions avoiding the method mentioned in the previous paragraph, and the solutions will have the form of power series expansions.

In Section 2 we treat the case $U=0$ and show that there exist solutions of (1. 1) depending only on r_1, r_2, r_3 and satisfying the three differential equations $\Delta_1\psi + \lambda_1\psi = 0$, $\Delta_2\psi + \lambda_2\psi = 0$, $\Delta_3\psi + \lambda_3\psi = 0$, where the λ_i's are suitable constants. These solutions satisfy a first order equation, too.

In Section 3 we discuss a particular case of the following more general question. Let φ^1, φ^2, φ^3 be functions of r_1, r_2, r_3 only; in which cases have the equations $(\Delta_i + \varphi^i)\psi = 0$ $(i=1, 2, 3)$ not identically vanishing common solutions depending

* Formulas of this type will be frequently written in the abbreviated form

$$r_1 = \overline{P_2 P_3} \qquad \text{(cycl.)}$$

where the abbreviation in parentheses refers to that the formula in question remains true after a cyclic permutation of the indices 1, 2, 3.

only on r_1, r_2, r_3? It will appear that the class of functions φ^i we consider is very restricted. It contains, however, among others, functions of the form $\varphi^i = a^i + \sum_k a^{ik} r_k^2$ where a^i and a^{ik} are constants. In this case the equation (1. 1) or $(\Sigma\mu_i\Delta_i + \Sigma\mu_i\varphi^i)\psi = 0$ is the Schrödinger equation of the above mentioned three-body harmonic oscillator. This case will be discussed in Section 4.

Finally in Section 5 it will be shown that if in (1. 1) U and ψ depend only on r_1, r_2, r_3, then by introducing suitable new variables x, y, z (1. 1) can be written in the form

$$\left(\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} + \frac{1}{z}\frac{\partial\psi}{\partial z}\right) + W(x, y, z)\psi = 0.$$

This transformation, already known to G. T. Gronwall [2, 3] in the special case $\mu_1 = \mu_2 = 1$, $\mu_3 = 0$, will enable us to write the Schrödinger equation of the above mentioned three-body harmonic oscillator problem in a form depending only on two parameters.

Instead of the variables r_i we shall use throughout this paper their squares

$$(1.\ 3) \qquad \varrho_i = r_i^2$$

and we shall use the notation

$$(1.\ 4) \qquad \sigma_1 = -\varrho_1 + \varrho_2 + \varrho_3 \qquad \text{(cycl.)}.$$

Supposing that ψ depends only on $\varrho_1, \varrho_2, \varrho_3$ we have

$$\Delta_i\psi = \Delta^i\psi$$

where

$$(1.\ 5) \qquad \Delta^1 = 4\left(\varrho_2\frac{\partial^2}{\partial\varrho_2^2} + \frac{3}{2}\frac{\partial}{\partial\varrho_2} + \sigma_1\frac{\partial^2}{\partial\varrho_2\,\partial\varrho_3} + \varrho_3\frac{\partial^2}{\partial\varrho_3^2} + \frac{3}{2}\frac{\partial}{\partial\varrho_3}\right) \qquad \text{(cycl.)}.$$

We introduce the first order differential operators

$$(1.\ 6) \qquad N_1 = (\varrho_3 - \varrho_2)\frac{\partial}{\partial\varrho_1} - \varrho_2\frac{\partial}{\partial\varrho_2} + \varrho_3\frac{\partial}{\partial\varrho_3} \qquad \text{(cycl.)}$$

and

$$(1.\ 7) \qquad H = \Sigma\varrho_i\frac{\partial}{\partial\varrho_i}.$$

The operators N_1, N_2, N_3 are not independent, they are connected by the relation

$$(1.\ 8) \qquad \Sigma\varrho_i N_i = 0$$

and satisfy the following relations:

$$(1.\ 9) \qquad N_1\left(\frac{\partial}{\partial\varrho_2} + \frac{\partial}{\partial\varrho_3}\right) + N_2\left(\frac{\partial}{\partial\varrho_3} + \frac{\partial}{\partial\varrho_1}\right) + N_3\left(\frac{\partial}{\partial\varrho_1} + \frac{\partial}{\partial\varrho_2}\right) = 0,$$

$$(1.\ 10) \qquad \varrho_3\Delta^2 - \varrho_2\Delta^3 = 4\left(H + \frac{1}{2}\right)N_1, \qquad \text{(cycl.)}$$

$$(1.\ 11) \qquad \Sigma N_i\Delta^i = 0.$$

Finally let A and B be two operators acting on a function ψ and $\{A, B\} = AB-BA$. Then

$$(1.12)\qquad \{\Delta^i, \Delta^k\} = 0 \qquad (i, k=1, 2, 3)$$

and if the operator f is a function of $\varrho_1, \varrho_2, \varrho_3$ only, then

$$(1.13)\qquad \{\Delta^1, f\} = (\Delta^1 f)+4\frac{\partial f}{\partial \varrho_2}(N_3+H)+4\frac{\partial f}{\partial \varrho_3}(-N_2+H) \qquad \text{(cycl.)}$$

where

$$(1.14)\qquad (\Delta^1 f) = 4\left(\varrho_2\frac{\partial^2 f}{\partial \varrho_2^2}+\frac{3}{2}\frac{\partial f}{\partial \varrho_2}+\sigma_1\frac{\partial^2 f}{\partial \varrho_2\,\partial \varrho_3}+\varrho_3\frac{\partial^2 f}{\partial \varrho_3^2}+\frac{3}{2}\frac{\partial f}{\partial \varrho_3}\right) \qquad \text{(cycl.)}.$$

Formulae (1. 8)—(1. 14) can be verified in a direct manner.

2. Let us consider the differential equation

$$(2.1)\qquad (\Sigma\mu_i\Delta^i+E)\psi = 0$$

where ψ depends only on $\varrho_1, \varrho_2, \varrho_3$. We put the question: does a not identically vanishing solution ψ of (2. 1) exist, which satisfies also the three equations

$$(2.2_i)\qquad (\Delta^i+\lambda_i)\psi = 0 \qquad (i=1, 2, 3)$$

where the λ_i's are appropriate constants? If such a ψ exists, then applying the operator N_i to the equation (2. 2_i) and summing with respect to i, we have by (1. 11)

$$(2.3)\qquad \Sigma\lambda_i N_i\psi = 0$$

and of course

$$(2.4)\qquad E = \Sigma\mu_i\lambda_i.$$

According to the general theory of the homogeneous partial differential equations of the first order, the general solution of (2. 3) is a differentiable function of its any two particular solutions. Now it is easy to verify that

$$(2.5)\qquad u = \Sigma b_i\varrho_i, \qquad b_1 = -\lambda_1+\lambda_2+\lambda_3 \qquad \text{(cycl.)}$$

and

$$(2.6)\qquad v = \sum_{1,2,3}(\varrho_1^2-2\varrho_2\varrho_3) = -\sum_{1,2,3}\sigma_1\sigma_2 \quad *$$

are particular solutions of (2. 3), hence any common solution of the system (2. 2_1), (2. 2_2), (2. 2_3), is of the form $f(u, v)$. Putting $f(u, v)$ into (2. 2_3), say, we have

$$(2.7)\qquad 4(-2\lambda_3 u+\Lambda\varrho_3)\frac{\partial^2 f}{\partial u^2}-16\lambda_3 v\frac{\partial^2 f}{\partial u\,\partial v}-16\varrho_3 v\frac{\partial^2 f}{\partial v^2}-$$
$$-12\lambda_3\frac{\partial f}{\partial u}-16\varrho_3\frac{\partial f}{\partial v}+\lambda_3 f = 0,$$

* By $\sum_{1,2,3}$ we shall denote a three term sum the terms of which differ only by cyclic interchanges of the numbers 1, 2, 3 occurring in one of the terms. In particular, the right-hand side of (2.6) is $-\sigma_1\sigma_2-\sigma_2\sigma_3-\sigma_3\sigma_1$.

where

$$\Lambda = \sum_{1,2,3} (\lambda_1^2 - 2\lambda_2 \lambda_3). \tag{2.8}$$

In (2. 7) ϱ_3 is a parameter, so this differential equation is equivalent to the following two equations:

$$2u \frac{\partial^2 f}{\partial u^2} + 4v \frac{\partial^2 f}{\partial u \, \partial v} + 3 \frac{\partial f}{\partial u} - \frac{1}{4} f = 0 \tag{2.9}$$

$$-\frac{\Lambda}{4} \frac{\partial^2 f}{\partial u^2} + v \frac{\partial^2 f}{\partial v^2} + \frac{\partial f}{\partial v} = 0. \tag{2.10}$$

Substituting $f(u, v)$ into (2. 2_1) and (2. 2_2) it is seen that these equations, too, are reduced to the same equations (2. 9) and (2. 10).

It remains to find a solution of the system (2. 9) and (2. 10). We do this only in the case $\Lambda \neq 0$. (The case $\Lambda = 0$ can be dealt with afterwards by a limiting process.) Setting

$$f = f(u, v) = \sum_{k=0}^{\infty} \left(\frac{\Lambda}{4} \right)^k u_k v^k \tag{2.11}$$

where u_k is a function of u only and substituting (2. 11) into (2. 9) and (2. 10) we get

$$2u u_k'' + (4k+3) u_k' - \frac{1}{4} u_k = 0 \tag{2.12}$$

$$u_{k-1}'' = k^2 u_k \tag{2.13}$$

where the dashes denote differentation with respect to u. This pair of sets of equations can be solved in the following manner. In the case $k = 0$ we have from (2. 12)

$$2u u_0'' + 3u_0' - \frac{1}{4} u_0 = 0 \tag{2.14}$$

a solution of which is $e^{-\sqrt{u/2}}/\sqrt{u/2}$. Differentiating (2. 14) $2k$ times we have

$$2u u_0^{(2k+2)} + (4k+3) u_0^{(2k+1)} - \frac{1}{4} u_0^{(2k)} = 0 \tag{2.15}$$

which shows that $c_k u_0^{(2k)}$ is a solution of (2. 12). Taking $c_k = (k!)^{-2}$ we see that (2. 13), too, is satisfied. Thus a solution of system (2. 9)—(2. 10), or of the system (2. 2_1), (2. 2_2), (2. 2_3) is

$$\psi = \sum_{k=0}^{\infty} \left(\frac{\Lambda}{4} \right)^k \frac{v^k}{k!^2} \frac{d^{2k}}{du^{2k}} \frac{e^{-\sqrt{u/2}}}{\sqrt{u/2}}. \tag{2.16}$$

With the help of the formulae

$$J_0(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!^2} \left(\frac{x}{2} \right)^{2k} = \frac{1}{2\pi} \int_0^{2\pi} e^{ix \sin \vartheta} \, d\vartheta$$

and of the symbolical formula

$$e^{h\frac{d}{dx}} f(x) = f(x+h) \tag{2.17}$$

(2. 16) can be transformed in the following way:

$$\psi = J_0\left(\sqrt{-\Lambda v}\,\frac{d}{du}\right)\frac{e^{-\sqrt{u/2}}}{\sqrt{u/2}} = \frac{1}{2\pi}\int_0^{2\pi}\frac{\exp\left[-\sqrt{(u+\sin\vartheta\sqrt{\Lambda v})/2}\right]}{\sqrt{(u+\sin\vartheta\sqrt{\Lambda v})/2}}\,d\vartheta. \tag{2.18}$$

The last integral converges for all finite values of u and v and a direct substitution shows that it satisfies the equations (2. 2_i).

We shall not discuss the properties of the function (2. 16), since, after all, the particular case of equation (1. 1) we dealt with in the present section is of no great importance. The above discussion leads, however, to the following more general question.

3. Let us consider the system of differential equations

$$(\Delta^i+\varphi^i)\psi = 0 \qquad (i=1, 2, 3) \tag{3.1}$$

where φ^1, φ^2, φ^3 are functions of ϱ_1, ϱ_2, ϱ_3. Under what circumstances have these equations a non-trivial common solution? This solution, if it exists, satisfies, of course (1. 1) with $U+E = \Sigma\mu_i\,\varphi^i$ and by virtue of (1. 11), also a first order differential equation

$$\Sigma N_i\varphi^i\psi = \Sigma[\varphi^i N_i\psi+(N_i\varphi^i)\psi] = 0, \text{ *} \tag{3.2}$$

where e.g.

$$(N_1\varphi^1) = (\varrho_3-\varrho_2)\frac{\partial\varphi^1}{\partial\varrho_1}-\varrho_2\frac{\partial\varphi^1}{\partial\varrho_2}+\varrho_3\frac{\partial\varphi^1}{\partial\varrho_3}. \tag{3.3}$$

The equations

$$\{\Delta^1+\varphi^1, \Delta^2+\varphi^2\}\psi = 0 \qquad \text{(cycl.)} \tag{3.4}$$

are also to be satisfied. By virtue of (1. 12) and (1. 13) these latter three equations are also of the first order.

We are not able to answer the above question in full generality. An additional condition will be needed to treat this problem: *we require that the equations* (3. 4) *should be linear consequences of* (3. 2). This condition is met in the special case treated in Section 2.

It will be seen that it is a comparatively narrow class of functions φ^i which meets the above requirements.

Using the notation $\partial F/\partial\varrho_i = F_i$ for the partial derivatives of a function F (higher partial derivatives will be denoted similarly) the explicit form of (3. 4) is according to (1. 13) and (1. 14)

$$\begin{aligned}4[\varphi^2_2N_3-\varphi^2_3N_2-\varphi^1_3N_1+\varphi^1_1N_3+(\varphi^2_2+\varphi^2_3-\varphi^1_3-\varphi^1_1)H]\psi+\\+[(\Delta^1\varphi^2)-(\Delta^2\varphi^1)]\psi = 0 \qquad \text{(cycl.)}\end{aligned} \tag{3.5}$$

* $N_i\varphi^i\psi$ means $N_i(\varphi^i\psi)$. Similar notations will be used throughout this paper.

Since this is to be a linear consequence of (3. 2) and the operator H cannot be expressed as a linear consequence of the N_i-s, the coefficient of H vanishes:

$$\varphi_2^2+\varphi_3^2 = \varphi_1^1+\varphi_3^1 \quad \text{(cycl.)}, \tag{3. 6}$$

or, introducing the variables σ_i defined by (1. 4) instead of $\varrho_1, \varrho_2, \varrho_3$,

$$\frac{\partial\varphi^1}{\partial\sigma_2} = \frac{\partial\varphi^2}{\partial\sigma_1} \quad \text{(cycl.)}. \tag{3. 7}$$

These three equations have a solution if and only if there exists a function φ, such that

$$\varphi^i = 2\frac{\partial\varphi}{\partial\sigma_i} \qquad (i=1, 2, 3) \tag{3. 8}$$

or, returning to the variables ϱ_i,

$$\varphi^1 = \varphi_2+\varphi_3 \quad \text{(cycl.)}. \tag{3. 9}$$

Using (1. 9) we have $\Sigma N_i\varphi^i=0$, thus equation (3. 2) can be written in the form

$$\Sigma\varphi^i N_i\psi = 0 \tag{3. 10}$$

showing that the coefficient of ψ vanishes in (3. 2).

Again, since we have assumed that (3. 2) and (3. 5) were equivalent equations, the coefficient of ψ in (3. 5) vanishes:

$$(\Delta^1\varphi^2) = (\Delta^2\varphi^1) \quad \text{(cycl.)}. \tag{3. 11}$$

Using (3. 9) these equations yield

$$\varrho_1\varphi_{11}^1+(\varrho_1-\varrho_2)\varphi_{13}^1 = \varrho_2\varphi_{22}^2+(\varrho_2-\varrho_1)\varphi_{23}^2 \quad \text{(cycl.)}. \tag{3. 12}$$

These equations can be transformed into a simpler form using the three equations (3. 6). Indeed differentiating (3. 6) with respect to ϱ_1 and ϱ_2 and by the aid of these new equations eliminating φ_{11}^1 and φ_{22}^2 from (3. 12) we have

$$\varrho_1^{-1}(\varphi_{12}^1+\varphi_{23}^1+\varphi_{31}^1) = \varrho_2^{-1}(\varphi_{12}^2+\varphi_{23}^2+\varphi_{31}^2) \quad \text{(cycl.)}. \tag{3. 13}$$

Let us now introduce the notation

$$\Phi = \varphi_{12}+\varphi_{23}+\varphi_{31} \tag{3. 14}$$

by the aid of which equations (3. 13) can be written in the form

$$\frac{\Phi_2+\Phi_3}{\varrho_1} = \frac{\Phi_3+\Phi_1}{\varrho_2} = \frac{\Phi_1+\Phi_2}{\varrho_3} \tag{3. 15}$$

(cf. (3. 9)), or alternatively $N_1\Phi=0$, $N_2\Phi=0$. Now $N_i v=0$, $N_i\sigma_i=0$, where v and σ_i are defined by (2. 6) and (1. 4), respectively. Thus the general solution of $N_1\Phi=0$ is $f(v, \sigma_1)$. Substituting this into $N_2\Phi=0$ we have $\partial f/\partial\sigma_1=0$, hence the general solution of (3. 15) is a differentiable function of v, say, $\frac{1}{2}g'(v)$ where the dash denotes differentation with respect to v:

$$\sum_{1,2,3} \varphi_{12} = \frac{1}{2} g'(v) \quad \text{or} \quad \sum_{1,2,3} \varphi_i^i = g'(v). \tag{3.16}$$

We remark that in view of (3. 16) and the relations (3. 6), (3. 11) the equations (3. 5) can be written in the simpler form

$$\{\sum_i \varphi_j^i N_i - g'(v) N_j\}\psi = 0 \qquad (j=1, 2, 3) \tag{3.17}$$

and we recall that our aim was to investigate under what circumstances is each of these equations equivalent to (3. 10) or (3. 2).

An equivalent form of equations (3. 17) may be given by introducing the notations

$$\chi = \varphi + \frac{1}{4} g(v), \tag{3.18}$$

$$\chi^1 = \chi_2 + \chi_3 = \varphi^1 - \varrho_1 g'(v) \qquad \text{(cycl.)} \tag{3.19}$$

and by eliminating with the aid of these the φ^i's from (3. 17). Then we have by (1. 8), i.e. by the differential equation

$$\Sigma \varrho_i N_i \psi = 0 \tag{3.20}$$

satisfied by all differentiable functions,

$$\sum_i \chi_j^i N_i \psi = 0 \qquad (j=1, 2, 3). \tag{3.21$_j$}$$

Also, we have from (3. 10) and (3. 20)

$$\Sigma \chi^i N_i \psi = 0 \tag{3.10$'$}$$

and from (3. 16)

$$\sum_{1,2,3} \chi_{12} + g'(v) = 0, \quad \text{or} \quad \sum \chi_i^i + 2g'(v) = 0. \tag{3.22}$$

Excluding the not very interesting case $N_1\psi = N_2\psi = N_3\psi = 0$ (when ψ is a function of v alone) * and introducing the three component vectors

$$\boldsymbol{\varrho} = \{\varrho_1, \varrho_2, \varrho_3\}, \quad \mathbf{w} = \{\chi^1, \chi^2, \chi^3\}, \quad \mathbf{w}^j = \{\chi_j^1, \chi_j^2, \chi_j^3\} \tag{3.23}$$

we can reformulate our assumption on the equivalence of (3. 21_j) and (3. 10′) as follows: *we require that the five vectors* $\mathbf{w}^1, \mathbf{w}^2, \mathbf{w}^3, \mathbf{w}, \boldsymbol{\varrho}$ *should be complanar* **, i.e.

$$D^j \equiv \begin{vmatrix} \chi_j^1 & \chi_j^2 & \chi_j^3 \\ \chi^1 & \chi^2 & \chi^3 \\ \varrho_1 & \varrho_2 & \varrho_3 \end{vmatrix} \equiv \begin{vmatrix} \chi_{1j} & \chi_{2j} & \chi_{3j} \\ \chi_1 & \chi_2 & \chi_3 \\ \sigma_1 & \sigma_2 & \sigma_3 \end{vmatrix} = 0 \qquad (j=1, 2, 3). \tag{3.24$_j$}$$

* If $\psi = \psi(v)$, then $\Delta^i \psi = -16\varrho_i(v\partial^2\psi/\partial v^2 + \partial\psi/\partial v)$ (cf. (2.7)) and by (3.1) $\varphi^i = \varrho_i \varphi^0$, where φ^0 is a function of v only.

** Equation (3.20) is, of course, to be satisfied, too.

This is a set of quadratic second order differential equations the function χ is to satisfy and we proceed to the solution of it.

First we derive three other second order equations in the following way:

$$-\frac{\partial D^1}{\partial \varrho_2}+\frac{\partial D^2}{\partial \varrho_1}=\begin{vmatrix}\chi_1^1 & \chi_1^3\\ \chi^1 & \chi^3\end{vmatrix}+\begin{vmatrix}\chi_2^2 & \chi_2^3\\ \chi^2 & \chi^3\end{vmatrix}=0 \qquad \text{(cycl.)}.$$

Here we used the complanarity of the vectors $\mathbf{w}^1$, $\mathbf{w}^2$, $\boldsymbol{\varrho}$. From these equations we have by (3. 22)

$$\sum_i \chi^i\chi_i^3+2g'(v)\chi^3=0 \qquad \text{(cycl.)} \tag{3.25}$$

or, equivalently, (cf. (3. 19))

$$\sum_i \chi^i\chi_{ij}+2g'(v)\chi_j=\frac{1}{2}\frac{\partial}{\partial\varrho_j}\sum\chi^i\chi_i+2g'(v)\chi_j=0 \qquad (j=1,2,3). \tag{3.26$_j$}$$

Differentiating (3. 26_1) with respect to ϱ_2, (3. 26_2) with respect to ϱ_1 and subtracting we have

$$g''(v)(v_2\chi_1-v_1\chi_2)=0 \qquad \text{(cycl.)}. \tag{3.27}$$

We distinguish now two cases, $g''(v)\not\equiv 0$ and $g''(v)\equiv 0$.

Case 1: $g''(v)\not\equiv 0$. Then

$$\begin{vmatrix}\chi_1 & \chi_2\\ v_1 & v_2\end{vmatrix}=0 \qquad \text{(cycl.)} \tag{3.28}$$

hence $\chi=h(v)$, say and by (3. 19)

$$\varphi^i=\varrho_i\{g'(v)-4h'(v)\}=\varrho_i k(v), \tag{3.29}$$

say. The system (3. 1) is now solvable: there exist solutions depending only on v. (See the first footnote on p. 473.)

Case 2: $g''(v)\equiv 0$, or $g'(v)=c$. Now from the three equations (3. 26_j)

$$\frac{1}{2}\Sigma\chi^i\chi_i+2c\chi=2d \tag{3.30}$$

where d is another constant and χ satisfies beside this the seven equations (3. 26_j), (3. 24_j) and (3. 22). This system of second order equations consists of at most six independent equations as it is seen by eliminating the derivatives χ_{11}, χ_{22}, χ_{33} from (3. 26_j) and (3. 24_j). The result of this elimination yields but two independent equations

$$\beta^2\chi_{31}-\beta^3\chi_{12}=-c^1, \quad -\beta^1\chi_{23}+\beta^3\chi_{12}=-c^2, \tag{3.31}$$

where

$$c^1=2c(\sigma_3\chi_2-\sigma_2\chi_3)\chi_1, \quad \beta^1=\chi_1\alpha-\sigma_1\beta \qquad \text{(cycl.)} \tag{3.32}$$

and

$$\alpha=\Sigma\sigma_i\chi^i=2\Sigma\varrho_i\chi_i, \quad \beta=\Sigma\chi^i\chi_i=4(d-c\chi). \tag{3.33}$$

Regarding the system (3. 31), (3. 22) as a linear algebraic system of equations with unknowns $\chi_{12}, \chi_{23}, \chi_{31}$, its determinant is $D = -\beta(\frac{1}{2}\alpha^2 + v\beta)$ and if we suppose that $\beta \neq 0$, a solution of it is

$$\chi_{23} = -2\frac{c}{\beta}\chi_2\chi_3 \qquad \text{(cycl.)} \tag{3.34}$$

and this is the only solution, if $D \neq 0$. We distinguish again several cases according to whether the determinant D vanishes or not.

Case 2. 1. 1: $D \neq 0$, $c \neq 0$. This is the most interesting case. Let us write

$$\chi^* = \chi - \frac{d}{c},$$

then we have from (3. 34)

$$2\chi^*\chi_{23}^* = \chi_2^*\chi_3^* \qquad \text{(cycl.)} \tag{3.35}$$

and the equations (3. 24_j) remain valid if we substitute in them χ by χ^*. From these six equations each second partial derivative — and hence each higher derivative of χ^* — can be expressed by the aid of χ^* and its first derivatives. These latter four quantities are connected by the relation (3. 30) or

$$\sum_{1,2,3} \chi_1^*\chi_2^* = -2c\chi^*. \tag{3.36}$$

This means that the general solution χ^* of this system can contain at most three constants, apart from c; these may be the values of χ^*, χ_1^*, χ_2^* at a fixed place.

Now a four parameter solution of (3. 35) is

$$\chi^* = -\frac{1}{2}(\varepsilon_0 + \Sigma\varepsilon_i\varrho_i)^2, \qquad \varepsilon_0, \varepsilon_1, \ldots, \varepsilon_3 = \text{const.}$$

It is readily seen that this function satisfies the system (3. 24_j) and also (3. 36) if

$$\sum_{1,2,3} \varepsilon_1\varepsilon_2 = c \neq 0.$$

We shall return to this case in the next section.

Case 2. 1. 2: $D \neq 0$, $c = 0$. Then we have $\beta \neq 0$, hence $d \neq 0$ and from (3. 34)

$$\chi_{23} = \chi_{31} = \chi_{12} = 0$$

hence $\chi = \Sigma f^i(\varrho_i)$. Substituting this into (3. 30) we have

$$\sum_{1,2,3} \frac{\partial f^1}{\partial\varrho_1}\frac{\partial f^2}{\partial\varrho_2} = d \neq 0.$$

Thus the f^i-s are linear functions and the φ^i-s occurring in (3. 1) are constants. This case was discussed in Section 2.

Case 2. 2. 1: $D = 0$, $\beta = 0$, $\alpha \neq 0$. Supposing $c \neq 0$ we have by (3. 33) that $\chi = \text{const.}$, and by (3. 22) that $c = g'(v) = 0$, a contradiction. Thus $c = 0$. Further, from (3. 31) and (3. 32)

$$\chi_1\chi_{23} = \chi_2\chi_{31} = \chi_3\chi_{12}. \tag{3.37}$$

Eliminating from (3. 24_1) the quantities χ_{21}, χ_{31} we have by the aid of (3. 37)

$$\begin{vmatrix} \chi_2 & \chi_3 \\ \sigma_2 & \sigma_3 \end{vmatrix} (\chi_2\chi_3\chi_{11} - \chi_1\chi_1\chi_{23}) = 0.$$

Supposing that the first factor vanishes, it is easily shown that the only common solution of the resulting differential equation $\chi_2\sigma_3 = \chi_3\sigma_2$ and $\chi_2\chi_{31} = \chi_3\chi_{12}$ is the trivial solution $\chi = 0$. So we consider only the case $\chi_2\chi_3\chi_{11} - \chi_1\chi_1\chi_{23} = 0$ (cycl.) which together with (3. 37) yields

$$\chi_3\chi_{11} = \chi_1\chi_{31}, \quad \chi_2\chi_{11} = \chi_1\chi_{12} \qquad \text{(cycl.)}.$$

These formulas show that the direction of grad χ is independent of the place in the space $\varrho_1, \varrho_2, \varrho_3$ or

$$\chi = f(\Sigma\vartheta^i\varrho_i), \quad \vartheta^i = \text{const.}$$

If χ is not a constant, then (3. 30) is satisfied in the case $d=0$ if $\Sigma\vartheta_1\vartheta_2 = 0$ and in the case $d \neq 0$ if f is a linear function of its argument, a case encountered earlier.

The remaining cases are

Case 2. 2. 2: $D=0$, $\beta=0$, $\alpha=0$ and

Case 2. 2. 3: $D=0$, $\beta\neq 0$.

In both of these cases we have

$$\text{(3. 38)} \qquad \frac{\alpha^2}{2} + v\beta = 2(\Sigma\varrho_i\chi_i)^2 + 4v(d - c\chi) = 0,$$

an additional first order equation the function χ is to satisfy. Considering also equations (3. 24_j) and (3. 22) it can be shown that the supposition $\chi \neq$ const. leads to a contradiction, hence we have χ = const. This leads by (3. 18) to $\varphi^j = 0$ ($j = 1, 2, 3$), a special case of which was already treated in Section 2.

4. We return now to case 2. 1. 1 of Section 3, where we have got that $\chi = -\frac{1}{2}(\varepsilon_0 + \Sigma\varepsilon_i\varrho_i)^2 + \text{const.}$, hence by (3. 9), (3. 18), by $g'(v) = c$ and by (3. 19)

$$\text{(4. 1)} \qquad \varphi^1 = -(\varepsilon_2 + \varepsilon_3)(\varepsilon_0 + \Sigma\varepsilon_i\varrho_i) + c\varrho_1 \qquad \text{(cycl.)},$$

where $c = \Sigma\varepsilon_1\varepsilon_2$. We are going to show that in this case the system (3. 1) has indeed a non-trivial solution, analytic in the whole space and vanishing at infinity provided the ε's satisfy some restrictions specified later.

Equation (3. 10) is by virtue of (3. 9) now equivalent to

$$\text{(4. 2)} \qquad \sum_{1,2,3} (\varepsilon_2 + \varepsilon_3) N_1 \psi = 0$$

and this in turn is the same as equation (2. 3) by putting $2\lambda_1 = \varepsilon_2 + \varepsilon_3$ (cycl.). Hence the general solution of (4. 2) is of the form $f(u, v)$ where u and v have the same meaning as in (2. 5) and (2. 6). We remark that

$$\text{(4. 3)} \qquad \varphi^1 = 2\lambda_1(\varepsilon_0 + u) + \Lambda\varrho_1 \qquad \text{(cycl.)}$$

where Λ is given by (2. 8) and $\Lambda = c$. Further, substituting $\psi = f(u, v)$ into the equations (3. 1) — where φ^i is given now by (4. 3) — we have as in Section 2 that $f(u, v)$

this time too is to satisfy but two distinct second order equations, namely

$$2uf_{uu}+4vf_{uv}+3f_u-\frac{1}{2}(\varepsilon_0+u)f=0 \tag{4.4}$$

and

$$-\frac{\Lambda}{4}f_{uu}+vf_{vv}+f_v+\frac{\Lambda}{16}f=0. \tag{4.5}$$

Let us again assume that $f(u, v)$ can be expanded into a series of ascending powers of v, this time in the form

$$f(u, v)=\sum_{k=0}^{\infty}\left(\frac{\Lambda}{4}\right)^k e^{-u/2}U_k(u)v^k$$

and let us substitute this expression into (4. 4) and (4. 5). Equating the coefficients of v^k to zero we obtain from (4. 4) and (4. 5)

$$uU_k''+\left(\frac{4k+3}{2}-u\right)U_k'-\frac{1}{4}(4k+3+\varepsilon_0)U_k=0, \tag{4.6}$$

$$(k+1)^2U_{k+1}=U_k''-U_k', \tag{4.7}$$

respectively.

We recall now that if n is a non-negative integer then the Laguerre polynomial $L_n^\alpha(x)$ is defined by

$$L_n^\alpha(x)=\sum_{\nu=0}^{n}\binom{n+\alpha}{n-\nu}\frac{(-x)^\nu}{\nu!} \tag{4.8}$$

[4, formula 5. 1. 6] or [1, formula 10. 12 (7)].

Let now be

$$-\frac{\varepsilon_0+3}{4}=n. \tag{4.9}$$

Then it is easily verified by [4, formulae 5. 1. 2, 5. 1. 13, 5. 1. 14] or [1, formulae 10. 12 (7) and (15)] that the functions

$$U_k=\begin{cases}k!^{-2}L_{n-k}^{2k+(1/2)}(u) & (0\leq k\leq n)\\ 0 & (k>n)\end{cases} \tag{4.10}$$

satisfy the system (4. 6), (4. 7). Hence $f(u, v)$ satisfies the system $(\Delta^i+\varphi^i)\psi=0$ with φ^i given by (4. 1) and the Schrödinger equation

$$(\Sigma\mu_i\Delta^i+\Sigma\mu_i\varphi^i)\psi=0,$$

too.

It remains to investigate the behaviour of these solutions at infinity. If we require that the solutions should vanish there, then the inequality

$$u=\Sigma\varepsilon_i r_i^2>0 \tag{4.11}$$

should be satisfied for all non-negative quantities r_1, r_2, r_3 satisfying the inequalities

$$(r_1-r_2)^2 \leqq r_3^2 \leqq (r_1+r_2)^2; \quad r_1+r_2>0 \qquad \text{(cycl.)}.$$

Since u is a linear function of r_3^2 it is necessary and sufficient that (4. 11) should be valid in the two extreme cases $r_3^2=(r_1+r_2)^2$ and $r_3^2=(r_1-r_2)^2$:

$$(\varepsilon_1+\varepsilon_3)r_1^2 \pm 2\varepsilon_3 r_1 r_2 + (\varepsilon_2+\varepsilon_3)r_2^2 > 0,$$

for $r_1 \geqq 0$, $r_2 \geqq 0$, $r_1+r_2>0$, or

$$(\varepsilon_1+\varepsilon_3)x^2+2\varepsilon_3 x+(\varepsilon_2+\varepsilon_3)>0$$

for all real x's. So (4. 11) holds if and only if

$$\Lambda = \sum_{1,2,3} \varepsilon_1\varepsilon_2 > 0, \qquad \varepsilon_1+\varepsilon_2>0 \qquad \text{(cycl.)}$$

in other words if the positive quantities $\sqrt{\lambda_1}, \sqrt{\lambda_2}, \sqrt{\lambda_3}$ are sides of a triangle having a positive area.

Finally we remark that the functions $f(u, v)$ defined above are not a complete orthogonal system of eigenfunctions of the problem of the three-body harmonic oscillator.

5. In the physically interesting cases of the equation $\Sigma\mu_i(\Delta^i+\varphi^i)=0$ the function $\Sigma\mu_i\varphi^i$ is the sum of three functions $U^i(\varrho_i)$ each depending on one variable only:

(5. 1) $$\sum \mu_i\varphi^i = \sum U^i(\varrho_i).$$

According to (3. 9) this is a differential equation for the function φ:

(5. 2) $$\sum_{1,2,3} (\mu_2+\mu_3)\varphi_1 = \sum U^i(\varrho_i).$$

Let $V^i(\varrho_i)$ be a primitive function of $U^i(\varrho_i)$ and let $\tilde\varphi$ be defined by

(5. 3) $$\varphi = \tilde\varphi + \sum_{1,2,3} \frac{V^1(\varrho_1)}{\mu_2+\mu_3},$$

further let be

(5. 4) $$\tilde\varphi^1 = \tilde\varphi_1+\tilde\varphi_2 \qquad \text{(cycl.)}.$$

Substituting (5. 3) into (5. 2) and (3. 11) we have

(5. 5) $$\sum (\mu_2+\mu_3)\tilde\varphi_1 = 0$$

and

(5. 6) $$\varrho_1^{-1}(\tilde\varphi_{12}^1+\tilde\varphi_{23}^1+\tilde\varphi_{31}^1) = \varrho_2^{-1}(\tilde\varphi_{12}^2+\tilde\varphi_{23}^2+\tilde\varphi_{31}^2) \qquad \text{(cycl.)},$$

respectively. The general solution of (5. 5) is $\tilde\varphi = F(p, q)$, where $p=\mu_3\sigma_1-\mu_1\sigma_3$, $q=\mu_3\sigma_2-\mu_2\sigma_3$. From the two distinct equations (5. 6) it follows that

(5. 7) $$\sum \tilde\varphi_{12}^i = 0 \qquad (i=1, 2, 3).$$

Indeed if it were not so, then from (5. 6) we should have, say,

$$\frac{\varrho_1}{\varrho_2} = \frac{\tilde\varphi_{12}^1+\tilde\varphi_{23}^1+\tilde\varphi_{31}^1}{\tilde\varphi_{12}^2+\tilde\varphi_{23}^2+\tilde\varphi_{31}^2} = G(p, q),$$

hence the Jacobian

$$\frac{\partial(\varrho_1/\varrho_2, p, q)}{\partial(\varrho_1, \varrho_2, \varrho_3)} = \frac{2\mu_3}{\varrho_2^2}[\varrho_2(\mu_2+\mu_3)-\varrho_1(\mu_3+\mu_1)]$$

of the three functions ϱ_1/ϱ_2, p, q would identically vanish. Since this is not true, (5. 7) follows and from these equations we have $\Sigma\tilde{\varphi}_{12}=\gamma=\text{const.}$, or in the special case $\gamma=0$ (putting $M = \mu_1\mu_2+\mu_2\mu_3+\mu_3\mu_1$)

$$(\mu_1+\mu_3)^2\frac{\partial^2 F}{\partial p^2}+2(M-\mu_3^2)\frac{\partial^2 F}{\partial p\,\partial q}+(\mu_2+\mu_3)^2\frac{\partial^2 F}{\partial q^2} = 0.$$

The general solution of this equation can be found by standard methods. If $M \neq 0$, it is of the form $F_+(x+iy)+F_-(x-iy)$ where the real quantities x and y are given e.g. by

$$x+iy = \frac{(\mu_2+iM^{1/2})^2\sigma_1+(\mu_1-iM^{1/2})^2\sigma_2+(\mu_1+\mu_2)^2\sigma_3}{2(\mu_1+\mu_2)M}.$$

Let us introduce two more functions, namely

$$z = \left(-\frac{v}{M}\right)^{1/2}$$

where v is given by (2. 6) and

$$r = \frac{\sum \mu_i\varrho_i}{M}.$$

Then we have the relations

(5. 8) $$x^2+y^2+z^2 = r^2$$

and — introducing the new variables x, y, z instead of ϱ_1, ϱ_2, ϱ_3 —

(5. 9) $$\sum \mu_i \Delta^i = r\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}+\frac{1}{z}\frac{\partial}{\partial z}\right)$$

generalizing thus formulae given by T. H. GRONWALL in the case $\mu_1=\mu_2=1$, $\mu_3=0$. In particular if we consider the Schrödinger equation

(5. 10) $$\left(\sum \mu_i\Delta^i+E+\sum c_i\varrho_i\right)\psi = 0$$

of the three-body harmonic oscillator containing 7 parameters, then by introducing the variables x, y, z instead of ϱ_1, ϱ_2, ϱ_3 it will be transformed into

$$\left\{r\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)+\frac{r}{z}\frac{\partial}{\partial z}+E+\alpha x+\beta y+\gamma r\right\}\psi = 0$$

where α, β, γ are constants. By leaving the z axis fixed and rotating the x and y axes, further by an appropriate change of the unity of length, the quantity $E+\alpha x+$ $+\beta y+\gamma r$ in the last equation can be replaced by $1+\alpha x+\gamma r$, say, reducing thus he constants of the differential equation (5. 10) to two.

REFERENCES

[1] ERDÉLYI, MAGNUS, OBERHETTINGER, TRICOMI: *Higher transcendental functions*, vol. 2 (McGraw—Hill, New York, Toronto, London, 1953).
[2] GRONWALL, T. H.: A special conformally Euclidean space of three dimensions, *Annals of Math*. **33** (1932), 279—293.
[3] GRONWALL, T. H.: The helium wave equation, *Physical Review* **51** (1937) 655—660.
[4] SZEGŐ, G.: *Orthogonal Polynomials*, 2nd edition (American Math. Soc. Colloquium Publications, New York, 1959).

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received November 20, 1970)

INDEX

Printed in Hungary

A kiadásért felel az Akadémiai Kiadó igazgatója – Műszaki szerkesztő: Várhelyi Tamás
A kézirat a nyomdába érkezett: 1971. III. 28. – Terjedelem: 21 (A/5) ív, 35 ábra

71-1655 – Szegedi Nyomda

Die *Studia Scientiarum Mathematicarum Hungarica* ist eine Halbjahrsschrift der Ungarischen Akademie der Wissenschaften. Sie veröffentlicht Originalbeiträge aus dem Bereiche der Mathematik in deutscher, englischer, französischer oder russischer Sprache. Es erscheint jährlich ein Band.

Adresse der Schriftleitung: Budapest V., Reáltanoda u. 13—15, Ungarn.
Technischer Redaktor: Gy. Petruska

Abonnementspreis pro Band (pro Jahr): $ 16.00. Bestellbar beim Buch- und Zeitungs-Aussenhandelsunternehmen *Kultúra* (Budapest 62, P. O. B. 149), oder bei den Vertretungen im Ausland.

Austauschabmachungen können mit der Bibliothek des Mathematischen Instituts (Budapest V., Reáltanoda u. 13—15) getroffen werden.

Die zur Veröffentlichung bestimmten Manuskripte sind in zwei Exemplaren an die Redaktion zu schicken.

Studia Scientiarum Mathematicarum Hungarica est une revue biannuelle de l'Académie Hongroise des Sciences publiant des essais originaux, en français, anglais, allemand ou russe, du domaine des mathématiques.

Rédaction: 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 Mathématiques (Budapest V., Reáltanoda u. 13—15).

On est prié d'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 английском, немецком, французском и русском языках. Отдельные выпуски составляют ежегодно один том.

Адрес редакции: Budapest V., Reáltanoda u. 13—15, Венгрия.
Технический редактор: Gy. Petruska.

Подпусная цена на год (за один том): $ 16.00. Подписка на журнал принимается Внешнеторговым предприятием „Культура" (Budapest 62, P. O. B. 149) или его представительствами за границей.

По поводу отношения обмена просим обращаться к Библиотеке Института Математики (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
Sidnev

AUSTRIA

Globus Buchvertrieb
Salzgries 16
Wien I.

BELGIUM

Office International de Libirairie
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 Smečkách 30
Praha II.
Poštová Novinová Sluzbá
Dovoz tisku
Vinohradská 46
Praha 2
Poštová Novinová Sluzbá
Dovoz tlače
Leningradska 14
Bratislava

DENMARK

Ejnar Munksgaard
Nörregade 6
Kopenhagen

FINLAND

Akateeminen Kirjakauppa
Keskuskatu 2
Helsinki

FRANCE

Office International de Documentation de Libraire
48, rue Gay Lussac
Paris 5

GERMAN DEMOCRATIC REPUBLIC

Deutscher Buch-Export und Impor
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 Streetf
London W. C. I.
Robert Maxwell and Co. Ltd.
4—5 Fitzroy Square
London W. 1

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 1374f
Bombay I.

ITALY

Santo Vanasia
71 Via M. Macchi
Milano
Libreria Commissionaria Sansoni
Via La Marmora 45
Firenze

JAPAN

Nauka Ltd.
2 Kanada-Zimbocho 2-chome
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 Grundt 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

28. IV. 1972. **Index: 26 748**