UDK 001 YU ISSN 0350-2910

Serbian Scientific Society

SCIENTIFIC REVIEW

Series:
Science and Engineering

19-20

Belgrade
1996

Publications of Serbian Scientific Society
Publikacije Naučnog društva Srbije

Series: Science and Engineering | **Serija: Prirodno-matematičke i tehničke nauke**

Scientific Review Office: Naučno društvo Srbije, Šafarikova 7, 11000 Beograd, Tel/Fax: 381 11 3224939

UDK 001 YU ISSN 0350-2910

Serbian Scientific Society

Научно друштво Србије

SCIENTIFIC REVIEW

Series:
Science and Engineering

19-20

Guest Editors

Vera Kovačević - Vujčić
Žarko Mijajlović

Belgrade
1996

Serbian Scientific Review is supported by:

- Ministry for Science and Technology of Serbia
- Federal Ministry for Development, Science and Ecology

EDITORIAL

An international mathematical symposium dedicated to the memory of academician Đuro Kurepa was held at the University of Belgrade, Yugoslavia, from 27-28 May, 1996. The Symposium was organized by the Serbian Scientific Society, the Faculty of Mathematics, the Mathematical Institute of the Serbian Academy of Science and Art, and Union of Mathematicians Societies of Yugoslavia. The joint Program and Organizing Committee was: S. Todorčević (chairman), D. Adamović, D. Adnađević, D. Cvetković, Lj. Ćirić, S. Dajović, Đ. Dugošija, N. Đuranović-Miličić, O. Hadžić, A. Ivić, Lj. Kočinac, V. Kovačević-Vujčić, Z. Marković, V. Mićić, Ž. Mijajlović, V. Perić, S. Perović, J. Petrić, Z. Šami, M. Tasković.

The ideas of Đuro Kurepa (1907-1993) have had a significant influence on modern mathematics, notably in set theory and foundations. The most well known are his works on ramified sets and set-theoretical trees. His considerations of various properties of trees (e.g. Suslin's Property) and his claims of the same to be postulates of the set theory were proved considerably later, with the advent of powerful new methods, through the works of Jech, Silver, Solovay, Tennenbaum, Todorčević and others. The purpose of the Symposium was to bring together researchers from the mathematical areas in which Kurepa worked. So, contributions were in the set theory, foundations, general topology, the number theory and other fields related to Kurepa's work. Sixty-eight papers were presented, and after a refereeing process thirty-one of them were accepted for publication in the Scientific Review.

The guest editors wish to express their gratitude to the Editorial Board of Scientific Review for the support in preparation of this issue and also to Mrs Alice Tošić for language editing, Mrs Nedeljka Vojnović for efficient secretarial assistance and to Mr Miroslav Živković for text processing.

Ž. M. and V.K.-V.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 1-5

ĐURO KUREPA
1907 – 1993

Đuro Kurepa was born on August 16, 1907 in Majske Poljane near Glina in Srpska Krajina as the fourteenth and last child of Rade and Anđelija Kurepa. He got his diploma in theoretical mathematics and physics at the Faculty of Philosophy of the University of Zagreb in 1931. Kurepa spent the years 1932-1935 in Paris at the Faculté des Sciences and the Collège de France. He obtained his doctoral diploma at the Sorbonne in 1935 before a committee whose members were Paul Montel, Maurice Fréchet and Arnaud Denjoy. He received his post-doctoral education at some of the world's best institutions: the University of Warsaw and the University of Paris, and after the Second World War he visited Cambridge (Massachusetts), the mathematical departments of the Universities of Chicago, Berkeley and Los Angeles, and the Institute of Advanced Studies in Princeton.

Kurepa's first employment was at the University of Zagreb in 1931, as an assistant in mathematics. He stayed in Zagreb as a professor until 1965 when he moved to Belgrade at the Faculty of Science. He remained there until his retirement in 1977. Meanwhile, he was a visiting professor at Columbia University in New York, and Boulder, Colorado. Besides his university teaching, Kurepa organized successfully scientific work, too. Professor Kurepa was the chairman of the Mathematical Department of the Faculty of Philosophy in Zagreb; then since 1970 till 1980 chairman of the Mathematical Seminar of the Institute of Mathematics of the Serbian Academy of Sciences and Arts. He was a full member of this Academy, the Academy of Science of Bosnia and Herzegovina, and a corresponding member of Yugoslav Academy of Sciences and Arts in Zagreb.

Professor Kurepa was the founder and president of the Society of Mathematicians and Physicists of Croatia, and president of the Union of Yugoslav Societies of Mathematicians, Physicists and Astronomers. He was also president of the Yugoslav National Committee for Mathematics, as well as president of the Balkan Mathematical Society. Furthermore, he was the founder and for many years the chief editor of the scientific mathematical journal *Mathematica Balkanica*, now published in Sofia. Kurepa was also a member of the editorial board of Belgrade's *Publications de l'Institut Mathématique*, and the German journal *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*.

Professor Kurepa earned many awards, honors and distinctions. He received the highest prize of former Yugoslavia, the Award of AVNOJ (1976). Also, he was a member of the Tesla Memorial Society of the U.S.A. and Canada (1982), the Bernhard Bolzano Charter, and the Gramata Marin Drinov of the Bulgarian Academy of Science (Sofia 1987).

The scientific output of Professor Kurepa was rather big. He published about 200 scientific papers, and more than 700 writings: books, articles, reviews. His papers were published in journals all around the world, and some of them in the most recognized mathematical journals. He delivered scientific talks at many universities of Europe, America and Asia; for example, at Warsaw, Paris, Moscow, Jerusalem, Istanbul, Cambridge, Boston, Chicago, Berkeley, Princeton and Peking. As Kurepa himself told once: "I lectured at almost each of nineteen universities of (former) Yugoslavia, then in almost every European country, then in Canada, Cuba, Israel and Iraq, and I gave at least ten lectures in each of the following countries: France, Italy, Germany, the Soviet Union and the United States." He participated at dozens of international scientific symposia, and many of them were organized by himself.

The influence of Professor Kurepa on the development of mathematics in Yugoslavia was great. As a professor of the University of Zagreb he introduced several mathematical disciplines, mainly concerning the foundations of mathematics and set theory. This is best witnessed by the following words of Kajetan Šeper, a professor of the University of Zagreb:

"Professor Kurepa was not only the professional mathematician and teacher, but he was a scientist, philosopher and humanist as well, in the true sense of these words. He was the founder and pioneer in mathematical logic and the foundations of mathematics in Croatia, and modern mathematical theories in Croatia and Yugoslavia. Generally speaking, he was the catalyzer, the initiator and the bearer of mathematical science.

His arrival to Belgrade in the mid-sixties, and the subsequent influence he had on the mathematical community there, may be described with almost the same words. Professor Kurepa exposed the newest results in diverse mathematical disciplines through many seminars, courses and talks which he delivered at the Faculty of Sciences and the Mathematical Institute. The topics of his lectures included: the construction of Cohen forcing, some questions concerning independence results in cardinal and ordinal arithmetic, ordered sets and general topology. But he was attracted to other mathematical topics, too. He gave valuable contributions to analysis, algebra, number theory, and even to those mathematical disciplines which were just appearing, as computer science, for example. The universality of his spirit is portrayed by the list of university courses he taught: Algebra, Analysis and Topology.

By publishing his doctoral dissertation *in extenso* in Belgrade's *Publications Mathématique de l'Université de Belgrade* in 1935, Kurepa made a first contact with the Belgrade mathematical community. In the beginning of the fifties these contacts became deeper and more frequent. Kurepa was invited already in 1952 to visit the University of Belgrade. On this occasion he gave talks on the theory of matrices, and held a seminar with topics in set theory, topology and algebra. By attending these seminars, many mathematicians gained ideas for their mathematical papers, while graduate students obtained themes for their master and doctoral theses. These works includes virtually all doctoral theses of the older generation of topologists, and many algebraists from all over Yugoslavia: Svetozar Kurepa, Zlatko Mamuzić, Sibe Mardešić, Pavle Papić, Viktor Sedmak, and a few years later, Ljubomir Ćirić, Rade Dacić, Milosav Marjanović, Veljko Perić, Milan Popadić, Ernest Stipanić and Mirko Stojaković. Professor Kurepa was supervisor (altogether 42 times) or member of examination boards for doctoral dissertations for many other mathematicians: Many of these mathematicians continued and developed further Kurepa's work, notably Stevo Todorčević

Professor Kurepa had contacts with many mathematicians of the highest rank from all around the world. Thanks to him some of these mathematicians visited Belgrade: A. Tarski, P. Alexandroff, M. Krasner, N.A. Shanin, K. Devlin and others. Professor Kurepa especially was proud of his encounter with Nikola Tesla, the great Serbian scientist and engineer, with whom Kurepa was fascinated.

Let me say few words about Kurepa's work in topology, set theory and number theory.

In topology Kurepa was interested in some generalizations (non-numerical) of distance functions. In this context there is a notion of Kurepa's pseudo-metric spaces. As it was already mentioned, in the middle of thirties Kurepa was a doctoral student in Paris, and in that time he was influenced by the French mathematical school, particularly by the work of M. Fréchet.

Along this, Kurepa took a new way to the notion of space. He defined the notion of pseudo-distancial space generalizing a class of Fréchet's spaces. In this approach, values of distance function range in a totally ordered set, instead of the set of positive reals, and the triangle condition on distance function is replaced by an interesting relation in ordered sets. Later Frechet came to the same notion and since then this class of abstract spaces are known under the name of "Kurepa-Fréchet spaces". It is interesting that Kurepa wrote last time about these spaces in 1992. Someone probably would recognize in this class of spaces the notion of Zadeh's fuzzy sets. Anyway, the phrase "sets of fuzzy structure" is contained in his book *Set theory* from 1951.

Trees, partially ordered sets in which every lower cone is a well-ordered set, may be considered as a natural generalization of ordinal numbers. They are special type of *ramified sets* which Kurepa introduced in his doctoral thesis. By widespread opinion, this capital work is a first systematic study on set-theoretical trees. In his thesis, and later in his papers, Kurepa introduced fundamental notions from the theory of infinite trees: Aronszajn tree, Suslin tree and Kurepa tree. Kurepa proved many interesting properties concerning these objects. Probably the best known is the following equivalent with the notable Suslin hypothesis:

$$\text{SH} \Leftrightarrow \textit{There is no Suslin tree.}$$

Suslin hypothesis says that there is no Suslin line, i.e. a linearly ordered set of the countable celularity, but which does not have a countable dense subset.

Lebesgue in his paper in 1905 identified implicitly analytic functions with Bair functions. In his proof he used an argument which was "simple and short, but wrong". The mistaken step in the proof was hidden in the proof of a lemma which he considered as trivial, namely that a projection of a Borel set is also a Borel set. Ten years later, Suslin, a young and talented Luzin's student discovered the mistake. Suslin introduced the notion of analytic sets, as projections of Borel sets, and he proved that there are analytic sets that are not Borel sets. So emerged Descriptive set theory, one of the deepest and most interesting parts of set theory. However, Suslin died soon (in 1919), and the formulation of Suslin hypothesis appear after his death in his paper a year later. This hypothesis will play the central role in the development of the theory of infinite trees, and in this progress Kurepa's work was of the principal importance. Namely, Kurepa was trying since 1935 to solve SH. He did not succeed, simply it was not possible to solve it in this time. Tools of the classical set theory were not adequate.

However, Kurepa was the first who understood the importance of trees in set theory. Using infinite trees, Kurepa also found examples of topological spaces with important and unusual properties.

Kurepa was not able to prove the existence or non existence of Suslin tree, neither of Kurepa tree. The postulate that there is a Kurepa tree was named Kurepa hypothesis, shortly KH. The complete solution of the problem of the existence of these trees was solved in the beginning of seventies, when the new method, Cohen's forcing, became a standard and prime tool in set theory. So Solovay, Tennenbaum and Jensen prove that SH is independent from ZFC+CH (Zermelo-Fraenkel set theory plus Continuum Hypothesis), while Devlin proved in 1978 that all logical combinations of CH, SH and KH are consistent with ZFC.

Kurepa had a distinguished ability to sense a good problem and a fine construction, especially if they are connected to ordered sets. We cannot mention all examples of this kind, but one problem from number theory deserves a special attention, as it was considered by several Yugoslav mathematicians and mathematicians from abroad. Kurepa formulated during a mathematical gathering in Ohrid in 1971 the following problem. First he defined an arithmetical function which he called "the left factorial function" as a sum of factorials of first $n-1$ non-negative integers.

Then the formulation of $!n$-hypothesis is stated as follows: The greatest common divisor of for $!n$ and $n!$ is 2. This hypothesis has a lot of interesting equivalences, and it was considered by many mathematicians. This hypothesis is stated in the book "Unsolved problems in number theory" by R. Guy under the number B44. The hypothesis was checked recently by use of computers for $n < 8.400.000$ Kurepa announced the solution (he ringed me up one early morning in the spring of 1992 to tell me this), but he never published the solution. R. Guy in a letter to me in 1991 mentioned that R. Bond from G. Britain might solved the left factorial hypothesis, but this proof did not appear either up to now.

Kurepa was attracted with many areas in mathematics, besides Set Theory, General Topology, Foundations, and Number Theory. His work include also themes in algebra (theory of matrices), numerical mathematics, computer science and fixed-point theory. Is is not possible to discuss here his full mathematical achievements, but in short we may say: Đuro Kurepa has great merits for the development of the foundations of set theory and mathematics in general.

Ž. Mijajlović
Faculty of Mathematics
University of Belgrade

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 7-13

SOUSLINEAN SPACES AND TYCHONOFF CUBES †

Stevo TODORČEVIĆ
Institute of Mathematics, Serbian Academy of Sciencie and Art
Kneza Mihaila 35, 11000 Belgrade, Yugoslavia
Department of Mathematics, University of Toronto,
Toronto, Canada M5S 3G3

This is a short survey of results that spread from the famous problem proposed by Michael Souslin in 1920 (see [23]):

(Souslin's Conjecture) The unit interval is the only ordered continuum which contains no uncountable family of pairwise disjoint open subintervals.

The first advance on this problem was made by Đ. Kurepa [13] when he showed that Souslin's Conjecture is actually equivalent to the following statement which deals with objects of a quite different sort from those appearing in continua theory:

(Kurepa's formulation) Every partially ordered set P satisfies one of the following four conditions:

(1) P is countable,

(2) P contains an uncountable chain,

(3) P contains an uncountable antichain,

(4) P contains the four-element poset $\diamond$

There are two major lines of investigations of the Souslin problem or problems motivated by it. The first line, started by Kurepa himself building on

†Research supported by NSERC and SFS.

his formulation of the Souslin problem, is purely set-theoretical in nature. It starts with the introduction of Souslin trees, the well-founded posets with minimal elements (roots) violating all four properties (1)–(4) above. These Souslin trees have received a considerable amount of attention throughout the years, and especially after the invention of the *Method of Forcing* by Paul Cohen in the early 60's ([2], [3]). Shortly afterwards, S. Tennenbaum [24] showed that there is a Cohen extension of the set-theoretical universe containing a Souslin tree. Independently, T. Jech [7] constructed another Cohen extension with the same property. Not long after that, R. Jensen [8] showed that Souslin trees exist even in the smallest set-theoretical sub-universe, the universe of constructible sets of Gödel [5]. This showed that Cohen's method of forcing was after all not needed if one just wants to produce a model with Souslin's trees, the model existed already in 1940. It turns out, however, that Cohen's method was quite essential in producing a model in which Souslin's Conjecture is true i.e., in showing its full independence from the ordinary axioms of set theory. In fact, Solovay and Tennenbaum [22] had to invent a considerable extension of the original method of Cohen, i.e., they had to invent the so-called *Iterated Forcing*. It should be noted that Kurepa's Souslin trees show up also in the Solovay–Tennenbaum construction, but now as forcing-notions to be iterated. A number of people including D.A. Martin noted that the Solovay–Tennenbaum construction gives a model in which a considerable extension of Souslin's Conjecture is true. This is the so-called *Martin's axiom* which states that in the class of forcing-notions satisfying the *Souslin condition* (every family of pairwise incompatible requirements must be countable) one can always have filters which are *generic* in some restricted sense, i.e., which meet families of dense open sets whenever they are not too large, e.g. if they have size smaller than continuum, or size not bigger than the first uncountable cardinal corresponding to two of the most frequently used versions of the axiom. (Our references to this axiom below will all concern the latter version.) This started an era of the so-called *Forcing Axioms*, an era of unifying various Forcing-constructions. Today, this represents one of the most vital fronts of research in this field of mathematics.

The second line of investigation is topological in nature and was again started by Kurepa [13] and later joined by Knaster [12], Marczewski (Szpilrajn) [16], Shanin [20], and others. The investigation is based on the observation that the Souslin problem is really a part of a more general problem which asks whether the Souslin condition is as strong as the stronger condition of separability in a given class of spaces. In fact there is a number of different *chain conditions* (see [4]) that one can put on a given compact space X and which in strength lie between separability and the Souslin condition. Here are some of the most prominent ones:

Souslin's condition: Every uncountable family $\mathcal{F}$ of open subsets of X contains two (distinct) sets U and V such that $U \cap V \neq \emptyset$.

Knaster's condition: Every uncountable family $\mathcal{F}$ of open subsets of X contains an uncountable Knaster family $\mathcal{K}$ i.e., an uncountable family $\mathcal{K} \subseteq \mathcal{F}$ such that $U \cap V \neq \emptyset$ for all U and V in $\mathcal{K}$.

Shanin's condition: Every uncountable family $\mathcal{F}$ of open subsets of X contains an uncountable subfamily $\mathcal{S}$ such that $\bigcap \mathcal{S} \neq \emptyset$.

Separability: The family $\mathcal{O}$ of all nonempty open subsets of X can be decomposed into countably many subfamilies $\mathcal{O}_n$ $(n = 1, 2, 3, \ldots)$ such that $\bigcap \mathcal{O}_n \neq \emptyset$ for all n.

The following relationship between these four chain conditions should be clear:

$$\text{Separable} \rightarrow \text{Shanin} \rightarrow \text{Knaster} \rightarrow \text{Souslin}$$

Investigating in which classes of compact spaces we have that some of these chain conditions are equivalent can be considered as a way of investigating the corresponding versions of the Souslin problem. For example, Knaster [12] proved that Knaster's condition is equivalent to the separability in the class of ordered continua, so the original Sousin problem reduces to the question whether Souslin's and Knaster's conditions are equivalent in this class of spaces. Another, interesting line of research was motivated by the following crucial question of Kurepa and Marczewski, a form of which already appears in the Scottish book (see [19; Problem 192]):

(Question of productiveness). Which of these four chain conditions are productive.

Here are some of the answers:

(I) The product of no more than continuum many separable spaces is separable, but no product of more than continuum many nontrivial spaces is separable (Hewitt–Marczewski–Pondiczery; [6], [17], [18]).

(II) Knaster's condition is preserved in products of any number of factors (Marczewski [16]).

(III) Shanin's condition is preserved in products of any number of factors (Shanin [20]).

(IV) If Souslin's condition is productive, then Souslin's conjecture is true (Kurepa [14]).

In order to prove (III) Shanin invented his famous *Delta-system Lemma* (see [21]) which turned out to be useful in many other contexts. For example, this lemma is used crucially in the Solovay–Tennenbaum iterated forcing construction discussed above. The following result can be considered as an explanation of the relationship between these two apparently quite different sets of results:

(V) Martins'a axiom is equivalent to the statement that Souslin's and Shanin's condition are equivalent in the case of all compact Hausdorff spaces (Todorčević- Veličković [25]).

In fact, one can go further and prove the following result which shows that Martin's axiom is nothing more than a Souslin's conjecture for a class of space not much more general than the class of ordered continua.

Theorem 1. *Martin's axiom is equivalent to the statement that every compact first countable Souslin space is separable.*

This leads to a possible line of strengthening the original Souslin's conjecture. To state this let us call (following M. Bell [1]) a compact space X a *Souslinean space* if X is Souslin but not separable. Thus, Souslin 's conjecture states that the pathology of Souslinean spaces does not occur in the class of ordered continua. Note that the pathology does occur in the class of all compact spaces: By the Hewitt–Marczewski -Pondiczery theorem every Tychonoff cube $[0,1]^A$ over an index-set A of size bigger than continuum is a Souslinean space. The fact that $[0,1]^A$ is Souslin (and in fact Shanin) follows from the theorems of Marczewski and Shanin mentioned above (see (II) and (III)). So the following assertion, stating that large Tychonoff cubes are essentially the only obstructions, can be considered as an ultimate form of Souslin's Conjecture.

(A) Every compact Souslinean space maps onto a Tychonoff cube of uncountable weight.

Since no first-countable compact space maps onto an uncountable Tychonoff cube, Theorem 1 tells us that (A) is at least as strong as Martin's axiom. But (A) seems to be quite different from Martin's axiom as the following fact shows (revealing also the close connection between the Souslin Problem and the Continuum Problem).

Theorem 2. *If (A) holds then the minimal size of an unbounded subset of* N^{N} *under the ordering of eventual dominance is equal to the second uncountable cardinal.*

Among the lines of investigations related to the Souslin Problem, we mention the following well-known problem of von Neumann, also appearing in the Scottish book and at about the same time as the problems of Knaster and Marczewski discussed above (see [19; Problem 163]):

(von Neumann's problem) Does weak-distributivity and the Souslin condition characterize measure algebras among all complete boolean algebras?

A *measure algebra* is a complete Boolean algebra A with a σ-additive measure $\mu: A \to [0,1]$ such that $\mu(a) > 0$ for all $a \neq 0$ in A. An algebra A is *weakly-distributive* if for every matrix $a_{nm}(n, m \in \mathrm{N})$ of positive elements of A such that

$$\sum_{m=1}^{\infty} a_{nm} = 1$$

for all n, there exists a sequence $\{k_n\} \subseteq \mathrm{N}$ such that

$$\prod_{n=1}^{\infty} \sum_{m=1}^{k_n} a_{nm} \neq 0$$

This problem was shortly afterwards analyzed by D. Maharam ([15]) who showed that the complete boolean algebra associated to the Souslin tree serves as a counterexample to von Neumann's question. However, the corresponding famous reformulation of von Neumann's question, also due to Maharam ([15]) and known under the name of *Control Measure Problem*, is still widely open. It has been reformulated to a purely combinatorial question about submeasures on the countable free algebra by Kalton and Roberts (see [9], [10]). It is interesting that this process of reformulating followed closely the chain-condition method described above. This is not so surprising if one knows that the existence of a strictly positive finitely additive measure on a boolean algebra A is simply a chain condition that also lies between separability and the Souslin condition. This is a result of J.L. Kelley [11] which says that there is a finitely additive strictly positive measure $\mu : A \to$ [0,1] if and only if there is a decomposition

$$A\backslash\{0\} = \bigcup_{n=1}^{\infty} A_n$$

such that $\mathrm{Int}(A_n) > 0$ for all n. Here, for a given $B \subseteq A\backslash\{0\}$, the *intersection number* $\mathrm{Int}(B)$ is defined to be equal to the infimum of all ratios of the form

$$\frac{\mathrm{cal}(b_1, ..., b_n)}{n}$$

where $(b_1, ..., b_n)$ is a finite sequence of (not necessarily distinct) elements of B and where $\mathrm{cal}(b_1, ..., b_n)$ is the maximal cardinality of a subset I of $\{1, ..., n\}$ such that

$$\prod_{i \in I} b_i \neq 0$$

REFERENCES

[1] M.G. Bell, *A compact ccc non-separable spaces and Martin's axiom*, preprint 1995.

[2] P.J. Cohen, *The independence of the continuum hypothesis*, Proc. Nat. Acad. Sci. U.S.A. 50 (1963), 1143–1148

[3] P.J. Cohen, *The independence of the continuum hypothesis*, Proc. Nat. Acad. Sci. U.S.A. 51 (1964), 105–110

[4] W.W. Comfort and S. Negrepontis, *Chain Conditions in Topology*, Cambridge Univ. Press, 1982.

[5] K. Gödel, *The consistency of the axiom of choice and of the generalized continuum hypothesis*, Annals of Mathematics Studies Vol. 3 (1940).

[6] E. Hewitt, *A remark on density characters*, Bill. Amer. Math. Soc. 52 (1946), 641–643.

[7] T.J. Jech, *Nonprovability of Souslin's hypothesis*, Comment. Math. Univ. Carolinae 8 (1967), 201–305.

[8] R.B. Jensen, *Souslin's hypothesis is incompatible with $V = L$* (Abstract), Notices American Math. Society 15 (1968), A-935.

[9] N.J. Kalton, J.W. Roberts, *Uniformly exhaustive submeasures and nearly additive set functions*, Trans. Amer. Math. Soc. 278 (1983), 803–816.

[10] N.J. Kalton, *The Maharam Problem, Seminaire Initiation á l'Analyse* (G. Choquet, G. Godefroy, M. Rogalski, J. Saint-Raymond), 28e Année, 1988/89 no. 18.

[11] J.L. Kelley, *Measures on boolean algebras*, Pacific J. Math. 9 (1959), 1165–1177.

[12] B. Knaster, *Sur une propriété caractéristique de l'ensemble des nombres réels*, Mat. Sbornik 16 (1945) 281–288.

[13] Đ. Kurepa, *Ensembles ordonnés et ramifiés*, Publ. Math. Univ. Belgrade 4 (1935), 1–138.

[14] Đ. Kurepa, *La condition de Souslin et une propriété caractéristique des nombres réels*, C.R. Acad. Sci. Paris 231 (1950), 1113–1114.

[15] D. Maharam, *An algebraic characterization of measure algebras*, Ann. of Math. (2) 48 (1947), 154–167.

[16] E. Marczewski (Szpilrajn), *Remarque sur les produits cartésiens d'espaces topologiques*, Dokl. Akad. Nauk SSSR 31 (1941), 525–527.

[17] E. Marczewski, *Séparabilité et multiplication cartésienne des espaces topologiques*, Fund Math. 34 (1947), 154–157.

[18] E.S. Pondiczery, *Power problem in abstract spaces*, Duke Math. J. 11 (1944), 835–837.

[19] *The Scottish Book*, (ed. R.D. Mauldin) Birkhauser, Boston, 1981.

[20] N.A. Shanin, *On intersection of open subsets in the product of topological spaces*, Doklady Akad Nauk SSSR. 53 (1946), 449–501.

[21] N.A. Shanin, *A theorem from the general theory of sets*, Doklady Akad Nauk SSSR. 53 (1946) 399–400.

[22] R. Solovay and S. Tennenbaum, *Iterated Cohen extensions and Souslin's problem*, Ann. Math. 94 (1971), 201–245.

[23] M. Souslin, *Problème 3*, Fund Math. 1 (1920), 225.

[24] S. Tennenbaum, *Souslin's problem*, Proc. Nat. Acad. Sci., U.S.A. 59 (1968), 60–63.

[25] S. Todorčević and B. Veličković, *Martin's axiom and partitions*, Compositio Math. 63 (1987), 391–408.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 15-16

THE VERY IDEA OF AN OUTCOME

Nuel BELNAP
Department of Philosophy, University of Pittsburgh,
Pittsburgh PA 15260

Abstract. We define, in branching-time structures, an algebra of "simple outcomes" and an algebra of "outcomes of a set of choice points." The former is boolean, the latter orthomodular.

Keywords: outcome, branching time, orthomodularity

What is an outcome? Outcomes make no sense if our world is fully deterministic. So we work in branching time, the simplest representation of indeterminism. $\langle Tree, \leq\rangle$ is a *branching-time (BT) structure* iff it is a partial ordering with no downward branching and with the existence of $glb(m_1, m_2)$ whenever $m_1, m_2 \in Tree$. Members of $Tree$ are called *moments*. Moments m_1, m_2 are *consistent* if either $m_1 \leq m_2$ or $m_2 \leq m_1$, and otherwise *inconsistent*.

We first define the algebra of simple outcomes, following a chain of ideas from von Neumann via [1]. Define $m_1 \perp m_2$ iff m_1 and m_2 are inconsistent. Then for a set M of moments, define $M^\perp = \{m_1 : \forall m_2[m_2 \in M \to m_1 \perp m_2]\}$. The set $M^\perp$ is the *orthocomplement* of M. Call M a *simple outcome* iff $M = M^{\perp\perp}$. The *algebra of simple outcomes* is given by $\langle \underline{O}, {}^\perp, \wedge, \vee\rangle$ where $\underline{O}$ is the set of simple outcomes, $M^\perp$ is as just defined, $M_1 \wedge M_2$ is the set meet $M_1 \cap M_2$, and $M_1 \vee M_2 = (M_1 \cup M_2)^{\perp\perp}$.

It is shown in [1] that any algebra following out these definitions from an irreflexive and symmetric relation is bound to be an ortholattice. The first result reported here is that the algebra of simple outcomes is boolean, although it is not a set algebra. More interesting is the nonmathematical fact that simple outcomes "look like" outcomes. For instance, if you are not in a simple outcome, then you can always find a path to its orthocomplement– and conversely: the orthocomplement of M is the outcome that has begun to happen exactly when M can no longer begin to happen.

Simple outcomes are "outcomes of something or other." Suppose, however, we fix on a special set E and ask for the "outcomes of E." For this we define $m_1 \perp_E m_2$ by: $m_1 \perp m_2$ and $glb(m_1, m_2) \in E$. So when $m_1 \perp_E m_2$, we know that E contains the exact place of splitting that makes at least one of m_1, m_2 henceforth impossible. Then we follow out the von Neumann chain of definitions exactly, keeping in mind the relativization to E, and ending with $\underline{O}_E = \langle \underline{O}_E, {}^{\perp_E}, \wedge_E, \vee_E \rangle$. This we call the *algebra of outcomes of* E.

Examples show that $\underline{O}_E$ is not boolean. Let e.g. $Tree = \{m_0, m_x, m_y, m_{x+}, m_{x-}, m_{y+}, m_{y-}\}$ begin with m_0, and split so that m_x and m_y are each just above m_0. Let m_x and m_y each split so that m_{x+} and m_{x-} are just above m_x, and m_{y+} and m_{y-} above m_y. Story: the physicist decides at m_0 whether to make the x or the y measurement, which are inconsistent (each can be made but not both). Then the measurements at m_x and m_y have each two possible results, $+$ or $-$. If we are interested in "the outcomes of the measurements" (unmixed with the decisions of the physicist), we let $E = \{m_x, m_y\}$. The resulting lattice of outcomes has just the six elements $\{\emptyset, \{m_{x+}\}, \{m_{x-}\}, \{m_{y+}\}, \{m_{y-}\}, Tree\}$ and is obviously not boolean.

Theorem. *Regardless of the tree structure and choice of* E, $\underline{O}_E$ *is orthomodular: If* $M_1, M_2 \in \underline{O}_E$ *and* $M_1 \subseteq M_2$, *then* $M_1 \vee_E (M_1^{\perp_E} \wedge_E M_2) = M_2$.

This second result is interesting because, while the present very simple intuitive basis involves indeterminism, there is no connection with quantum-mechanical "funny business," to which nonboolean orthomodularity is usually attributed.

Hardly anything additional is known about how the choice of *Tree* and E influences the structure of $\underline{O}_E$; almost all questions are open.

REFERENCES

[1] Goldblatt R.H., *Semantic analysis of orthologic*, Journal of Philosophical Logic 3 (1974), 19–35.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 17-29

A CONJECTURE ABOUT IDENTITY, INTERPOLATION AND THE LENGTH OF PROOFS IN $\mathbf{TW}_{\rightarrow}$–ID[†]

Aleksandar KRON
University of Belgrade

Abstract. It is known that in the system $\mathbf{TW}_{\rightarrow}$–ID of purely implicational relevance logic there is no theorem of the form $(A \rightarrow A)$ (NOID). We conjecture that NOID is equivalent to the following interpolation property: for any theorem $(A \rightarrow C)$ of $\mathbf{TW}_{\rightarrow}$–ID, if there is a formula B such that $(A \rightarrow B)$ and $(B \rightarrow C)$ are theorems of $\mathbf{TW}_{\rightarrow}$–ID, then there is a natural number n such that for any natural number m and any formulas $B_1, \ldots, B_m$, if

$$(A \rightarrow B_1), (B_1 \rightarrow B_2), \ldots, (B_{m-1} \rightarrow B_m), (B_m \rightarrow C)$$

are theorems of $\mathbf{TW}_{\rightarrow}$–ID, then $m \leq n$.

We show that the interpolation property stated above is equivalent to the following provability property: for any formula A, if A is a theorem of $\mathbf{TW}_{\rightarrow}$–ID, then there is a natural number n such that the length of proofs of A in $\mathbf{TW}_{\rightarrow}$–ID is not greater than n.

1. THE SYSTEMS $\mathbf{TW}_{\rightarrow}$–ID, $\mathbf{TRW}_{\rightarrow}$–ID AND $\mathbf{TW}_{\rightarrow}^{01}$–ID

The only connective in the language of $\mathbf{TW}_{\rightarrow}$–ID is $\rightarrow$. If A and B are formulas, so is $(A \rightarrow B)$. We shall write (AB) instead. The parentheses are omitted as usual, with the association to the left. Also, we write $A.BC$ for $A(BC)$.

The only rule of $\mathbf{TW}_{\rightarrow}$–ID is modus ponens (MP).

[†]This research has been supported by the Science Fund of Serbia (grant number 0401A) through the Mathematical Institute in Belgrade.

The axioms of $\mathbf{TW}_{\rightarrow}$–ID are all formulas of the form

ASU $\qquad AB.BC.AC$

APR $\qquad BC.AB.AC$

By a *proof tree* of a formula C in $\mathbf{TW}_{\rightarrow}$–ID we understand a finite tree such that:

(a) C is at the origin of the tree;

(b) if B is a node that is not an end node, then there is an application of MP with the premises A and AB such that A and AB are nodes immediately above B;

(c) if B is an end node, then it is an axiom of $\mathbf{TW}_{\rightarrow}$–ID.

By the *length* of a proof tree of C we understand the number of nodes in the longest branch of the tree.

Theorem 1. *Let there be a proof tree of C of length n, $n > 1$, and let $B_1, \ldots, B_{n-1}, C$ be a branch of length n; then there is a node of this proof of the form $A_n. \ldots .A_1C$ such that $A_1, \ldots, A_n$ are theorems of* $\mathbf{TW}_{\rightarrow}$–ID.

Proof. It is clear that any B_i, $1 \leq i \leq n-1$, is either a major or a minor premise in an application of MP. If B_i is the major premise, then $B_i = B_i'B_{i+1}$, where B_i' is the corresponding minor premise; if B_{n-1} is the major premise, then $B_{n-1} = B_{n-1}'C$.

Let us define a branch $A_{n-1}, \ldots, A_1, C$ of the given proof of C as follows: if B_{n-1} is the major premise in the application of MP, then $A_1 = B_{n-1}$; if B_{n-1} is the minor premise, then A_1 is the corresponding major premise, $A_1 = B_{n-1}C$.

Suppose that $A_k, \ldots, A_1, C$ have been defined; then $A_{k+1} = B_{n-k-1}$, if B_{n-k-1} is the major premise in an application of MP; furthermore, $A_{k+1} = B_{n-k-1}A_k$, if B_{n-k-1} is the minor premise of MP, where, obviously, B_{n-k-1} B_{n-k} is the corresponding major premise.

By induction on k, $1 \leq k < n-1$, it can be shown that $A_{n-1}, \ldots, A_1, C$ is a branch of the given proof of C. Moreover, A_1 is of the form $D_{n-1}C$. Suppose that A_k is of the form $D_kD_{k-1}. \ldots .D_1C$ and let us consider A_{k+1}. If $A_{k+1} = B_{n-k-1}$ is the major premise in an application of MP, then $A_{k+1} = B_{n-k-1}'A_k = B_{n-k-1}'.D_k.D_{k-1} \ldots .D_1C$; if B_{n-k-1} is the minor premise in an application of MP, then A_{k+1} is the corresponding major premise; hence, $A_{k+1} = B_{n-k-1}A_k = B_{n-k-1}.D_k \ldots .D_1C$.

By the construction, $A_1, \ldots, A_n$ are theorems of $\mathbf{TW}_{\rightarrow}$–ID.

This proves the theorem.

There is an alternative formulation of $\mathbf{TW}_{\rightarrow}$-ID called herein $\mathbf{TRW}_{\rightarrow}$-ID. It has the axiom-schemata ASU and APR, but the rules are:

SU From AB to infer $BC.AC$

PR From BC to infer $AB.AC$

TR From AB and BC to infer AC

It is clear that the theorems of $\mathbf{TRW}_{\rightarrow}$-ID are theorems of $\mathbf{TW}_{\rightarrow}$-ID. That the theorems of $\mathbf{TW}_{\rightarrow}$-ID are theorems of $\mathbf{TRW}_{\rightarrow}$-ID follows by an inductive argument showing that $\mathbf{TRW}_{\rightarrow}$-ID is closed under MP.

The proofs of theorems in $\mathbf{TRW}_{\rightarrow}$-ID can be written in a normal form.

Theorem 2. *For any proof of a theorem of* $\mathbf{TRW}_{\rightarrow}$-ID *containing n applications of* TR *there is a proof of the same theorem containing n applications of* TR *such that no application of TR precedes an application of either SU or PR.*

Proof. Suppose that (c) AC is obtained from (a) AB and (b) BC by TR, and that (d) $CD.AD$ is obtained from (c) by SU.

Obviously, we can apply SU to (b) and (a) first, to obtain (b') $CD.BD$ and (a') $BD.AD$, and then to apply TR to (b') and (a') to prove (d).

Suppose that (c) AC is obtained from (a) AB and (b) BC by TR, and that (d) $DA.DC$ is obtained from (c) by PR.

We can apply PR to (a) and (b) first, to obtain (a') $DA.DB$ and (b') $DB.DC$, and then to apply TR to (b') and (a') to prove (d).

It is clear that if the first proof has n applications of TR, so does the second.

Theorem 3. *If (a) AB and (b) BC are proved in* $\mathbf{TRW}_{\rightarrow}$-ID *such that (a) is obtained by an application of SU (of PR) in the last step and (b) is obtained by an application of PR (of SU) in the last step, then there is a proof of AC by TR such that the left premiss (a') in this application of TR is obtained by PR (SU) in the last step and the right premiss (b') is obtained by SU (PR) in the last step.*

Proof. Suppose that (a) is $DE.FE$, obtained from (a') FD by SU, and that (b) is $FE.FG$, obtained from (b') EG by PR. It is clear that from (b') we can derive $DE.DG$ by PR and that from (a') we can prove $DG.FG$ by SU.

Suppose that (a) is $ED.EF$, obtained from (a') DF by PR, and that (b) is $EF.GF$, obtained from (b') GE by SU. Now from (b') we obtain $ED.GD$ by SU and from (a') we can prove $GD.GF$ by PR.

Theorem 4. *If (a) AB is obtained either by SU or by PR and (b) BC is an axiom then there is a proof of AC by TR such that the left premiss (b') of this application of TR is an axiom.*

Proof. Let (a) be $DE.FE$, obtained from (a') FD by SU, and let (b) be the axiom $FE.EG.FG$. We can take (b') to be $DE.EG.DG$, then we apply SU to (a'), to obtain (a") $DG.FG$, and then we use PR to prove (b") $(EG)(DG).EG.FG$.

If (b) is the axiom $FE.GF.GE$, we take (b') to be $DE.GD.GE$, and we apply PR to (a') to obtain $GF.GD$; now we use SU to prove (b") $(GD)(GE).GF.GE$.

Let (a) be $DE.DF$, obtained from EF by PR, and let (b) be the axiom $DF.FG.DG$. We can take (b') $DE.EG.DG$, and apply SU to (a') to obtain $FG.EG$, and then SU to prove (b") $(EG)(DG).FG.DG$.

If (b) is the axiom $DF.GD.GF$, we take (b') to be $DE.GD.GE$, and apply PR to (a') to prove $GE.GF$, and then PR again, to obtain $(GD)(GE).GD.GF$.

This completes the proof of the theorem.

Theorem 5. *If AC is a theorem of* **TRW**$_{\rightarrow}$-ID, *then either AC is an axiom or there are non-negative integers k, m, n and formulas $B_1, \ldots, B_k$, $D_1, \ldots D_m$, and $E_1, \ldots, E_n$ such that at least one of k, m, and n is not 0, and*

$$AB_1, \ldots B_{k-1}B_k, B_kD_1, \ldots, D_{m-1}D_m, D_mE_1, \ldots, E_{n-1}E_n, E_nC$$

are theorems of **TRW**$_{\rightarrow}$-ID *satisfying the following conditions:*

(1) TR is not used in the proof of any of them;

(2) $AB_1, \ldots, B_{k-1}B_k$ are axioms;

(3) $B_kD_1, \ldots, D_{m-1}D_m$ are obtained by an application of SU (of PR) in the last step;

(4) $D_mE_1, \ldots, E_{n-1}E_n, E_nC$ are obtained by an application of PR (of SU) in the last step.

Proof. By Theorems 1 - 3.

The proof of AC in $\mathbf{TRW}_{\rightarrow}$–ID having the form described in Theorem 4 is called *normal*. Hence, for any theorem of $\mathbf{TRW}_{\rightarrow}$–ID there is a normal proof.

The most important theorem of $\mathbf{TW}_{\rightarrow}$–ID and hence of $\mathbf{TRW}_{\rightarrow}$–ID is NOID:

Theorem 6. *There is no theorem of* $\mathbf{TW}_{\rightarrow}$–ID *and hence of* $\mathbf{TRW}_{\rightarrow}$–ID *either of the form AA or of the form ABB or of the form $A.ABB$ or of the form $ABBA$.*

Proof. Cf. [2], [3], [4] and [5].

By a *code* we understand a finite sequence of 0's and 1's. The *empty code* is $\emptyset$. The letters a, b, c, ... range over the set of codes. By the *length* $l(a)$ of a code a we understand the number of 0's and 1's in a. If $a = a_1 \dots a_m$ and $b = b_1 \dots b_n$, then $ab = a_1 \dots a_m b_1 \dots b_n$.

With each occurrence of a subformula B in a formula A we associate a code $f(B, A)$ as follows:

(i) $f(A, A) = 0$;

(ii) if $f(BC, A) = a0$, then $f(B, A) = a01$ and $f(C, A) = a00$;

(iii) if $f(BC, A) = a1$, then $f(B, A) = a10$ and $f(C, A) = a11$.

If the code a is the initial segment of a code b, we write $a \preceq b$; if $a \preceq b$ and $a \neq b$, we write $a \prec b$.

Following [1] we shall call an occurrence of a subformula B in a formula A *antecedent* or *consequent* if $f(B, A) = a1$ or $f(B, A) = a0$, respectively, for some code a.

For any code a we define the code a^{-1}. If $a = \emptyset$, then $a^{-1} = \emptyset$; furthermore, $0^{-1} = 1$, $1^{-1} = 0$ and $(ab)^{-1} = a^{-1}b^{-1}$.

Let us define the *depth* of an occurrence of a subformula B in a formula A as follows: if $f(B, A) = a$, then $l(a) - 1$ is the depth of B in A.

Let B be a subformula of A with a code a in A; then we shall write $A[B, a]$ for A. We have $(A_1A_2)[B, 01a] = A_1[B, 0a^{-1}]A_2$ and $(A_1A_2)[B, 00a] = A_1A_2[B, 0a]$.

If $A = A[B, a]$, we shall write $A[B/C, a]$ for the result of the substitution of an occurrence of C for the particular occurrence of B with the code a in A. Obviously, $A[B/C, a] = C$ if $a = 0$ and $A[B, 0] = B$.

Theorem 7. *Let AB be a theorem of* $\mathbf{TW}_{\rightarrow}$–ID; *then so are*

$$\text{(a) } D[A, a0]D[A/B, a0] \text{ and } \text{(b) } D[A/B, a1]D[A, a1].$$

Proof. By induction on the depth of A in D.

Let the depth of A in D be 0; then $A = D$ and $B = D[A/B, 0]$. Obviously, $D[A, 0]D[A/B, 0] = AB$.

Let $D = D_1D_2$.

If $D[A, 01a0] = D_1[A, 0a^{-1}1]D_2$, then $D_1[A/B, 0a^{-1}1]D_1[A, 0a^{-1}1]$ is a theorem by induction hypothesis. Hence, by ASU and MP we obtain

$$D_1[A, 0a^{-1}1]D_2.D_1[A/B, 0a^{-1}1]D_2,$$

i.e. $D[A, 01a0]D[A/B, 01a0]$.

If $D[A, 00a0] = D_1D_2[A, 0a0]$, then $D_2[A, 0a0]D_2[A/B, 0a0]$ is a theorem of $\mathbf{TW}_{\rightarrow}$–ID by induction hypothesis. Hence, by APR and MP we obtain

$$D_1(D_2[A, 0a0]).D_1(D_2[A/B, 0a0]),$$

i.e. $D[A, 00a0]D[A/B, 00a0]$.

If $D[A, 01a1] = D_1[A, 0a^{-1}0]D_2$, then $D_1[A, 0a^{-1}0]D_1[A/B, 0a^{-1}0]$ is a theorem of $\mathbf{TW}_{\rightarrow}$–ID by induction hypothesis. Hence, by ASU and MP we get

$$D_1[A/B, 0a^{-1}0]D_2.D_1[A, 0a^{-1}0]D_2,$$

i.e. $D[A/B, 01a1]D[A, 01a1]$.

If $D[A, 00a1] = D_1(D_2[A, 0a1])$, then $D_2[A/B, 0a1]D_2[A, 0a1]$ is a theorem by induction hypothesis. Hence, by APR and MP we get

$$D_1(D_2[A/B, 0a1]).D_1(D_2[A, 0a1]),$$

i.e. $D[A/B, 00a1]D[A, 00a1]$

The system $\mathbf{TW}^{01}_{\rightarrow}$–ID was defined by K. Došen and the present author. It has no axioms and the rules are:

SU^0	From $D[AB, a0]$ to infer $D[AB/BC.AC, a0]$
PR^1	From $D[BC, a0]$ to infer $D[BC/AB.AC, a0]$
SU^1	From $D[BC.AC, a1]$ to infer $D[BC.AC/AB, a1]$
PR^1	From $D[AB.AC, a1]$ to infer $D[AB.AC/BC, a1]$

SU^0 and PR^0 are called 0-rules; SU^1 and PR^1 are called 1-rules.

Suppose that we have started applying these rules to a formula A, and that, after a finite nonzero number of applications of these rules we have obtained the formula B; then we shall write $A \longrightarrow B$ to denote both this fact and the derivation of B from A. Also, we shall write $A \longrightarrow B \longrightarrow C$ if $A \longrightarrow B$ and $B \longrightarrow C$.

Let us define the set of theorems of $\mathbf{TW}^{01}_{\rightarrow}$–ID: suppose that $A \longrightarrow B$; then AB is a theorem of $\mathbf{TW}^{01}_{\rightarrow}$–ID.

K. Došen has remarked that in $\mathbf{TW}^{01}_{\rightarrow}$–ID every derivation can be written in a normal form.

Theorem 8. *If DF is a theorem of $\mathbf{TW}^{01}_{\rightarrow}$–ID, then there is a derivation of DF in $\mathbf{TW}^{01}_{\rightarrow}$–ID such that no application of a 1-rule precedes an application of a 0-rule.*

Proof. If an application of a 1-rule precedes an application of a 0-rule, there is the possibility of reversing that order. Let us see how this can be done.

Let SU^1 be applied to E in a derivation of F from D, let E' be the result of this application, and then let a 0-rule (say SU^0) be applied to E' with the result E''. Thus, $E = E[BC.AC, a1]$ and $E' = E[BC.AC/AB, a1]$. If now $HI.GI$ is substituted for a consequent occurrence of GH in E', then this consequent occurrence of GH exists in E as well, and hence SU^0 can be applied to E first, and then we may apply SU^1 to derive E''.

We take care of the remaining rules in a similar way.

Theorem 9. *If $A \longrightarrow B$, then there is a derivation of AB in $\mathbf{TW}^{01}_{\rightarrow}$–ID such that every application of a 0-rule (if any) at depth d precedes all applications of the 0-rules at depth d', $d < d'$.*

Proof. If an application of a 0-rule at depth d' precedes an application of a 0-rule at depth d, $d < d'$, there is the possibility of reversing that order. For example, let SU^0 be applied to E, $E = E[AB, a0]$, at depth d', in a derivation of F from D, let E' be the result of this application, $E' = E[AB/BC.AC, a0]$, and then let SU^0 be applied to E' at depth d, $d < d'$, with the result E'', as follows: let the displayed occurrence of $BC.AC$ in E' be in a consequent occurrence of a subformula GH' of E', say, $H = H[AB, b0]$ and $H' = H[AB/BC.AC, b0]$. Suppose, moreover, that SU^0 is applied to GH' in E' at depth d, $d < d'$ and $b \prec a$, to derive E'', where $E'' = E'[GH'/H'I.GI, b0]$. It is clear that we may apply SU^0 to E at depth

d first, to obtain $E[GH/HI.GI, b0]$, and then SU^0 again, at depth d', to obtain E''.

We take care of the remaining cases in a similar way.

Theorem 10. *If $A \longrightarrow B$, then there is a derivation $A \longrightarrow B$ in $\mathbf{TW}^{01}_{\rightarrow}$–ID such that every application of a 1-rule (if any) at depth d precedes all applications of the 1-rules at depth d', $d > d'$.*

Proof. If an application of a 1-rule at depth d' precedes an application of a 1-rule at depth d, $d > d'$, there is the possibility of reversing that order.

Let SU^1 be applied to E at depth d' (in E) in a derivation of F from D, let E' be the result of this application, and then let a 1-rule be applied to E' at depth d (in E'), $d > d'$, with the result E''. Thus, say, $E = E[BC.AC, a1]$ and $E' = E[BC.AC/AB, a1]$. If now GH is substituted for an antecedent occurrence of $HI.GI$ in E', then this antecedent occurrence of $HI.GI$ exists in E as well, and hence SU^1 can be applied to E first, at depth d, and then we may apply SU^1 at depth d' to derive E''.

We take care of the remaining cases in a similar way.

Let us call a derivation $A \longrightarrow B$ *normal* if

(1) no application of a 1-rule precedes an application of a 0-rule;

(2) no application of a 0-rule at greater depth precedes an application of a 0-rule at a smaller depth;

(3) no application of a 1-rule at smaller depth precedes an application of a 1-rule at a greater depth.

Theorems 7 – 9 show that for any theorem AB of $\mathbf{TW}^{01}_{\rightarrow}$–ID there is a derivation $A \longrightarrow B$ in a normal form. In the sequel we shall assume that all derivations are in normal form.

We define an application of a 0-rule to be *principal* if the depth of the occurrence of the subformula to be substituted is 0.

A derivation $A \longrightarrow B$ is called *principal* if it consists of principal applications of 0-rules at depth 0 only.

If there is a principal derivation $A \longrightarrow B$, we shall write $A \Longrightarrow B$ to denote this fact and the principal derivation of B from A.

A derivation $A \longrightarrow B$ is called *non-principal* if it is not principal.

A derivation of $A \longrightarrow B$ is called *inner* iff it contains no principal application of a 0-rule.

Theorem 11. *$AB \longrightarrow CD$ is an inner derivation iff either (1) $A = C$ and $B \longrightarrow D$ or (2) $B = D$ and $C \longrightarrow A$ or (3) $B \longrightarrow D$ and $C \longrightarrow A$.*

Proof. By the normal form theorem we have the "if part" of the theorem.

On the other hand, in each of the cases (1) - (3) we prove $AB \longrightarrow CD$.

Theorem 12. $\mathbf{TW}^{01}_{\rightarrow}$–ID *is equivalent to* $\mathbf{TW}_{\rightarrow}$–ID.

Proof. If $D[AB, a0] \longrightarrow D[AB/BC.AC, a0]$ in $\mathbf{TW}^{01}_{\rightarrow}$–ID contains a single application of SU^0, then, by Theorem 7, $D[AB, a0]D[AB/BC.AC, a0]$ is a theorem of $\mathbf{TW}_{\rightarrow}$–ID. This shows that $\mathbf{TW}_{\rightarrow}$–ID is closed under SU^0.

In a similar way, by using Theorem 7, we take care of the remaining rules.

Hence, if A is a theorem of $\mathbf{TW}^{01}_{\rightarrow}$–ID, then it is a theorem of $\mathbf{TW}_{\rightarrow}$–ID.

The axioms of $\mathbf{TW}_{\rightarrow}$–ID are easily derived in $\mathbf{TW}^{01}_{\rightarrow}$–ID.

It is obvious that in $\mathbf{TW}^{01}_{\rightarrow}$–ID we have neither $p \longrightarrow A$ nor $A \longrightarrow p$ nor $A \longrightarrow pp$. Therefore, we may write MP in the form: if $A \longrightarrow B$ and $AB \longrightarrow CD$, then $C \longrightarrow D$.

Let us show that $\mathbf{TW}^{01}_{\rightarrow}$–ID is closed under MP.

Suppose that $A \longrightarrow B$ and $AB \longrightarrow CD$.

If $AB \longrightarrow CD$ is an inner derivation, then either $A = C$ and $C \longrightarrow B \longrightarrow D$ or $B = D$ and $C \longrightarrow A \longrightarrow D$ or $C \longrightarrow A \longrightarrow B \longrightarrow D$, by Theorem 1.11.

Let $AB \Longrightarrow EF.GH \longrightarrow CD$ and $AB \neq EF.GH$.

By induction on the number of principal applications of 0-rules we can prove that $EF \longrightarrow GH$.

Suppose that there is only one principal application of a 0-rule; then either (a) $B = E$, $G = A$ and $F = H$ or else (b) $E = G$, $F = A$ and $H = B$, say (a).

We have $A \longrightarrow B$; this means that starting with A and applying the 0-and 1-rules we eventually obtain B. Let us start with BF; in this formula every consequent occurrence of a subformula in B is an antecedent occurrence in BF, and conversely, every antecedent occurrence of a subformula in B is a consequent occurrence in BF. It is easy to see that AF can be obtained from BF by applying the same rules that lead from A to B in reverse order. This means that AF is obtained from BF by applying a 0-rule instead of the corresponding 1-rule and a 1-rule instead of the corresponding 0-rule.

Hence, if $A \longrightarrow B$, then $BF \longrightarrow AF$.

We proceed in case (b) in a similar way.

Suppose that the number of principal derivations in $AB \Longrightarrow EF.GH$ is greater than 1. If $EF.GH$ is obtained by SU^0 from GE, then $F = H$ and $G \longrightarrow E$, by induction hypothesis. As above, we see that $EF \longrightarrow GH$.

We proceed in a similar way if $EF.GH$ is obtained by PR^0.

As a consequence, if $CD = EF.GH$, then $C \longrightarrow D$. Otherwise, we have $EF.GH \longrightarrow CD$; hence, by Theorem 1.11, either $C = EF$ and $C \longrightarrow GH \longrightarrow D$ or $D = GH$ and $C \longrightarrow EF \longrightarrow D$ or $C \longrightarrow EF \longrightarrow GH \longrightarrow D$.

Therefore, $\mathbf{TW}^{01}_{\rightarrow}$–ID is closed under MP and every theorem of $\mathbf{TW}_{\rightarrow}$–ID is a theorem of $\mathbf{TW}^{01}_{\rightarrow}$–ID.

In the sequel we shall write "theorem" instead of "theorem of $\mathbf{TW}_{\rightarrow}$–ID (of $\mathbf{TRW}_{\rightarrow}$–ID, of $\mathbf{TW}^{01}_{\rightarrow}$–ID)" if it is not important to refer to the system under consideration.

The following theorems will be useful in the next section.

Theorem 13. *If A_1B_1, ..., A_mB_m and $A_1B_1. \dots .A_mB_m$ are theorems, $m \geq 1$, then the following conditions are satisfied, $m \geq 1$, for any $k, l \in \{1, \dots m\}$:*

(a) *either $A_k = A_m$ or $A_m \longrightarrow A_k$;*
(b) *if $k < l$ and $A_l \neq A_m$, then $A_l \longrightarrow A_k$;*
(c) *either $B_k = B_m$ or $B_k \longrightarrow B_m$;*
(d) *if $k < l$ and $B_k \neq B_m$, then $B_k \longrightarrow B_l$;*
(e) *$A_m = A_k$ for at most one $k \in \{1, \dots, m-1\}$;*
(f) *$B_m = B_k$ for at most one $k \in \{1, \dots, m-1\}$.*

Proof. Let A_1B_1, ..., A_mB_m be theorems.

By induction on the length of proof of $A_1B_1. \dots .A_mB_m$ in $\mathbf{TW}_{\rightarrow}$–ID we can prove that the conditions (a) - (f) are satisfied.

Suppose that $A_0B_0.A_1B_1.A_2B_2$ is an instance of ASU; then $A_2 = A_0$, $B_0 = A_1$ and $B_1 = B_2$. By NOID, $A_2 \neq A_1$ and $B_2 \neq B_0$. If it is an instance of APR, then $A_2 = A_1$, $A_0 = B_1$ and $B_0 = B_2$. By NOID, $A_2 \neq A_0$ and $B_2 \neq B_1$. Hence, (a) - (f) hold.

Suppose that A_0B_0 and $A_0B_0. \dots .A_mB_m$ are theorems such that

$$A_1B_1. \dots .A_mB_m$$

is obtained by MP. By induction hypothesis, (a) - (f) hold for A_0B_0, A_1B_1, ..., A_mB_m, and $A_0B_0.A_1B_1. \dots .A_mB_m$; it is easy to check that they hold for A_0B_0, A_1B_1, ..., A_mB_m, and $A_1B_1. \dots .A_mB_m$ as well.

Theorem 14. *If (a) $A \longrightarrow B_1 \longrightarrow \dots \longrightarrow B_n \longrightarrow C$, $n \geq 1$, then (b) $AB_1.B_1B_2. \dots .B_kB_{k+1}. \dots .B_{n-1}B_n.B_nC.AC$ is a theorem.*

Proof. Suppose that (a) and proceed by induction on n. If $n = 1$, then $AB_1.B_1C.AC$ is an axiom. If $n > 1$, suppose that D is a theorem, where

$D = AB_1.B_1B_2. \dots .B_kB_{k+1}. \dots .B_{n-1}B_n.AB_n$. Since B_nC is a theorem, by PR we obtain $D.AB_1.B_1B_2. \dots .B_kB_{k+1}. \dots .B_{n-1}B_n.AC$. Hence, by MP we have (b).

2. THE CONJECTURE

Let $A \longrightarrow B \longrightarrow C$; then B is called an *interpolant* for AC.

Let $A \longrightarrow B_1 \longrightarrow B_2 \longrightarrow \dots \longrightarrow B_{n-1} \longrightarrow B_n \longrightarrow C$; the sequence $B_1, \dots, B_n$ is called a *chain of length* n from A to C.

Let there be a chain from A to C; if there is a natural number n such that any chain from A to C is of length not greater than n, then n is called the *interpolation number* for the theorem AC. If there is no interpolant for AC, then the interpolation number for AC is zero.

The interpolation number for a theorem AC, if it exists, is the unique upper limit of length of any chain from A to C.

It is clear that if any theorem AC has an interpolation number, then there is no theorem of $\mathbf{TW}^{01}_{\rightarrow}$-ID of the form AA.

On the other hand, by using NOID we can prove that for some theorems AC there is only a finite set of chains from A to C. Hence, for some theorems NOID and the existence of at most a finite number of finite chains for these theorems of $\mathbf{TW}^{01}_{\rightarrow}$-ID are equivalent.

If the set of chains from A to C is finite, then there is an interpolation number for AC. On the other hand, a *priori* there might be an interpolation number for AC and yet the set of chains from A to C to be infinite.

Now we state our conjecture:

CONJ$_1$ *for any theorem* AC *of* $\mathbf{TW}^{01}_{\rightarrow}$-ID *there is an interpolation number.*

The conjecture is non-trivial. Suppose that

$$A \longrightarrow B_1 \longrightarrow \dots \longrightarrow B_m \longrightarrow B_{m+1} \longrightarrow \dots \longrightarrow B_{m+n} \longrightarrow C$$

such that each B_i, $1 \leq i \leq m$ is obtained from its immediate predecessor by a 0-rule, and that each B_j (and C), $m < j \leq m + n$, is obtained from its immediate predecessor by a 1-rule, in a derivation in normal form. It is clear that each B_i is of a degree greater than the degree of its immediate predecessor, and that each B_j (and C) is of a degree smaller than the degree of its immediate predecessor. If the conjecture is false, then there is a theorem AC of $\mathbf{TW}^{01}_{\rightarrow}$-ID such that for any m and n there are interpolants $B'_1 \dots B'_{m+1}, B'_1 \dots B'_{m+n+1}$ for AC.

Hence, the conjecture means that if $A \longrightarrow C$ in $\mathbf{TW}^{01}_{\rightarrow}$–ID, then there is a natural number n such that the number of applications of 0-and-1 rules in any derivation of $A \longrightarrow C$ is smaller than n.

Let us consider the conjecture in $\mathbf{TRW}_{\rightarrow}$–ID. Suppose that there is a proof of AC containing n applications of TR. By the normal form theorem, there are formulas $B_1 \ldots B_n$ and a normal form of this proof

$$AB_1, B_1B_2, \ldots B_{n-1}B_n, B_nC$$

containing n applications of TR. Of course, there might be *possible* applications of TR in the proof of AC that were not realized – we might have used TR more than n times, but we didn't. Let us call such applications of TR *hidden*. Now our conjecture is equivalent to the claim:

CONJ_2 *The number of hidden applications of* TR *in a proof of* AC *in* $\mathbf{TRW}_{\rightarrow}$–ID, *for any* A *and* C, *is finite.*

This is in $\mathbf{TRW}_{\rightarrow}$–ID equivalent to the claim:

$\mathrm{CONJ'}_2$ *For any* A *and* C *there is a natural number* n *such that in any normal proof of* AC *the number of interpolants for* AC *in such a proof is smaller than* n.

The conjecture in $\mathbf{TW}_{\rightarrow}$–ID can be stated is the following form:

CONJ_3 *If* AC *is a theorem, then there is a natural number* n *such that any proof of* AC *in* $\mathbf{TW}_{\rightarrow}$–ID *is of length smaller than* n.

Let us show that CONJ_3 is equivalent to CONJ_1.

Suppose that AC is a theorem and that there is a natural number n such that any proof of AC in $\mathbf{TW}_{\rightarrow}$–ID is of length smaller than n, but that, on the other hand, there is a chain $B_1, \ldots, B_m$, $n < m$, from A to C. By Theorem 14, $AB_1.B_1B_2. \ldots .B_{m-1}B_m.B_mC.AC$ is a theorem. Since AB_1, B_1B_2, …, $B_{m-1}B_m$, B_mC are theorems as well, there is a proof of AC in $\mathbf{TW}_{\rightarrow}$–ID of length m, contrary to CONJ_3.

Suppose that n is the interpolation number for AC in $\mathbf{TW}^{01}_{\rightarrow}$–ID and that there is a proof of AC in $\mathbf{TW}_{\rightarrow}$–ID of length $m > n$. By Theorem 1 there is a node of this proof of the form $C_1. \ldots .C_m.AC$ such that $C_1. \ldots .C_m$ are theorems. It is clear that $C_k = A_kB_k$ for some A_k and B_k, $1 \leq k < m$. Let us apply Theorem 13. Suppose that there are $1 \leq i, j \leq m$ such that $A = A_i$ and $C = B_j$. We have $A \longrightarrow A_m \longrightarrow \ldots \longrightarrow A_{k+1} \longrightarrow A_{k-1} \longrightarrow A_1 \longrightarrow B_1 \longrightarrow B_{j-1} \longrightarrow B_m \longrightarrow C$ – a chain of length greater than n from A to C, contrary to CONJ_1.

REFERENCES

[1] Anderson A.R., Belnap N.D., *Entailment, the Logic of Relevance and Necessity*, Vol. I, Princeton Univ. Press, 1975, §8.11.

[2] Komori Y., *Syntactical investigations into BI logic and $BB'I$ logic*, Studia Logica 53 (1994), 397–416.

[3] Kron A., *A constructive proof of a theorem in relevance logic*, Z. Math. Logik Grund. Math. 31 (1985), 423–430.

[4] Kron A., *Identity and permutation*, Publications de l' Institut Mathématique, Nouvelle série, 57 (71) (1995), 165–178. Đuro Kurepa memorial volume.

[5] Martin E., Meyer R.K., *Solution to the P - W problem*, J. Symbolic Logic 47 (1982), 869–887.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 31-34

A NOTE ON KUREPA'S HYPOTHESES IN ALGEBRA

Radoslav M. DIMITRIĆ
Institute of Mathematics, Serbian Academy of Science and Arts
Knez Mihailova 35, Beograd, Yugoslavia

Abstract. This short note will indicate the role Kurepa's trees have assumed in algebra and will lead the reader to the sources that discuss relationships of Kurepa's hypotheses with algebraic constructions, such as p-groups (both commutative and non-commutative) and valuated vector spaces. More details may be found in the forthcoming [3] and the references therein. The relevant initial source is Kurepa's dissertation, (reprinted in [5]).

Keywords: valuated vector space, Kurepa's hypothesis, abelian p-group, C_{ω_1}-group, balanced projective dimension, extraspecial groups, classes $\mathcal{Z}_\kappa$ and $\mathcal{Y}_\kappa$.

AMS Subject Classification (1991): Primary 04A20, 20K10, 20F24; Secondary 20K40, 20K45, 20F50, 20D25.

1. NON-ABELIAN GROUPS

The relationship between Kurepa's trees and non-abelian groups was essentially in the making when B.H. Neumann brought about various characterisations of finite-by-abelian and center-by-finite groups (as in [6]). The author did not anticipate the relationship at the time; it sprang out of later generalizations of the results in this paper.

Proposition 1. *(Brendle, [1]) The following are equivalent*

(1) There is a Kurepa tree

(2) There is an extraspecial p-group which is $\mathcal{Z}_{\omega_1}$ but not $\mathcal{Z}_{\omega_2}$

(3) There is an FC group which is $\mathcal{Z}_{\omega_1}$, but not $\mathcal{Z}_{\omega_2}$

Proposition 2. *(Brendle, [1]) The following are equiconsistent:*

(1) $ZFC + \neg\ KH$

(2) $ZFC+$ *for any* FC*-group* G *and* $\kappa = \omega_1, \omega_2$ *if* $|G/Z(G)| = \kappa$, *then there is an abelian subgroup* $A \leq G$, *with* $[G : N_G(A)] = \kappa$.

(3) $ZFC+$ *any extraspecial p-group of size* ω_2 *has an abelian subgroup* A *with* $[G : N_G(A)] = \omega_2$.

2. VALUATED VECTOR SPACES

We begin with a statement about the existence of κ-Kurepa trees in terms of valuated vector spaces:

Theorem 1. *(Cutler, Dimitrić, [2]) Let* κ *be an uncountable regular cardinal and* $\aleph$ *a cardinal greater than* κ. *Then there is a* κ*-Kurepa tree with at least* $\aleph$ κ*-branches if and only if, for every field* F *of cardinality* $< \kappa$, *there exists a valuated* F*-vector space* V *with the following properties:*

(a) $|V| = \kappa$,
(b) $V(\kappa) = 0$,
(c) for every $i < \kappa$, $|V/V(i)| < \kappa$,
(d) the completion $\hat{V}$ *of* V *in the* κ*-topology has cardinality* $\geq \aleph$.

Remark: If in the theorem above $|\hat{V}| = \aleph$, then the constructed κ-Kurepa tree T has exactly $\aleph$ κ-branches.

The following is a strengthened version of Theorem 1 (from the existence of a certain valuated vector space to the existence of a Kurepa family).

Theorem 2. *(Cutler, Dimitrić, [2]) Let* κ *be an uncountable regular cardinal and* $\aleph$ *a cardinal greater than* κ. *Then there is a* κ*-Kurepa family of cardinality* $\geq \aleph$, *if and only if there exists a valuated vector space* V *of cardinality* κ, *over a field of cardinality* $< \kappa$, *with the following properties:*

(a) $V(\kappa) = 0$,
(b) for every (limit) $i < \kappa$, $V/V(i)$ *has the* (i, κ)*-closure property.*
(c) the completion $\hat{V}$ *of* V *in the* κ*-topology has cardinality* $\geq \aleph$.

3. ABELIAN P-GROUPS

Theorems 1 and 2 are used to give proofs of two Keef's results, utilizing some of the techniques similar to those he used. The approach in [2] is different in that it deals only with the socles of the groups in question whenever possible.

Proposition 3. *(Keef, [4]; Cutler, Dimitrić, [2]) Kurepa's hypothesis is equivalent to the existence of a C_{ω_1}-group G of length ω_1 and cardinality $> \aleph_1$ with a p^{ω_1}-pure subgroup A of cardinality $\aleph_1$ such that the closure of A in G in the ω_1-topology has cardinality $\kappa > \aleph_1$.*

A group G with properties as in the last proposition, is called a *κ-Kurepa extension* of A, whereas A is a *κ-Kurepa subgroup* of G (all for $\kappa \geq \aleph_2$). It may be shown that if such an extension exists, then there is one satisfying $|G| = \kappa$. Thus, there exists a κ-Kurepa extension if and only if there exists a Kurepa family of cardinality κ.

Theorem 3. *(Keef, [4]; Cutler, Dimitrić, [2]) Kurepa's hypothesis is equivalent to the existence of a C_{ω_1}-group of length ω_1 and balanced projective dimension 2.*

Theorem 4. *(Keef, [4]) The following are equivalent*

(1) $\neg KH$

(2) Every C_{ω_1}-group of length ω_1 has a balanced projective dimension of at most 1.

(3) For any C_{ω_1}-groups A, B of length ω_1, $\operatorname{Tor}(A, B)$ *is a direct sum of countables.*

(4) For every C_{ω_1}-group A of length ω_1, and every group B, the group $\operatorname{Ext}(A, B)/\operatorname{Ext}(A, B)(\omega_1)$ *is complete in the ω_1-topology*

(5) The class of ω_1-dsc's coincides with the class of pseudo-dsc's

(6) An ω_1-dsc is a dsc iff it is complete in its ω_1-topology.

REFERENCES

[1] Brendle, J., *Set-theoretic aspects of periodic FC-groups – Extraspecial p-groups and Kurepa trees*, J. Algebra 162 (1993), 259–286.

[2] Cutler D., Dimitrić R., *Valuated vector spaces, Kurepa's hypothesis and abelian p-groups*, Rocky Mountain J. Math 23 (1993), 1253–1266.

[3] Dimitrić R., *Algebraic equivalents of Kurepa's hypothesis*, Submitted (1996).

[4] Keef, P., *On set theory and the balanced-projective dimension of C_Ω groups*, Contemporary Math. 87 (1989), 31–41.

[5] Kurepa Đ., *Ensembles ordonnés et ramifiés*, Publ. Math. Univ. Belgrade 4 (1935), 1–138.

[6] Neumann B.H., *Groups with finite classes of conjugate subgroups*, Math Z. 63 (1955), 76–96.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 35-39

COMPLETENESS THEOREM FOR PROBABILITY MODELS WITH FINITELY MANY VALUED MEASURE †

Miodrag D. RAŠKOVIĆ, Radosav S. ĐORĐEVIĆ
Faculty of Natural Sciencies and Mathematics, University of Kragujevac, Radoja Domanovića 12, 34000 Kragujevac, Yugoslavia

Abstract. The aim of the paper is to prove the completeness theorem for probability models with finitely many valued measure.

Keywords: probability models, finitely valued measure

AMS Subject Classification (1991): 03C70.

Let $\mathcal{A}$ be a countable admissible set such that $\mathcal{A} \subseteq \mathrm{HC}$ and $\omega \in \mathcal{A}$. The probability logic $L_{\mathcal{AP}}$ was introduced in [4] by H.J. Keisler. This logic is similar to the infinitary logic $L_{\mathcal{A}}$, except that probability quantifiers $P\vec{x} \geq r$ ($\vec{x}$ is a finite sequence of variables) are used instead of the usual $\forall x$ and $\exists x$. A model of this logic is a classical model without operations with a probability measure on the universe, such that each relation is measurable. The formula $(P\vec{x} \geq r)\varphi(\vec{x})$ means that the set $\{\, \vec{x} : \varphi(\vec{x}) \,\}$ has a probability greater than or equal to r.

The axioms and rules of inference of $L_{\mathcal{AP}}$ are listed in [5].

The logic $L_{\mathcal{AP}}^{\mathrm{fin}}$ is similar to the probability logic $L_{\mathcal{AP}}$. The only difference is that the list of axioms of $L_{\mathcal{AP}}^{\mathrm{fin}}$ has the following *axiom of finitely many valued measure* added:

$$\bigvee\nolimits_{c\in Q^+} \bigwedge\nolimits_{\varphi\in\Phi_n} \big((P\vec{x} > 0)\varphi(\vec{x}) \to (P\vec{x} > c)\varphi(\vec{x})\big),$$

where $\Phi_n \in \mathcal{A}$ and $\Phi_n = \{\, \varphi : \varphi \text{ has } n \text{ free variables} \,\}$.

†Research supported by the Science Foundation of Serbia grant number 0401A, through Math. Inst. Serbian Academy of Science and Art

The probability structure for $L_{\mathcal{AP}}^{\text{fin}}$ is a structure $\langle \mathcal{U}, \mu \rangle$ such that $\mathcal{U}$ is a classical first-order structure without operations and μ is a finitely many valued probability measure.

We shall prove that this axiomatization is complete for Σ_1 definable theories with respect to the class of probability models with finitely many valued measure, by combining a consistency property argument, such as that of Keisler [5] or Hoover [3], and a weak-middle-strong model construction, such as that of Rašković [6].

We shall introduce two sorts of auxiliary structures.

Definition 1. *(i) A weak structure for $L_{\mathcal{AP}}^{fin}$ is a structure $\langle \mathcal{U}, \mu_n \rangle_{n \geq 1}$ such that each μ_n is a finitely additive probability measure on A^n with each singleton measurable and the set $\varphi_{\vec{a}} = \{ \vec{b} \in A^n : \langle \mathcal{U}, \mu_n \rangle \models \varphi[\vec{a}, \vec{b}] \}$ is μ_n-measurable for each $\varphi(\vec{x}, \vec{y}) \in L_{\mathcal{AP}}^{fin}$ and $\vec{a} \in A^m$.*

(ii) A middle structure for $L_{\mathcal{AP}}^{fin}$ is a weak structure $\langle \mathcal{U}, \mu_n \rangle$ such that the following is true: There is a $c > 0$ such that for each formula $\varphi(\vec{x}, \vec{y}) \in L_{\mathcal{AP}}^{fin}$ and each $\vec{a} \in A^m$, if $\mu_n(\varphi_{\vec{a}}) > 0$, then $\mu_n(\varphi_{\vec{a}}) > c$.

By means of a consistency property argument similarly as in [3] or [5] we prove that a Σ_1 definable theory of $L_{\mathcal{AP}}^{\text{fin}}$ is consistent if and only if it has a weak model in which each theorem of $L_{\mathcal{AP}}^{\text{fin}}$ is true. Let $C \in \mathcal{A}$ be a set of new constant symbols introduced in this Henkin construction and let $K = L \cup C$.

Theorem 1. (Middle Completeness Theorem) *A Σ_1 definable theory T of $K_{\mathcal{AP}}^{fin}$ is consistent if and only if it has a middle model in which each theorem of $K_{\mathcal{AP}}^{fin}$ is true.*

Proof. In order to prove that consistent Σ_1 definable theory T of $K_{\mathcal{AP}}^{\text{fin}}$ has a middle model, we introduce language M with three sorts of variables, such as that of Rašković [6]: $X, Y, Z, \ldots$ variables for sets, $x, y, z, \ldots$ variables for urelements and $r, s, t, \ldots$ variables for reals from $[0, 1]$. The predicates of M are $\leq$ for reals, $E_n(\vec{x}, X)$ for $n \geq 1$ and $\vec{x} = x_1, \ldots, x_n$ (with the canonical meaning $\vec{x} \in X$) and $\mu(X, r)$ (with the meaning $\mu(X) = r$). The constant symbols are set constant symbol A_φ for each $\varphi \in K_{\mathcal{AP}}^{\text{fin}}$ and $\mathbf{r}$ for each $r \in [0, 1] \cap \mathcal{A}$. The functional symbols are $+$ and $\cdot$ for reals.

Let S be the first-order theory of $M_{\mathcal{A}}$ which has the following list of formulas: axiom of well-definedness, an axiom of extensionality, axioms of satisfaction, axioms which tell us that μ is an additive function, axioms for an Archimedean field and axioms which are transformations of axioms of

$K_{\mathcal{AP}}^{\mathrm{fin}}$ as listed in Rašković [6] (with the remark that $\mu_1 = \mu_2 = \mu$), together with the *axiom of realizability of all sentences* φ *in* T

$$(\forall x)E_1(x, A_\varphi),$$

and the *axiom of finitely many valued measure*

$$(F) \qquad (\exists c > 0)(\forall X)\big(\mu(X) > \mathbf{0} \to \mu(X) > c\big),$$

where $\mu(X) > \mathbf{r} \overset{\mathrm{def}}{\Longleftrightarrow} (\exists s)\big(s > \mathbf{r} \wedge \mu(X, s)\big)$.

A *standard structure* for $M_{\mathcal{A}}$ is the structure

$$\mathcal{B} = \langle B, P, E_n^{\mathcal{B}}, \mu^{\mathcal{B}}, +, \cdot, \leq, A_\varphi^{\mathcal{B}}, r\rangle_{n\geq 1, \varphi\in K', r\in F},$$

where $P \subseteq \bigcup_{n\geq 1} \mathcal{P}(\mathcal{B}^\setminus)$, $E_n^{\mathcal{B}} \subseteq B^n \times P$, $F = F' \cap [0,1]$, $F' \subseteq R$ is a field, $\mu^{\mathcal{B}}: P \to F$, $+, \cdot: F^2 \to F$, $\leq \subseteq F^2$, $A_\varphi^{\mathcal{B}} \in P$ and $K' \subseteq K_{\mathcal{AP}}^{\mathrm{fin}}$.

The theory S is Σ_1 definable over $\mathcal{A}$. To prove that S is consistent it is enough, by the Barwise Compactness Theorem (see [1]), to show that $S_0 \subseteq S$, $S_0 \in \mathcal{A}$ has a standard model. First, note that a weak structure $\langle \mathcal{U}, \mu_n\rangle$ for $K_{\mathcal{AP}}^{\mathrm{fin}}$ can be transformed into a standard structure by taking: $A_\varphi^{\mathcal{B}} = \{\, \vec{a} \in A^n : \langle \mathcal{U}, \mu_n\rangle \models \varphi[\vec{a}]\,\}$ and $P = \{\, A_\varphi^{\mathcal{B}} : \varphi \in K_{\mathcal{AP}}^{\mathrm{fin}}\,\}$. Since the axiom

$$\bigvee\nolimits_{c\in Q+} \bigwedge\nolimits_{\varphi\in(S_0')_n} \big((P\vec{x} > 0)\varphi(\vec{x}) \to (P\vec{x} > c)\varphi(\vec{x})\big)$$

holds in the weak model $\langle \mathcal{U}, \mu_n\rangle$, where $S_0 \subseteq S_0'$, $S_0' \in \mathcal{A}$ is the closure for the substitution of constant symbols from C and disjunction and $(S_0')_n = \{\, \varphi \in S_0' : \varphi \text{has } n \text{ free variables}\,\}$, it follows that

$\big\langle A, \big\{\{\vec{a} \in A^n : \langle \mathcal{U}, \mu_n\rangle \models \varphi[\vec{a}, \vec{c}]\} : \varphi \in S_0\big\}, E_n^{\mathcal{U}}, \mu^{\mathcal{U}}, +, \cdot, \leq, \{\vec{a} \in A^n : \langle \mathcal{U}, \mu_n\rangle \models \varphi[\vec{a}, \vec{c}]\}, r\big\rangle_{n\geq 1, \varphi\in S_0, r\in[0,1]\cap\mathcal{A}}$

is the standard model for S_0' and S_0, too.

Lastly, note that a standard model $\mathcal{B}$ of S can be transformed into a middle model $\overline{\mathcal{B}}$ of T by taking:

$$\vec{x} \in R^{\overline{\mathcal{B}}} \text{ iff } E_n^{\mathcal{B}}(\vec{x}, A_R) \quad \text{for an } n\text{-ary relational symbol } R \in L,$$

$$\mu_n^{\overline{\mathcal{B}}}(X) = r \text{ iff } \mu^{\mathcal{B}}(X, r) \quad \text{for } X \in \mathcal{P}(\mathcal{B}^\setminus).$$

In order to prove the main result, we shall use the following theorem (see [2]).

Theorem 2. *Let $\mathcal{F}$ be a field of subsets of a set Ω. Then μ is a finitely many valued probability measure on $\mathcal{F}$ if and only if there is a real number $c > 0$ such that $\mu(A) > c$ whenever $A \in \mathcal{F}$ and $\mu(A) > 0$.*

It follows from the Loeb-Hoover-Keisler construction (see [3] and [5]) that the axiom of finitely many valued measure implies that (F) holds for all internal sets in the nonstandard superstructure. The property (F) also holds for all Loeb measurable sets because these can be approximated by internal ones. Thus, it follows from Theorem 2 that each middle model in which all theorems of $L_{\mathcal{AP}}^{\mathrm{fin}}$ hold is elementarily equivalent to a probability model for $L_{\mathcal{AP}}^{\mathrm{fin}}$. As a consequence of the preceding, we obtain

Theorem 3. (Completeness Theorem for $L_{\mathcal{AP}}^{\mathrm{fin}}$) *A Σ_1 definable theory T of $L_{\mathcal{AP}}^{fin}$ is consistent if and only if T has a probability model with finitely many valued measure.*

Finally, let us note that the structure $\langle \mathcal{U}, \mu \rangle$, where μ is a finitely many valued probability measure, cannot be axiomatized so that the extended completeness theorem holds. The following example (see [6]) of a countable consistent theory T in $L_{\mathcal{AP}}^{\mathrm{fin}}$ does not have a probability model with finitely many valued measure.

Example Let $L = \{ R_1(x), R_2(x), \ldots \}$ be a Δ_1 definable set which is not a subset of an element of $\mathcal{A}$, and let $\varphi_1, \varphi_2, \ldots$ be an enumeration of all formulas from $L_{\mathcal{AP}}^{\mathrm{fin}}$. Then there exists the first predicate, denoted by $S_n(x)$, not occurring in $\varphi_1, \ldots, \varphi_n$; otherwise $L \subseteq \mathrm{TC}(\varphi_1) \cup \ldots \cup \mathrm{TC}(\varphi_n) \in \mathcal{A}$, which would imply that $L \in \mathcal{A}$ as a Δ_1 definable set.

It is obvious that the countable theory

$$T = \{ (Pxy \geq 1) x \neq y \} \cup \{ (Px > 0)\big(S_1(x) \wedge \ldots \wedge S_n(x)\big) : n \in \omega \} \\ \cup \{ (Px < 1/2^n) S_n(x) : n \in \omega \}$$

does not have any probability model with finitely many valued measure. We prove that T is consistent in $L_{\mathcal{AP}}^{\mathrm{fin}}$.

Let A be a unit interval $[0,1]$ and let μ be a Lebesque measure on $[0,1]$. For $B_n = [0, 1/2^{n+1})$ we have $0 < \mu(B_n) < 1/2^n$. Let $A_{n_1 \ldots n_k}^{i_1 \ldots i_k} = B_{n_1}^{i_1} \cap \ldots \cap B_{n_k}^{i_k}$ be a Boolean atom, where $B_n^i = B_n$ for $i = 1$ and $B_n^i = A \setminus B_n$ for $i = -1$.

We interprete the predicates by taking $\mathrm{R}_{\mathrm{n}}^{\langle \mathcal{U}, \mu \rangle} = \begin{cases} B_n, & \text{if } R_n = S_m \text{ for some } m \\ B_1, & \text{otherwise.} \end{cases}$

Since only finitely many predicates $S_n(x)$ can occur in an element of $\mathcal{A}$, it follows that the set $\{ A_{n_1 \ldots n_k}^{i_1 \ldots i_k} : \mu\left(A_{n_1 \ldots n_k}^{i_1 \ldots i_k}\right) > 0 \}$ is finite. The theory T and all axioms of $L_{\mathcal{AP}}^{\mathrm{fin}}$ are satisfied except perhaps the axiom of finitely many valued measure. But for $c = \min\{ \mu\left(A_{n_1 \ldots n_k}^{i_1 \ldots i_k}\right) : \mu\left(A_{n_1 \ldots n_k}^{i_1 \ldots i_k}\right) > 0 \}$, it follows that this axiom holds as well. Thus, T is consistent in $L_{\mathcal{AP}}^{\mathrm{fin}}$.

REFERENCES

[1] Barwise J. *Admissible Sets and Structures* , Springer–Verlag, Berlin, 1975.

[2] Bhaskara Rao K.P.S., M. Bhaskara Rao, *Theory of Charges*, Academic Press, New York, 1983.

[3] Hoover D., *Probability logic*, Ann. Math. Logic 14 (1978), 287–313.

[4] Keisler H.J., *Hyperfinite model theory*, in: *Logic Colloquium '76*, North–Holland, Amsterdam, 1977, 5–110.

[5] Keisler H.J., *Probability quantifiers*, in: J. Barwise, S. Feferman (eds.), *Model Theoretic Logics*, Springer–Verlag, Berlin–Heidelberg–New York, 1985, 509–556.

[6] Rašković M., *Completeness theorem for biprobability models*, J. Symbolic Logic 51 (1986), 586–590.

;CIENTIFIC REVIEW (1996), Number 19-20, pp 41-53

INTENSIONAL SEMANTICS APPROACH TO LOCATIVE QUALIFIERS

Marica PREŠIĆ, Milan BOŽIĆ

Faculty of Mathematics, University of Belgrade, Studentski trg 16, 11000 Belgrade, Yugoslavia

In the natural language there are a great number of words and expressions, more or less complex, to describe location and direction. In the actual world, the location of a non-moving individual or direction of its movement, should the individual be moving, can generally be determined in two ways:

(I) **Absolute:** Using some sort of coordinate system or some system of addresses, for example: country - town - street - house - apartment - individual.

(II) **Relative:** Using some specific individual as the main point of locative reference and determining the position of all other moving or non-moving individuals relative to this point. Usually the point of locative reference is: speaker, listener, a person being the topic of conversation, or some object. Practically any individual can be the main point of locative reference in some context, which gives the consideration of locations and directions the highest degree of complexity.

For expressing relative position of an individual (moving or non-moving) many natural languages usually have a list of **prepositions** which play the key role in the description of locations and directions. For example, English has the following list [5]:

- *at, to from, toward(s), away from, against*
- *on, on to, off*
- *in, into, out off, inside, outside, within*
- *in front of, behind, from behind, beyond, before, after*

- *by, past, beside, near, opposite*
- *along, alongside*
- *across*
- *around/round, about*
- *between, among*
- *above, below, beneath, underneath, up, up to, down, under, from under*
- *over*

By combining prepositions with common noun or noun phrases (in the corresponding noun case) a long list of locative and directional expressions can be built. For example:

> *at the table, on the table, in the table, across the road, along the road*
> *near me, away from me, before me, after me*

Apart from this there are also **adverbs** or **adverbial expressions** such as: *here, there* (in Serbian there are three possibilities: *ovde, tu, onde*), *to here, to there* (in Serbian *ovamo, tamo, onamo*), *from here, from there* (in Serbian *odavde, odatle, odande*), *left, right, this way, next to* and similar.

In addition to these simple words or expression, location and direction can be determined by means of very complex natural language locutions and transformations: relativization, nominalization, all sorts of embedded sentences and other kinds of syntactic transformations. Just to feel the complexity of the topic we give several examples:

> *<u>The book which is on the table</u> was given to me by Mary*
>
> *My brother proposed that we all go <u>to a theater performance</u>*
>
> *There was <u>a theater performance to which my brother proposed that we all go</u>*
>
> *<u>The orchard where we are picnicking</u> belongs to my uncle*

Intensional logic [6], [10] seems to be good enough for the treatment of locations and directions. It suffices to add to the basic types the types of locations and directions. Thus the appropriate list of basic type would be o, ι, α, β, π, δ, τ, ω (truth values, individuals, speakers, listeners, locations, directions, time moments, possible worlds). Compound types are built up in the standard functional way.

Locative and directional qualifiers and qualifying expressions have different roles in the syntactic and semantic structure of a sentence. The <u>first</u> role is the one of verb, noun or adjective ***inner participant*** (obligatory or optional). In this case it belongs to the framework of the verb, noun or adjective and is reflected in its type. Verbs of such kind are:

live (somewhere), *come* (somewhere), *arrive* (somewhere), *arrive* (from somewhere), *return* (somewhere), *come* (from somewhere), *be* (somewhere), *travel* (somewhere)
put (something somewhere), *take* (something from somewhere)
move (something from somewhere to somewhere)

The types of the verbs *live* (somewhere), *put* (something somewhere) would then be:

$o(\iota\ \pi)(\alpha)(\beta)(\pi)(\delta)(\tau)(\omega)$, $o(\iota\ \iota\ \delta)(\alpha)(\beta)(\pi)(\delta)(\tau)(\omega)$

The second role is that of ***adverbial modifiers*** (of verb, noun, adjective, adverb, sentence). In this case locative and directional qualifiers (or qualifying expressions) do not change grammatical category (and type) of the expression they modify. For example, locative verb modifiers of unary and binary verbs are of the types:

$$(o(\iota)(o(\iota)(\alpha)(\beta)(\pi)(\delta)(\tau)(\omega))(\alpha)(\beta)(\pi)(\delta)(\tau)(\omega)$$

$$(o(\iota\ \iota)(o(\iota\ \iota\ )(\alpha)(\beta)(\pi)(\delta)(\tau)(\omega))(\alpha)(\beta)(\pi)(\delta)(\tau)(\omega)$$

respectively, and locative sentence modifiers are of the type:

$$(o(o(\omega))(\omega)$$

Following the ideas developed in [9] and [7] locative and directional deep case operators such as these:

$\mathrm{Loc}_{\ in}$, $\mathrm{Loc}_{\ on}$, $\mathrm{Dir}_{\ to}$, $\mathrm{Dir}_{\ from}$, $\mathrm{Dir}_{\ way}$

can built in the standard intensional logic. For example the sentences:

John came to Belgrade, John came to the capital of Serbia, John walks in a park

have the following logic structures:

came(John, to Belgrade), came(John, to the capital of Serbia)
$\exists\, x[park(x) \wedge in\ x(walk)(\ John)]$

Using locative and directional deep case operators the corresponding translations would be:

$$\lambda a\ \lambda b\ \lambda p\ \lambda d\ \lambda t\ \lambda w\ CA'\ (J, \mathrm{Dir}_{\ to}(B)),\ \ \lambda a\ \lambda b\ \lambda p\ \lambda d\ \lambda t\ \lambda w\ CA'\ (J, \mathrm{Dir}_{\ to}(CS)'\ )$$

$$\lambda a\ \lambda b\ \lambda p\ \lambda d\ \lambda t\ \lambda w\ (\exists\ x)(PA'\ (x) \wedge \mathrm{Loc}_{\ in}(x)'\ (WA'\ )(J))$$

where J, B, CS, PA, CA, WA are abbreviations for *John*, *Belgrade*, *the capital of Serbia*, *par.* *came*, *walks*, R' is the abbreviation for R(w)(t)(d)(p)(b)(a) where R is a symbol having the typ of form $\eta(\alpha)(\beta)(\pi)(\delta)(\tau)(\omega)$ (η being any type) and a, b, p, d, t, w are variables of types α, β, π δ, τ, ω respectively.

In the subsequent text we confine ourselves to locatives. The intensional logic we use is TIL [10 in which the basic types are pairwise disjoint non-empty sets o, ι, α, β, π, τ, ω (truth values individuals, speakers, listeners, locations, time moments, possible worlds, where o={T, $\perp$}). Th type π of locations is a subset of the partitative set P(E), where E is the domain of Euclidean space **E**. Thus locations are sets of vectors. The idea is based on the fact that solid objects in the actual world (for a given state of affairs) can be treated as sets of material points each having a specific position determined by coordinates in some three dimensional Euclidean space. The exact position of the object is then determined by the set of coordinates of all its points and the less precise position is determined by any set containing the set of these coordinates as its subset.

We give some examples of sentences with locative expressions and their translations into the intensional logic TIL. The atomic constructions of TIL corresponding to natural language words are built by the capitalization of the first letter and transformation of the italic format into non-italic. For example, for the sentence: *John lives in Paris*, the corresponding TIL atomic constructions are: John, Live, In, Paris. The type τ of the TIL expression S is usually noted in th following way: S / τ.

Example 1. The sentence:

John lives in Paris

is built from the binary verb *live* (somewhere), preposition *in*, and two proper names: *John* *Paris*. The corresponding atomic constructions in TIL and their types are the following:

John / ι, Paris / ι, a / α, b / β, p / π, t / τ, w / ω, Lives / $o(\iota\,\pi)(\alpha)(\beta)(\pi)(\tau)(\omega)$,

In / $\pi(\iota)$

To simplify the notations we use the following abbreviations:

$(_\omega)$ stands for $(\alpha)(\beta)(\pi)(\tau)(\omega)$, $_\lambda\mathbf{w}$ stands for $\lambda a\,\lambda b\,\lambda p\,\lambda t\,\lambda w$

R' stands for R(w)(t)(d)(b)(a) (R is of type $\eta(_\omega)$ for any η)

The construction corresponding to the above sentence is then the following:

$_\lambda\mathbf{w}$ Live'(John, In(Paris))

Similar to the previous is the sentence:

John lives in a village

in which, however, we have the determiner *a*. In the corresponding TIL construction the common noun symbol Village (having the type $o(\iota)(_\omega)$) and the existential quantifier $\exists$ occur:

$_\lambda\mathbf{w}$ $[(\exists z)([\text{Village}'(z) \wedge \text{Live}'(\text{John}, \text{In}(z))]]$

Of the same sort are the sentences:

John lives somewhere, John lives nowhere

having the following constructions:

$_\lambda\mathbf{w}$ $[(\exists p)\text{Live}'(\text{John}, p)]$, $_\lambda\mathbf{w}$ $[\neg(\exists p)\text{Live}'(\text{John}, p)]$

where p is a variable of type π.

The construction corresponding to the sentence *John lives in a village* can be written in a more natural way using the operator So having the type $(o(o(\iota)(_\omega)))(o(\iota)(_\omega))$:

$_\lambda\mathbf{w}$ $[\text{So}(\text{Village})(_\lambda\mathbf{w}\ [\lambda z\ \text{Live}'(\text{John}, \text{In}(z))]])]$

We recall that the definitions of operators So, Ev, both having the type $(o(o(\iota)(_\omega)))(o(\iota)(_\omega))$, read:

$\text{So}(X)(Y) \Leftrightarrow (\exists z)[X'(z) \wedge Y'(z)]$, $\text{Ev}(X)(Y) \Leftrightarrow \forall z[X'(z) \Rightarrow Y'(z)]$

where X,Y are variables of type $o(\iota)(_\omega)$, and z is a variable of type ι.

The sentence:

John lives in a small village

differs from the sentence *John lives in a village* only in the adjective *small*. The corresponding TIL symbol Small is of the type $((o(\iota))(o(\iota)(_\omega)))(_\omega)$, so that it suffices to built first the expression Small'(Village) which is of the type $o(\iota)$, and after that to precede the construction the same way as in the case of the sentence *John lives in a village*. The result will be:

$_\lambda\mathbf{w}$ $[\text{So}(_\lambda\mathbf{w}\ [\text{Small}'(\text{Village})])(_\lambda\mathbf{w}\ [\lambda z\ \text{Live}'(\text{John}, \text{In}(z))])]$

Example 2. The sentences:

It is dark in the park, In this room it is very cold have similar TIL constructions. The corresponding atomic constructions in TIL are:

$\text{Park} / o(\iota)_{(\omega)}$, $\text{Room} / o(\iota)(_{\omega})$, $\text{Dark} / o(\pi)(_{\omega})$, $\text{Cold} / o(\pi)(_{\omega})$,

$\text{Very} / (o(\pi)(o(\pi)(_{\omega})))(_{\omega})$

$\text{The} / (o(o(\iota)(_{\omega})))(o(\iota)(_{\omega}))$, $\text{This} / (o(o(\iota)(_{\omega})))(o(\iota)(_{\omega}))$, x / ι, $\text{In} / \pi(\iota)$

where Dark and Cold have the meaning "be dark somewhere", "be cold somewhere". The determiners The and This have the same type as So, Ev considered in the previous example. The constructions:

The(Park), This(Room)

are both of type $o(o(\iota)_{(\omega)})$, so that they can be applied on constructions of type $o(\iota)_{(\omega)}$. The constructions:

Dark'(In(x)), Very'(Cold)(In(z))

are formulae, i.e. they are of type 0, and the constructions:

$_{\lambda}\mathbf{w}$ [λz [Dark'(In(x)]], λ**w** [λz [Very'(Cold)(In(z))]]

are both of type $o(\iota)_{(\omega)}$ so that the constructions The(Park), This(Room) can be applied on them. The new constructions both of type o are obtained in the next step:

The(Park)($_{\lambda}\mathbf{w}$ [λz [Dark'(In(x)]]), This(Room)($_{\lambda}\mathbf{w}$ [λz [Very'(Cold)(In(z))]])

The application of λ operator on a, b, p, t, w, is all that remains:

$_{\lambda}\mathbf{w}$ [The(Park)($_{\lambda}\mathbf{w}$ [λz [Dark'(In(x)]])] ,

$_{\lambda}\mathbf{w}$ [This(Room)($_{\lambda}\mathbf{w}$ [λz [Very'(Cold)(In(z))]])]

Example 3. The construction corresponding to the sentence:

She wrote a letter with a pencil in a small café

is built from the atomic constructions:

$\text{Wrote} / o(\iota\ \iota\ \pi)_{(\omega)}$, $\text{Letter} / o(\iota)_{(\omega)}$, $\text{Pencil} / o(\iota)_{(\omega)}$, $\text{Café} / o(\iota)_{(\omega)}$,

$\text{Small} / (o(\iota)(o(\iota)_{(\omega)}))_{(\omega)}$

$\text{So} / (o(o(\iota)_{(\omega)}))(o(\iota)_{(\omega)})$, She / ι, x / ι, y / ι, z / ι, $\text{In} / _{\pi}(_{\iota})$

Applying the operator So three times on the formula:

$$\text{Wrote}'(\text{She}, x,y, \text{In}(z))$$

more precisely applying So($\lambda\mathbf{w}[\text{Small}'(\text{Café})]$), So(Pencil), So(Letter) on the corresponding constructions of type $o(\iota)_{(\omega)}$:

$$_{\lambda}\mathbf{w}\ [\lambda z\ [\text{Wrote}'(\text{She}, x, y, \text{In}(z))]],$$

$$_{\lambda}\mathbf{w}\ [\lambda y\ [\text{So}(_{\lambda}\mathbf{w}\ [\text{Small}'(\text{Café})])(_{\lambda}\mathbf{w}\ [\lambda z\ [\text{Wrote}'(\text{She}, x, y, \text{In}(z))]])]]$$

$$_{\lambda}\mathbf{w}\ [\lambda x\ [\text{So}(\text{Pencil})(_{\lambda}\mathbf{w}\ [\lambda y[\text{So}(_{\lambda}\mathbf{w}\ [\text{Small}'(\text{Café})])(_{\lambda}\mathbf{w}[_{\lambda}z[\text{Wrote}'(\text{She}, x, y, \text{In}(z))]])]])]]$$

respectively, we obtain:

$$\text{So}(\text{Letter})\ (_{\lambda}\mathbf{w}\ [\lambda x\ [\text{So}(\text{Pencil})(_{\lambda}\mathbf{w}\ [ly\ [\text{So}(_{\lambda}\mathbf{w}\ [\text{Small}'(\text{Café})])\ (_{\lambda}\mathbf{w}\ [\lambda z\ [\text{Wrote}'(\text{She}, x, y, \text{In}(z))]])]])]])$$

Finally, after applying $_{\lambda}\mathbf{w}$ the following construction:

$$_{\lambda}\mathbf{w}\ [\text{So}(\text{Letter})\ (_{\lambda}\mathbf{w}\ [\lambda x\ [\text{So}(\text{Pencil})(_{\lambda}\mathbf{w}\ [\lambda y\ [\text{So}(_{\lambda}\mathbf{w}\ [\text{Small}'(\text{Café})])\ (_{\lambda}\mathbf{w}\ [\lambda z\ [\text{Wrote}'(\text{She}, x, y, \text{In}(z))]])]])]])]$$

will be obtained. If we want to reduce the whole construction to the usage of the existential quantifier $\exists$ only, the result is the following:

$$_{\lambda}\mathbf{w}\ [\exists\ x\ y\ z\ [\text{Letter}'(x) \wedge \text{Pencil}'(y) \wedge \text{Small}'(\text{Café})(z) \wedge \text{Wrote}'(\text{She}, x, y, \text{In}(z))]]$$

Example 4. In the sentence:

Peter saw two young people there

the quantifier Two, having the expected type $(o(o(\iota)_{(\omega)}))(o(\iota)_{(\omega)})$, occurs. For any type ξ it can be defined as:

$$\text{Two}(X)(Y) \Leftrightarrow \exists\ z_1 z_2[z_1 \neq z_2 \wedge X'(z_1) \wedge X'(z_2) \wedge Y(z_1) \wedge Y(z_2)] \wedge Y' = \lambda x[x = z_1 \vee x = z_2]]$$

where X,Y are variables of type $o(\xi)_{(\omega)}$ and x, z_1, z_2 are variables of type ι.

Atomic constructions for the given sentence are:

Saw / $o(\iota\ \iota\ \pi)_{(\omega)}$, People/ $o(\iota)_{(\omega)}$, Young / $(o(\iota)(o(\iota)_{(\omega)}))_{(\omega)}$

There / $\pi_{(\omega)}$, Two / $((o(o(\iota)_{(\omega)}))(o(\iota)_{(\omega)})$, Peter / ι, x / ι, In / $\pi(\iota)$

First we construct Young'(People) which is of type $o(\iota)$ and after that we apply on it the operator Two. The obtained construction:

Two(Young'(People))

is of type $o(o(\iota)_{(\omega)})$ and it can be applied on the formula-in-intension:

$\lambda\mathbf{w}$ [λx [Saw'(Peter, x, There')]]

which yields the formula:

Two(Young'(People))($\lambda\mathbf{w}$ [λx [Saw'(Peter, x, There')]])

In the last step we apply the operator $\lambda\mathbf{w}$ and obtain the construction:

$\lambda\mathbf{w}$ [Two(Young'(People))($\lambda\mathbf{w}$ [λx [Saw'(Peter, x, There')]])]

Similar constructions can be obtained for the sentences:

Peter met two poor students in New York
My friend met many good linguists in Prague
Our professor visited a number of good galleries in New York

For locative adverbial modifiers of any type η we introduce the binary operation + of type $\eta(\eta\ \eta)$. The purpose of this operation is to introduce the compound system of addresses such as:

In Belgrade, Terazije street.

In the actual world he operation $\rho+\sigma$ has its full meaning when σ is a sublocation of ρ.

Example 5. In the TIL construction corresponding to the sentence:

John lives in Paris, Kennedy street

the operation + of type $\pi(\pi\ \pi)$ occurs. The atomic constructions for this sentence are:

Live / $o(\iota\ \pi)_{(\omega)}$, Paris / ι, Kennedy-Street / ι, John / ι, / $\pi(\iota)$

The whole construction reads:

$\lambda\mathbf{w}$ [Live'(John, In(Paris)+(Kennedy Street))]

The following sentence has a similar construction

He lives in a small village, the main street

$\lambda\mathbf{w}$ [So($\lambda\mathbf{w}$ [Small'(Village)])($\lambda\mathbf{w}$ [(λx [The($\lambda\mathbf{w}$ [Main'(Street)])
($\lambda\mathbf{w}$ [λy [Live'(He,In (x)+(y))]])])])]

It has been built from the following atomic constructions:

Live / $o(\iota\ \pi)_{(\omega)}$, Small / $(o(\iota)(o(\iota)_{(\omega)}))_{(\omega)}$, Main / $(o(\iota)(o(\iota)_{(\omega)}))_{(\omega)}$,

He / ι, In / $\pi(\iota)$

Street / $o(\iota)_{(\omega)}$, Village / $o(\iota)_{(\omega)}$, So / $(o(o(\iota)_{(\omega)}))(o(\iota)_{(\omega)})$,

The / $(o(o(\iota)_{(\omega)}))(o(\iota)_{(\omega)})$

Example 6. In the sentences:

In New York, every man is nervous

In this country everybody runs

the locative modifiers *in New York*, *in this country* modify the sentences:

Every man is nervous, Everybody runs

which means that the corresponding TIL constructions corresponding to the above locative modifiers are of the type $(o(o_{(\omega)})_{(\omega)}$. For the first sentence this is the construction In(New York), and for the second the the construction In(u). The atomic constructions are the following:

Run / $o(\iota)_{(\omega)}$, Is-nervous / $o(\iota)_{(\omega)}$, Person / $o(\iota)_{(\omega)}$, Man / $o(\iota)_{(\omega)}$,

Country / $o(\iota)_{(\omega)}$

New York / ι, So /$((o(o(\iota)_{(\omega)}))(o(\iota)_{(\omega)})$, This /$(o(o(\iota)_{(\omega)}))(o(\iota)_{(\omega)})$,

In /$((o(o_{(\omega)})_{(\omega)})(\iota)$

and the complete constructions:

λ**w** [(In(New York))'[λ**w** [Ev(Man)(Is-nervous)])]]

λ**w** [This(Country)(λu[(In(u))'(λ**w** [Ev(Person)(λ**w** [λx Run'(x)])])])]

Example 7. For the sentence:

In every small town, in the suburb, everybody knows everybody

the compound construction In(z)+In(u) plays the role of locative adverbial sentence modifier. The atomic constructions are the following:

Know /$o(\iota\ \iota)_{(\omega)}$, Person /$o(\iota)_{(\omega)}$, Town /$o(\iota)_{(\omega)}$, Suburb /$o(\iota\ \iota)_{(\omega)}$,

x / ι, y / ι, z / ι, u / ι

Ev / $(o(o(\iota)_{(\omega)}))(o(\iota)_{(\omega)})$, The / $(o(o(\iota)_{(\omega)}))(o(\iota)_{(\omega)})$, In $/(o(o_{(\omega)})_{(\omega)})(\iota)$

and the construction corresponding to the whole sentence:

$$\lambda\mathbf{w}\ [So(\lambda\mathbf{w}\ [Small'(Town)])(\lambda\mathbf{w}\ [\lambda z\ [The(\lambda\mathbf{w}\ [\lambda u\ [Suburb'(u,z)]])\ ((In(z)+In(u))'(\lambda\mathbf{w}\ [Ev(Person)(\lambda\mathbf{w}\ [\lambda x\ [Ev(Person)\ (\lambda\mathbf{w}\ [\lambda y\ [Know'(x,y)]])]])]))])]$$

Example 8. In the sentence:

In this country, one finds book-shops on all major streets in every town

we have both a locative sentence modifier (*in this country*) and locative inner participants (two of them: *on all major streets, in every town*). The atomic constructions are:

Find / $o(\iota\ \pi)_{(\omega)}$, Major / $(o(\iota)(o(\iota)_{(\omega)}))_{(\omega)}$

x / ι, y / ι, z / ι, u / ι, One / ι

In / $\pi(\iota)$, In / $((o(o_{(\omega)})_{(\omega)})(\iota)$

Country / $o(\iota)_{(\omega)}$, Town / $o(\iota)_{(\omega)}$, Book-shop / $o(\iota)_{(\omega)}$, Street / $o(\iota)_{(\omega)}$

So / $(o(o(\iota)_{(\omega)}))(o(\iota)_{(\omega)})$, Ev / $(o(o(\iota)_{(\omega)}))(o(\iota)_{(\omega)})$,

This / $(o(o(\iota)_{(\omega)}))(o(\iota)_{(\omega)})$

The complete construction reads:

$$This(Country)(\lambda\mathbf{w}\ [\lambda u\ [(In(u))'(\lambda\mathbf{w}\ [So(Book\text{-}shop)\ (\lambda\mathbf{w}\ [\lambda x[Ev(Town)\ (\lambda\mathbf{w}\ [\lambda y[Ev(\lambda\mathbf{w}\ [Major'(Street)])\ (\lambda\mathbf{w}\ [\lambda z\ [Find(One,\ x,\ In(y)+On(z))]])]])]])])]])$$

As far as the foundation of the *meaning* of locatives is concerned there are several relations which have been recognized and explored. The most important is the relation:

being on that and that place

which can determine:

- location of an individual (expressed by a noun or noun phrase)
- location of an action (expressed by a verb or predicate)
- location of a property (expressed by an adjective or adjective phrase)
- location of an event (expressed by a formula) etc.

On the basis of the ideas of Resher's topological logic [8] we give some of the meaning postulates which characterize the location of an event (expressed by a formula of TIL) and the location of an action (expressed by a unary verb).

If F,G are any formulae and ρ,s are locative formula modifiers (having the type $(o(o_{(\omega)}))_{(\omega)}$), then the following meaning postulates seem to be acceptable:

$$\rho'(\lambda_{\omega}(\neg F)) \Rightarrow \neg\rho\ni(\lambda_{\omega}F)$$
$$\forall\rho[\rho'(\lambda \mathbf{w}[\sigma'(\lambda \mathbf{w}\ F)])] \Leftrightarrow \sigma'(\lambda \mathbf{w}[\mathrm{Evh}'(\lambda \mathbf{w}\ F)])$$
$$\mathrm{Evh}'(\lambda \mathbf{w}\ F) \Rightarrow F$$
$$\rho'(\lambda \mathbf{w}\ [\sigma'(\lambda \mathbf{w}\ F)]) \Leftrightarrow \sigma'(\lambda \mathbf{w}\ F)$$

or, instead of the last postulate:

$$\rho'(\lambda \mathbf{w}\ [\sigma'(\lambda \mathbf{w}\ F)]) \Leftrightarrow (\rho+\sigma)'(\lambda \mathbf{w}\ F)$$

The constant locative modifier Evh, as well as the modifier Soh (meaning *everywhere*, *somewhere*), both of type $(o(o_{(\omega)}))_{(\omega)}$ are introduced by the obvious definitions:

$$\mathrm{Evh}'(\lambda \mathbf{w}\ F) \Leftrightarrow \forall\ \rho[\rho'(\lambda \mathbf{w}\ F)],\ \ \mathrm{Soh}'(\lambda \mathbf{w}\ F) \Leftrightarrow \exists\ \rho[\rho'(\lambda \mathbf{w}\ F)]$$

The only rule of interference is:

$$\frac{\mathbf{F}}{\rho'(\lambda \mathbf{w}\mathbf{F})}$$

It does not seem acceptable to have as a meaning postulate the converse of the implication:

$$\rho'(\lambda \mathbf{w}(\neg F)) \Rightarrow \neg\rho'(\lambda \mathbf{w}\ F)$$

The reason is the following: If it is the case $\neg\rho'(\lambda \mathbf{w}\ F)$, i.e. if it is not the case $\rho'(\lambda \mathbf{w}\ F)$, then one possibility is that $\rho'(\lambda \mathbf{w}(\neg F))$ (i.e. on location ρ takes place $\neg F$), but it could also happen that F takes place on some other location which is not on ρ.

For the location of an action the corresponding meaning postulates would be the following

$$\forall\ \rho[\sigma'(\lambda \mathbf{w}\ (\rho'(X))(\nu\ ))] \Leftrightarrow \sigma'(\lambda \mathbf{w}\ (\mathrm{Evh}'(X))(\nu\ )$$
$$(\mathrm{Evh}'(X)(\nu\ )) \Rightarrow X(\nu\ )$$
$$\rho'(\lambda \mathbf{w}\ [\sigma'(X)])(y) \Leftrightarrow \sigma'(X)(y)$$

or, instead of the last postulate:

$$\rho'(\lambda \mathbf{w}\ [\sigma'(X)])(y) \Leftrightarrow (\rho+\sigma)'(X)(y)$$

where ρ, σ are variables having the type of adverbial verb modifiers $((o(\iota))(o(\iota)_{(\omega)}))_{(\omega)}$, X is a variable denoting an unary verb, ρ' is already introduced abbreviation for $\rho(w)$ (i.e for ρ(w)(t)(d)(p)(b)(a), where a, b, p, d, t, w are variables of types α, β, π, δ, τ, ω respectively), y is a variable of type ι, Evh (as well as Soh) is a constant adverbial verb modifier meaning *everywhere* (*somewhere*) introduced by definitions:

$$\mathrm{Evh}'(X)(y) \Leftrightarrow \forall\, \rho[\rho'(X)(y)],\ \ \mathrm{Soh}'(X)(y) \Leftrightarrow \exists\, \rho[\rho'(X)(y)]$$

The corresponding rule of interference is:

$$\frac{X(y)}{\rho'(X)(y)}$$

Adding some new meaning postulates, for example those related to the Boolean or non Boolean conjunction of locative modifiers, requires much more investigation in the light of results [1], [4].

REFERENCES

[1] Jaap van der Does, *Applied Qualifier Logic*,
Haveka, B.V. Abblasserdam, 1992.

[2] E. Koktová, *Are local adverbs derived from temporal sources ?*,
Prague Bulletin of Mathematical Linguistics 27 (1977), 45-63

[3] Е Кондрашина, Л.В. Литвицева, Д.А. Поспелов, Представление знаний о времени и пространстве в интелектуалъньіх системах,
Наука, Москва,1989.

[4] Manfred Krifka, *Boolean and non-Boolean 'AND'*,
in: *Papers from the Second Symposium on Logic and Language* held in Hajdúszobszló 1989, Akadémiai Kiadó, Budapest, 161-188

[5] Jovanka Z. Lazarevska, *Predlozi za mesto u engleskom i makedonskom jeziku* (magistarski rad), Beograd, 1993.

[6] Richard Montague, *Formal Philosophy* (selected papers),
R. Thomason (ed.), London, 1974.

[7] Marica D. Prešić, *Intensional logic with deep cases*, Publ. Ins. Math.
Beograd, t. 54 (68) (1993), 1-12

[8] N. Rescher, *Topological Logic, Topics in Philosophical Logic* , B. Reidel

Publishing company, Dordrecht, 229-244, 1968.
[9] Petr Sgal, Eva Hajičová, Jarmila Panevová, *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects,* Academia, Prague, 1986.
[10] P. Tichy, *The logic of temporal discourse,* Linguistics and Philosophy, 3 (1980), 343-369

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 55-59

ULTRAFILTER CARDINALITY JUMPS AND THE CONTINUUM PROBLEM

Aleksandar JOVANOVIĆ

Institute of Mathematics,
Knez Mihajlova 35, 11000 Beograd, Yugoslavia

Abstract. The continuum problem and the existence of ultrafilters with jumping cardinality over smaller cardinals are related via the two cardinal problems.

Keywords: continuum problem, nonregular ultrafilters, two cardinal problem.

In ZFC all cardinals are on the aleph line which can be expressed by

$$2^{\aleph_\alpha} = \aleph_{\alpha+f(\alpha)},\ \alpha \in Ord,$$

where $f : Ord \longrightarrow Ord$, (we call it a continuum displacement function). By Cantor's theorem $f(\alpha) \geq 1$, $\alpha \in Ord$. Also f has to by monotonic and $cf(\aleph_{\alpha+f(\alpha)}) > \aleph_\alpha$. The behavior of f has been investigated in detail; we list some exciting possibilities:

$(\forall\alpha)f(\alpha) = 1 \leftrightarrow GCH$, is consistent with ZFC, Gödel;

$f(0) = 2$, consistent with ZFC, Cohen '63;

"On regular cardinals f can have arbitrary values", (respecting monotony and the cofinality condition), Easton '67;

$ZFC + \kappa$ is a measurable cardinal $\vdash (\forall\alpha < \kappa)f(\alpha) = 1 \rightarrow f(\kappa) = 1$, Scott '61;

$ZFC \vdash (\forall \alpha < \omega_1) f(\alpha) = 1 \rightarrow f(\omega_1) = 1$, Silver '74;

$ZFC \vdash (\forall \alpha < \omega) f(\alpha) = 1 \rightarrow f(\omega) \leq 4$, the same holds for all singular cardinals of the uncountable cofinality;

$ZFC + 2^{\aleph_0}$ is $RVM \vdash (\forall \kappa < 2^{\aleph_0}) 2^\kappa = 2^{\aleph_0}$, Prikry '74

If $2^{\aleph_0}$ is a real valued measurable then it is weakly inaccessible, hence

$$2^{\aleph_0} = \omega_{f(0)} = \omega_{2^{\aleph_0}}, \text{ i.e. } f(0) = 2^{\aleph_0}.$$

$(\forall n) f(n) = 1 \wedge f(\omega) > 1$, consistent from the existence of huge cardinals, Magidor '77;

Chang's conjecture implies: if $\aleph_1$ is a strong limit then $2^{\aleph_{\omega_1}} < \aleph_{\omega_2}$, Magidor '77;

There is an $\aleph_2$ saturated ideal over ω_1 which implies $2^{\aleph_0} = \aleph_1 \rightarrow 2^{\aleph_1} = \aleph_2$, Prikry, Jech '74;

When $f(\sigma) < 2^{\aleph_\sigma}$ we say that f is bounded at σ, and when $f(\sigma) = 2^{\aleph_\sigma}$ we say that f is (very) unbounded at σ. Let T be a first order theory with unary predicate symbol U. T admits pair (α, β) if there is a model of T of cardinality α in which the cardinality of U is β.

Definition 1. *The pair (α, β) is (when exists) :*

Left Large Gap (LLG) for T iff T admits (α, β) and does not admit any (α', β) for $\alpha' \geq \alpha$.

Right Large Gap (RLG) for T iff T admits (α, β) and does not admit any (α, β') for $\beta' \leq \beta$.

Large Gap (LG) for T iff (α, β) is LLG and RLG.

It is clear that we can correspond partial cardinal functions Λ_i, $i \in 2$, to any theory T, so that (when exists):

$(\Lambda_0(\kappa), \kappa)$ is LLG;

$(\kappa, \Lambda_1(\kappa))$ is RLG;

for all κ.

Lemma 1. *Let Λ_0 and Λ_0' be cardinal operations defining LLG for theories T_1 and T_2 , for many κ. If one of the following holds:*

1. GCH;

2. Λ_0 is monotonous;

3. $|\Lambda_0'(\kappa)|^\omega = \Lambda_0'(\kappa)$;

then there is a theory T such that $((\Lambda_0 \circ \Lambda_0')(\kappa), \kappa)$ is LLG for T. Similar is true for RLG.

Proof. Let $\mathfrak{A} = \langle A, V, ...\rangle$ be a model for T_1 such that V is an interpretation of predicate symbol Q and $|V| = \Lambda_0'(\kappa)$, $|A| = (\Lambda_0 \circ \Lambda_0')(\kappa)$.

Let $\mathfrak{B} = \langle B, U, ...\rangle$ be a model for T_2 such that U is an interpretation of predicate symbol P and $|U| = \kappa$, $|B| = \Lambda_0'(\kappa)$. We may suppose that $L_{T_1} \cap L_{T_2} = \emptyset$ and that T_2 has universally closed axioms. Consider the extension T of T_1 obtained in the following way. First, take T' to be a theory in the language $L_{T_1} \cup L_{T_2}$ with the axioms of T_1. Extend T' to T adding interpretations of axioms of T_2 in the language $L_{T_1} \cup L_{T_2}$. The interpretation is defined in the following way. On L_{T_2} it is the identity. In the axioms of T_2 every subformula of the form $(\exists x)\varphi$ is replaced with the formula $(\exists x)Q(x) \wedge \varphi$. The universe of the interpretation is Q, so we introduce the axiom $(\exists x)Q(x)$. If F is an n-ary function symbol of the language then the axiom of T is formula $Q(x_1) \wedge ... \wedge Q(x_n) \rightarrow Q(F(x_1, ..., x_n))$. Using the bijection $|B| = |V|$, a model $\mathfrak{A}'$, an expansion of $\mathfrak{A}$, is constructed. Let f be such a bijection. Extend f to an isomorphism. For $c \in L_{T_2}$ define $c^{\mathfrak{A}} = f(c)$. If F is a function symbol in L_{T_2} define:

$$F^{\mathfrak{A}}(a_1, ..., a_n) = \begin{cases} F^{\mathfrak{B}}(f^{-1}(a_1), ..., f^{-1}(a_n)) & \text{if } a_1, ..., a_n \in V, \\ \text{arbitrary} & \text{otherwise.} \end{cases}$$

If R is a predicate symbol in L_{T_2} define the interpretation of R

$$R^{\mathfrak{A}}(a_1, ..., a_n) \text{ iff } R^{\mathfrak{B}}(f^{-1}(a_1), ..., f^{-1}(a_n))$$

Now consider predicate symbol P in L_T and model $\mathfrak{A}'' = \langle A, U, ...\rangle$ for T. $\mathfrak{A}''$ is $((\Lambda_0 \circ \Lambda_0')(\kappa), \kappa)$ model. Let $\mathfrak{A} = \langle A, V, U, ...\rangle$ be some model for T with $|U| = \kappa$. Let V and U be interpretations for Q and P, respectively. Let $L_{T_2'}$ be (the reduction of L_T) $\{Q\} \cup L_{T_1}$ and let T_2' have the interpretations of axioms of T as the only axioms. It follows that $|V| \le \Lambda_0'(\kappa)$. Using the hypothesis in a similar way, we get $|A| \le (\Lambda_0 \circ \Lambda_0')(\kappa)$.

Corollary 1. *Let T_1 and T_2 have $(\Lambda(\kappa), \kappa))$ and $(\kappa, \Gamma(\kappa))$ as LLG and RLG, for many κ, respectively, Then there is theory T for which $(\Lambda(\kappa), \Gamma(\kappa))$ is LG.*

Corollary 2. *With the hypothesis of the corollary the following is true for any ultrafilter D:*

$$\Gamma(|\prod_D \kappa|) \le |\prod_D \Gamma(\kappa)| \le |\prod_D \kappa| \le |\prod_D \Lambda(\kappa)| \le \Lambda(|\prod_D \kappa|)$$

There are theories [1] T_1 and T_2 such that for all κ:

$$(\kappa^+, \kappa) \textit{ is } LLG \textit{ for } T_1; \;\; (2^\kappa, \kappa) \textit{ is } LLG \textit{ for } T_2.$$

From **Lemma 1.** we get:

> for all λ, $(\omega_n(\lambda), \lambda)$ and $(\mathfrak{I}_n(\lambda), \lambda)$ are LLG; if Γ is any finite combination of $\omega_n()$, $\mathfrak{I}_n()$ then for all λ, $(\Gamma(\lambda), \lambda)$ is LLG.

Lemma 2. *Let D be a uniform ultrafilter over λ. If $\lambda^{<\lambda} = \lambda$ then* $|\prod_D \lambda| = 2^\lambda$.

For any uniform ultrafilter D over a cardinal α define its cardinal trace by:

$$ctr_D = \{\ \lambda \mid \omega \le \lambda \le 2^\alpha \ \wedge \ \lambda = |\prod_D \alpha_i|, \textit{ for some } \alpha_i \le \alpha,\ i \in \alpha\ \}.$$

If $|ctr_D| \ge 1$ we say that D has cardinal jumps or D is jumping. The first examples of jumping ultrafilters over small cardinals are the nonregular ultrafilters of Magidor [2].

Let $(\Gamma(\kappa), \kappa) = (\omega_\xi(\kappa), \kappa) = (\omega_\sigma, \kappa)$ be a LLG for some theory T. Let D be a uniform ultrafilter over ω_σ with jumps after κ, i.e.

$$|\prod_D \kappa| \ < \ |\prod_D \omega_\sigma|,$$

and let $\omega_\sigma^{<\omega_\sigma} = \omega_\sigma$. Then using the above conditions and **Corollary 2.**

$$\begin{aligned} |\prod_D \kappa| = \omega_\eta < \ |\prod_D \omega_\sigma| \ &= 2^{\omega_\sigma} = \omega_{\sigma+f(\sigma)} \\ &= |\prod_D \omega_\xi(\kappa)| \ \le \ \omega_\xi(|\prod_D \kappa|) = \omega_{\eta+\xi}, \end{aligned}$$

$$\text{i.e. } \eta < \sigma + f(\sigma) \le \eta + \xi \le \eta + \sigma$$

this means that $2^{\aleph_\sigma}$ is limited by the sum of $\min ctr_D$ and the length ξ of LLG, thus the continuum function is well-bounded. On the contrary, the large $f(\sigma)$ reduces the possibility of jumps, for instance $\max\{\omega, |\sigma|\} <$

$|f(\sigma)|$, makes $\eta < f(\sigma) \leq \eta + \sigma$ impossible. We can formulate the former as follows.

Lemma 3. *Let D be a uniform ultrafilter over λ, jumping after κ. If there is a theory for which $(\Gamma(\alpha), \alpha)$ is LLG for sufficiently many α and $\Gamma(\kappa) = \lambda$ then $\lambda^{<\lambda} = \lambda$ implies $f(\lambda)$ is well-bounded.*

Some examples with the above notation:

1. (D nicely separates κ and ω_σ) $|\prod_D \kappa| \leq \omega_\sigma \rightarrow f(\sigma) \leq \sigma$.

2. $(\forall \xi \in \omega)(\forall \kappa)$ $(\omega_\xi(\kappa), \kappa)$ is LLG. Let $\omega_\xi(\kappa)^{<\omega_\xi(\kappa)} = \omega_\xi(\kappa)$. If there is a uniform ultrafilter over $\omega_\xi(\kappa)$ jumping after κ then $f(\sigma)$ is a successor ordinal and well-bounded.

3. Let D be a jumping ultrafilter over $\aleph_{15}$ and $\aleph_{15}^{<\aleph_{15}} = \aleph_{15}$.

   (a) If $|\prod_D \omega| \leq \aleph_{15}$ then $2^{\aleph_{15}} \leq \aleph_{30}$.

   (b) If $2^{\aleph_{15}} = \aleph_{\omega+1}$ then there is no jumping ultrafilter over $\aleph_{15}$.

4. Let $2^{\aleph_{15}} = \aleph_{\omega_1+1}$ and $\aleph_{15}^{\aleph_{14}} = \aleph_{15}$. Then if there is a jumping ultrafilter over $\aleph_{15}$, there would have to be $|\prod_D \omega| = \aleph_{\omega_1}$, thus singular. Is it possible?

Questions.

1. Is condition $\lambda^{<\lambda} = \lambda$ in **Lemma 2.** necessary?

2. Can **Lemma 1.** be generalized to combine more than finitely many LLG's into LLG?

3. Are there examples of LLG $(\Gamma(\kappa), \kappa)$ with $\Gamma(\kappa) \geq \aleph_\omega(\kappa)$ for sufficiently many κ?

REFERENCES

[1] C. C. Chang, H. J. Keisler, *Model Theory*, North Holland, Amsterdam, 1973.

[2] M. Magidor, *On the existence of nonregular ultrafilters and the cardinality of ultrapowers*, Math. Tech Report 176, 1977.

[3] Ž. Mijajlović, Doctoral Dissertation, University of Belgrade, Faculty of Sciences, 1976.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 61-63

ON A PROBLEM OF MATHIAS

Ilijas FARAH
Department of Mathematics, University of Toronto,
Toronto, Canada M5S 3G3
Institute of Mathematics, Serbian Academy of Science and Art,
Kneza Mihajla 35, 11 000 Beograd, Yugoslavia

The main purpose of this note is to give an answer to a question of A.R.D. Mathias (see [3; 8.6]). We assume that the reader is familiar with the basic theory of Forcing. Recall that for a filter $\mathcal{U}$ on the set of positive integers $\mathbf{N}$ the *Prikry poset*, $\mathcal{P}_{\mathcal{U}}$, is defined as follows: a typical condition is $\langle s, A\rangle$, where s is a finite set, and A is an element of $\mathcal{U}$. The ordering is defined by

(1) $\langle s, A\rangle \leq \langle t, B\rangle$ if and only if t is an initial segment of s, $A \supseteq B$, and $s \setminus t$ is included in B.

Recall that $[A]^2$ denotes the family of all unordered pairs of elements of the set A. An ultrafilter $\mathcal{U}$ on $\mathbf{N}$ is *selective* (or *Ramsey*) if and only if for every partition of the family $[\mathbf{N}]^2$ into two pieces there is an $A \in \mathcal{U}$ which is *homogeneous* for this partition, i.e. the set $[A]^2$ is included in a single piece. The following remarkable theorem was proved by Mathias in [3; Theorem 2.0]:

Theorem 1. *If M is a transitive model of a large enough fragment of Set Theory and $\mathcal{U}$ is a selective ultrafilter in M, then a set X is $\mathcal{P}_{\mathcal{U}}$-generic over M if and only if it is included modulo finite in every $A \in \mathcal{U} \cap M$.*

This result is the main ingredient of the proof that in Solovay's model in which all sets of reals are Lebesgue measurable the partition relation $\omega \to (\omega)^\omega$ holds; see [3]. In [3; 8.6] Mathias asks the following (recall that L denotes the constructible universe):

Question 1. Is there a criterion similar to that of Theorem 2.0 for a pair (X, Y) of reals to be $\mathcal{P}_{\mathcal{U}} \times \mathcal{P}_{\mathcal{U}}$-generic over L?

The motivation for this question comes from the Recursion Theory, and the question about the complexity of the set of all hyperarithmetically encodable subsets of $\mathbf{N}$; see [3; page 101]. Below we will give such a criterion, even without using an assumption that the ground model is the subuniverse of all constructible sets. Let $\mathcal{C}$ denote the standard poset for adding a Cohen real. We say that posets $\mathcal{P}$ and $\mathcal{Q}$ are *equivalent* if their regular open algebras are isomorphic, i.e. if they yield the same forcing extensions. By [1; VII.7.11], if $\mathcal{Q}$ is *densely embedded* into $\mathcal{P}$ (see [1; VII.7.7]) then $\mathcal{P}$ and $\mathcal{Q}$ are equivalent.

Theorem 2. *For every filter $\mathcal{U}$, poset $\mathcal{P_U} \times \mathcal{P_U}$ is equivalent to $\mathcal{P_U} \times \mathcal{C}$.*

Proof. Let $\mathcal{PP_U}$ be the following poset: a typical condition is $p = \langle s_p, t_p, A_p \rangle$, where $\langle s_p, A_p \rangle$ and $\langle t_p, A_p \rangle$ are conditions in $\mathcal{P_U}$ and the ordering is defined by $p \leq q$ if

$$\langle s_p, A_p \rangle \leq_{\mathcal{P_U}} \langle s_q, A_q \rangle \text{ and } \langle t_p, A_p \rangle \leq_{\mathcal{P_U}} \langle t_q, A_q \rangle. \tag{2}$$

Obviously $\mathcal{PP_U}$ is densely embedded into $\mathcal{P_U} \times \mathcal{P_U}$ by

$$\langle s, t, A \rangle \mapsto \langle \langle s, A \rangle, \langle t, A \rangle \rangle.$$

Now consider the product $\mathcal{P_U} \times \mathcal{C}$. We can assume $\mathcal{C}$ is the poset of all finite partial functions from $\mathbf{N}$ into $\{0, 1, 2\}$ ordered by the extension, because all nonatomic countable posets are equivalent (see [1; VII.(C3)]). Let $\mathcal{PC_U}$ denote the set of all conditions in $\mathcal{P_U} \times \mathcal{C}$ of the form

$$p = \langle \langle s, A \rangle, \sigma \rangle, \qquad \text{where } \mathrm{dom}\sigma = |s|.$$

Obviously $\mathcal{PC_U}$ is dense in $\mathcal{P_U} \times \mathcal{C}$, so it suffices to prove that it is isomorphic to $\mathcal{PP_U}$. The mapping from $\mathcal{PC_U}$ into $\mathcal{PP_U}$ defined by

$$\langle \{n_1, \ldots, n_k\}, A, \sigma \rangle \mapsto \langle \{n_i | \sigma(i) \neq 0\}, \{n_i | \sigma(i) \neq 1\}, A \rangle$$

is the desired isomorphism. This proves theorem.

Corollary. *A finite product of two or more copies of $\mathcal{P_U}$ is equivalent to $\mathcal{P_U} \times \mathcal{P_U}$.*

Note that the proof of Theorem 2 gives the following more precise statement (χ_X is a characteristic function of the set X):

Theorem 3. *If M is a transitive model of a large enough fragment of Set Theory and $\mathcal{U}$ is a selective ultrafilter in M, then a pair $\langle X, Y \rangle$ is $\mathcal{P_U} \times \mathcal{P_U}$-generic over M if and only if:*

(a) $X \cup Y$ *is* $\mathcal{P}_{\mathcal{U}}$*-generic over* M, *and*

(b) function $\chi_X + \chi_Y : X \cup Y \to \{0,1,2\}$ *is* $\mathcal{C}$*-generic over* $M[X \cup Y]$.

So by using Theorem 1 and the well-known characterization of Cohen-generic reals (see [2]), we can answer the above question of Mathias.

Corollary. *If* M *is a transitive model of a large enough fragment of Set Theory and* $\mathcal{U}$ *is a selective ultrafilter in* M, *then a pair* $\langle X, Y\rangle$ *is* $\mathcal{P}_{\mathcal{U}} \times \mathcal{P}_{\mathcal{U}}$*-generic over* M *if and only if:*

(a) $X \cup Y$ *is included modulo finite in very* $A \in \mathcal{U} \cap M$, *and*

(b) function $\chi_X + \chi_Y : X \cup Y \to \{0,1,2\}$ *avoids all meager* F_σ *subsets of* $\{0,1,3\}^{X\cup Y}$ *coded in* $M[X \cup Y]$.

REFERENCES

[1] K. Kunen, *An Introduction to Independence Proofs*, North–Holland, 1980.

[2] A. Kanamori, *The Higher Infinite*, Springer–Verlag, 1995.

[3] A.R.D. Mathias *Happy families*, Ann. Math. Logic 12 (1977), 59–111.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 65-73

SOME TOPOLOGICAL PROPERTIES OF THE LAMBDA CALCULUS

Silvia GHILEZAN
Faculty of Engineering, University of Novi Sad
Trg Dositeja Obradovića 6, 21000 Novi Sad, Yugoslavia
e-mail: gsilvia@uns.ns.ac.yu

Abstract. We give an overview and a comparison of topologies in the untyped lambda calculus. The basic operation of the lambda calculus, application, is continuous, unsolvable terms are compactification points and normal forms are isolated points with respect to the topologies considered. Similarities and differences between these topologies will be pointed out as well.

Keywords: lambda calculus

1. INTRODUCTION

The *type topology* in untyped lambda calculus is introduced via typability in intersection type systems. Intersection type systems are extensions of simply typed lambda calculus with an additional type forming operator, intersection, and a special type ω which is given the property to type every lambda term. The sets $\mathcal{V}_{\Gamma,\sigma} = \{N \in \Lambda \mid \Gamma \vdash N : \sigma\}$, where $\Gamma \vdash N : \sigma$ means that the lambda term N is typable by the type σ in the context Γ in the intersection type system, form a basis for this topology. The set $\mathcal{V}_{\Gamma,\sigma}$ consists of all lambda terms typable by the same type σ in the given context Γ.

Tree topology is based on the notion of Böhm trees (see [1]). The tree topology on the set of all untyped lambda terms Λ is the smallest one that makes the mapping $BT : \Lambda \to \mathcal{B}$ continuous, $\mathcal{B}$ is the set of all Böhm-like trees, meaning that the open sets in Λ are of the form $BT^{-1}(\mathcal{O})$, with $\mathcal{O}$ open with reference to (w.r.t.) the Scott topology on $\mathcal{B}$. The sets $\mathcal{O}_{M,k} =$

$\{N \in \Lambda \mid M^{(k)} \sqsubseteq N\}$, form a basis w.r.t. this topology on Λ, where $M^{(k)}$ is the lambda term reconstructed from $BT^k(M)$, the Böhm tree of the lambda term M of length k. The set $\mathcal{O}_{M,k}$ consists of all lambda terms whose Böhm trees contain the same Böhm tree of $M^{(k)}$.

Once again intersection type systems are tied with topology. This time it is the *filter topology* on the set of all untyped lambda terms Λ. The set of filters $\mathcal{F}$ is defined on the set of intersection types. The filter topology is the smallest topology that makes the mapping $\| \quad \| : \Lambda \to \mathcal{F}$, where $\|M\| = \{\sigma | \Gamma \vdash M : \sigma\}$, continuous. The open sets in Λ are of the form $\|\mathcal{O}\|^{-1}$, with $\mathcal{O}$ open w.r.t. the Scott topology on the set of filters $\mathcal{F}$.

Using topological tools, basic lambda calculus concepts can be expressed topologically. Unsolvable terms are compactification points and normal forms are isolated points. The basic operation of the lambda calculus, application, is continuous, w.r.t. these topologies.

In Section 2 the type topology is presented. In Section 3 topological properties of the lambda calculus are investigated. Section 4 is a brief overview of two other topologies, tree topology and filter topology. Similarities and differences between them are discussed.

2. TYPE TOPOLOGY

The types of an intersection type assignment system are propositional formulae with the connectives $\to$ and $\cap$, where $\cap$ is a *specific conjunction*, called *intersection*, whose properties are in accordance with its interpretation as an intersection of types. The basic notions and properties of intersection type assignment systems are given in [3] and can be found in the surveys of typed lambda calculi in [2] and [5].

The set of intersection types T is defined in the following way:

Definition 1. *(i)* $V = \{\alpha, \beta, \gamma, \alpha_1, \ldots\} \subset T$ *(V is a denumerable set of propositional variables).*

(ii) $\omega \in T$.

(iii) *If* $\sigma, \tau \in T$, *then* $(\sigma \to \tau) \in T$.

(iv) *If* $\sigma, \tau \in T$, *then* $(\sigma \cap \tau) \in T$.

Let $\alpha, \beta, \gamma, \alpha_1, \ldots$ be schematic letters for type variables, and let $\delta, \epsilon, \sigma, \tau$, $\delta_1, \ldots$ be schematic letters for types.

Definition 2. *(i) A pre-order $\leq$ is introduced on T in the following way:*

1. $\sigma \leq \sigma$	*5.* $(\sigma \to \rho) \cap (\sigma \to \tau) \leq \sigma \to (\rho \cap \tau)$
2. $\sigma \leq \tau, \quad \tau \leq \rho \Rightarrow \sigma \leq \rho$	*6.* $\sigma \cap \tau \leq \sigma, \quad \sigma \cap \tau \leq \tau$
3. $\sigma \leq \omega$	*7.* $\sigma \leq \tau, \quad \sigma \leq \rho \Rightarrow \sigma \leq \tau \cap \rho$
4. $\omega \leq \omega \to \omega$	*8.* $\sigma \leq \sigma_1, \quad \tau \leq \tau_1 \Rightarrow \sigma_1 \to \tau \leq \sigma \to \tau_1$.

(ii) $\sigma \sim \tau$ *if and only if* $\sigma \leq \tau$ *and* $\tau \leq \sigma$.

Let Λ be the set of *untyped (type-free) lambda terms*, and let $x, y, z, x_1, \ldots$ be schematic letters for term variables and $M, N, P, Q, M_1, \ldots$ schematic letters for lambda terms. The expression $M : \sigma$, called a *statement*, where $M \in \Lambda, \sigma \in T$ links the terms of Λ and the types of T. M is the *subject* and σ is the *predicate* of the statement $M : \sigma$. If $x \in V$, then $x : \tau$ is a *basic statement*. A *context* is a set of basic statements. $\Gamma, \Delta, \Gamma_1, \ldots$ are used as schematic letters for contexts. **B** will denote the set of all contexts.

Definition 3. *The following rules determine the intersection type assignment system* $\mathcal{D}\Omega_{\leq}$*:*

$$(start \;\; rule) \quad \frac{(x : \sigma) \in \Gamma}{\Gamma \vdash x : \sigma};$$

$$(\to E) \quad \frac{\Gamma \vdash M : \sigma \to \tau \quad \Gamma \vdash N : \sigma}{\Gamma \vdash MN : \tau}; \quad (\to I) \quad \frac{\Gamma, x : \sigma \vdash M : \tau}{\Gamma \vdash (\lambda x.M) : \sigma \to \tau};$$

$$(\cap E) \quad \frac{\Gamma \vdash M : \sigma \cap \tau}{\Gamma \vdash M : \sigma}, \quad \frac{\Gamma \vdash M : \sigma \cap \tau}{\Gamma \vdash M : \tau}; \quad (\cap I) \quad \frac{\Gamma \vdash M : \sigma \quad \Gamma \vdash M : \tau}{\Gamma \vdash M : \sigma \cap \tau};$$

$$(\omega) \quad \frac{}{\Gamma \vdash M : \omega};$$

$$(\leq) \quad \frac{\Gamma \vdash M : \sigma \quad \sigma \leq \tau}{\Gamma \vdash M : \tau}.$$

Some examples of typability in this system are:

$$\vdash \lambda x.xx : (\sigma \cap (\sigma \to \tau)) \to \tau;$$

$$\vdash \Omega \equiv (\lambda x.xx)(\lambda x.xx) : \omega.$$

Every term in normal form has a principal type in the system $\mathcal{D}\Omega_{\leq}$. The most important property of the principal type of a normal form N is that all other types that type N can be obtained from the principal type by some operations on types (see [6]). We will need later the following property, which is proved in [5].

Proposition 1. *(i) Let* N *be a normal form, and let* $\Gamma \vdash N : \pi$ *be a principal typing of* N*. Then, if* $\Gamma \vdash M : \pi$*, there is an* η*-reduct* Q *of* N *such that* $M \twoheadrightarrow_\beta Q$*.*

(ii) *Let N be a $\beta\eta$–normal form, and let $\Gamma \vdash N : \pi$ be a principal typing of N. If $\Gamma \vdash M : \pi$, then $M \twoheadrightarrow_\beta N$.*

It is possible to introduce a topology on the set of lambda terms Λ via the typability of lambda terms in $\mathcal{D}\Omega_{\leq}$. We shall call it *type topology.*

Let us consider a set of all lambda terms that can be typable by the same type in the same context, say

$$\mathcal{V}_{\Gamma,\sigma} = \{M \in \Lambda | \Gamma \vdash M : \sigma, \}.$$

If Γ is empty, meaning that the terms typable by σ are closed, i.e. do not contain free variables, then we will write just $\mathcal{V}_\sigma$.

$\Gamma \cup \Delta$ will denote the context obtained from the contexts Γ and Δ, such that $x : \sigma \in \Gamma \cup \Delta$ if and only if $x : \sigma \in \Gamma$, or $x : \sigma \in \Delta$, or $x : \sigma' \in \Gamma$, $x : \sigma'' \in \Delta$ and $\sigma \equiv \sigma' \cap \sigma''$.

Lema 1. *(i) σ is inhabited if and only if $\mathcal{V}_{\Gamma,\sigma} \neq \emptyset$ for some context Γ.*

(ii) $\mathcal{V}_{\Gamma,\sigma} \cap \mathcal{V}_{\Delta,\tau} \subseteq \mathcal{V}_{\Gamma\cup\Delta,\sigma\cap\tau}$.

(iii) $\mathcal{V}_{\Gamma,\omega} = \Lambda$ for any context Γ.

Proof.

(i) Obvious.

(ii) From $M \in \mathcal{V}_{\Gamma,\sigma} \cap \mathcal{V}_{\Delta,\tau}$ it follows that $\Gamma \vdash M : \sigma$ and $\Delta \vdash M : \tau$ for some contexts Γ and Δ. Obviously, $\Gamma \cup \Delta \vdash M : \sigma$ and $\Gamma \cup \Delta \vdash M : \tau$. Hence, by $(\cap I)$ we obtain that $\Gamma \cup \Delta \vdash M : \sigma \cap \tau$. The converse does not hold, e.g. if $x : \sigma \vdash x : \sigma$, $x : \tau \vdash x : \tau$, then $x : \sigma \cap \tau \vdash x : \sigma$, $x : \sigma \cap \tau \vdash x : \tau$ and hence $x : \sigma \cap \tau \vdash x : \sigma \cap \tau$ by $(\cap I)$. However, from $x : \sigma \cap \tau \vdash x : \sigma \cap \tau$ we cannot conclude $x : \sigma \vdash x : \sigma$ and $x : \tau \vdash x : \tau$.

(iii) Obvious, since every lambda term is trivially typable by ω in any context.

Proposition 2. *The family $\{\mathcal{V}_{\Gamma,\sigma}\}_{\Gamma\in\mathbf{B},\sigma\in T}$ form a basis for a topology on Λ.*

Proof.

(i) Every lambda term is typable in $\mathcal{D}\Omega_{\leq}$; so for every lambda term $M \in \Lambda$ there is a context Γ and a type $\tau \in T$ such that $\Gamma \vdash M : \tau$. That is $M \in \mathcal{V}_{\Gamma,\tau}$.

(ii) For every two sets $\mathcal{V}_{\Gamma,\sigma}$ and $\mathcal{V}_{\Delta,\tau}$, we have that $\mathcal{V}_{\Gamma,\sigma} \cap \mathcal{V}_{\Delta,\tau} \subseteq \mathcal{V}_{\Gamma\cup\Delta,\sigma\cap\tau}$, according to Lemma 1(ii).

The introduced topology will be called *type topology. Open sets* in the type topology are defined in the usual way:

Definition 4. *A set $\mathcal{O} \subseteq \Lambda$ is open if for any term $M \in \mathcal{O}$ there is a set $\mathcal{V}_{\Gamma,\rho}$ such that*

$$M \in \mathcal{V}_{\Gamma,\rho} \quad \textit{and} \quad \mathcal{V}_{\Gamma,\rho} \subseteq \mathcal{O}.$$

3. TOPOLOGICAL PROPERTIES

Familiar lambda calculus concepts can be expressed topologically. Now we can show the continuity of application with respect to the type topology. We will show the following properties of the type topology:

- Application is continuous with respect to the type topology.
- Unsolvable terms are compactification points (bottoms).
- $\beta\eta$–normal forms are isolated points.

Theorem 1. (Continuity Theorem)
Given $F \in \Lambda$. Then the map $M \mapsto FM$ is continuous (w.r.t. the type topology).

Proof. We have to show that

$$(\forall \mathcal{V}_{\Gamma,\varepsilon} \ni FM)(\exists \mathcal{V}_{\Gamma,\delta} \ni M)(Q \in \mathcal{V}_{\Gamma,\delta} \Rightarrow FQ \in \mathcal{V}_{\Gamma,\varepsilon}).$$

If $FM \in \mathcal{V}_{\Gamma,\varepsilon}$, then $\Gamma \vdash FM : \varepsilon$. By the structural property of the intersection type system $\mathcal{D}\Omega_{\leq}$ given in [3] there is a type δ such that

$$\Gamma \vdash F : \delta \to \varepsilon \text{ and } \Gamma \vdash M : \delta.$$

Hence, there is $\mathcal{V}_{\Gamma,\delta}$ to which M belongs. If $Q \in \mathcal{V}_{\Gamma,\delta}$, i.e., $\Gamma \vdash Q : \delta$, then by $(\to E)$ we have that $\Gamma \vdash FQ : \varepsilon$.

Let us consider some properties of the introduced topology that are related to β- and η- reduction.

Lema 2. *Let $M, N \in \Lambda$.*

(i) If $M \twoheadrightarrow_\beta N$, then $\forall \sigma \in T\ \forall \Gamma \in \mathbf{B}\ (M \in \mathcal{V}_{\Gamma,\sigma} \Leftrightarrow N \in \mathcal{V}_{\Gamma,\sigma})$.

(ii) If $M \twoheadrightarrow_\eta N$, *then* $\forall \sigma \in T\ \forall \Gamma \in \mathbf{B}\ \ (M \in \mathcal{V}_{\Gamma,\sigma} \Rightarrow N \in \mathcal{V}_{\Gamma,\sigma})$.

(iii) If $N \in \Lambda$ *is a* $\beta\eta$*–normal form, then*

$$M \twoheadrightarrow_\beta N \ \textit{ if and only if } \forall \sigma \in T\ \forall \Gamma \in \mathbf{B}\ (M \in \mathcal{V}_{\Gamma,\sigma} \Leftrightarrow N \in \mathcal{V}_{\Gamma,\sigma}).$$

Proof.

(i) The subject reduction property holds for the intersection type system and therefore if $M \twoheadrightarrow_\beta N$, then $\forall \sigma \in T\ \forall \Gamma \in \mathbf{B}\ \ (M \in \mathcal{V}_{\Gamma,\sigma} \Rightarrow N \in \mathcal{V}_{\Gamma,\sigma})$. But it implies that $\forall \sigma \in T\ \forall \Gamma \in \mathbf{B}\ (N \in \mathcal{V}_{\Gamma,\sigma} \Rightarrow M \in \mathcal{V}_{\Gamma,\sigma})$ as well, since $\mathcal{D}\Omega_{\leq}$ is closed under β-expansion.

(ii) $M \twoheadrightarrow_\eta N$ implies that $\forall \sigma \in T\ \forall \Gamma \in \mathbf{B}\ (M \in \mathcal{V}_{\Gamma,\sigma} \Rightarrow N \in \mathcal{V}_{\Gamma,\sigma})$ since $\mathcal{D}\Omega_{\leq}$ is closed under η-reduction, but it does not imply the other implication because it is not closed under η-expansion.

A counter-example is $1 \equiv \lambda xy.xy \rightarrow_\eta \lambda x.x \equiv \mathsf{I}$, but there are types of I, which are not types of 1 such as $\alpha \rightarrow \alpha$, where α is a type variable, since

$$\vdash \lambda x.x : \alpha \rightarrow \alpha, \text{ but } \not\vdash \lambda xy.xy : \alpha \rightarrow \alpha.$$

Hence, $\mathsf{I} \in \mathcal{V}_{\alpha\rightarrow\alpha}$, but $1 \notin \mathcal{V}_{\alpha\rightarrow\alpha}$.

(iii) $(\Rightarrow)$ By (i).
$(\Leftarrow)$ Let $\Gamma \vdash N : \pi$ be a principal typing of N. Hence, $\Gamma \vdash M : \pi$. Therefore by Proposition 1(ii) $M \twoheadrightarrow_\beta N$.

β-equal terms cannot be separated. Lambda terms that are β-equal belong to the same open sets therefore we shall identify them. By "up to β-equality" we mean that if a normal form N is in a set, then all the terms that are β-equal to N are in the same set.

Proposition 3. *(i) Unsolvable terms are compactification points (bottoms).*

(ii) If N *is a normal form, then there is a type* $\rho \in T$ *and a context* Γ *such that*
$\mathcal{V}_{\Gamma,\rho} = \{P | N \twoheadrightarrow_\eta P\}$(*up to* $=_\beta$).

(iii) $\beta\eta$*-normal forms are isolated points (up to* $=_\beta$*).*

Proof.

(i) If M is an unsolvable term, then it is typable only by a type, which is equivalent to ω, therefore if $M \in \mathcal{V}_{\Gamma,\sigma}$, then $\sigma \sim \omega$. However, $\mathcal{V}_{\Gamma,\sigma} = \mathcal{V}_{\Gamma,\omega} = \Lambda$ by Lemma 1(iii). Therefore Λ is the only open set containing unsolvable terms, and hence they are compactification points.

(ii) If N is a normal form, then there is a principal typing $\Gamma \vdash N : \pi$. For every term M for which $\Gamma \vdash M : \pi$, i.e., $M \in \mathcal{V}_{\Gamma,\pi}$ by Proposition 1(i) there is an η-reduct P of N such that $M \twoheadrightarrow_\beta P$. By Lemma 2(i) $P \in \mathcal{V}_{\Gamma,\pi}$, also. P is a normal form as well. Hence $\mathcal{V}_{\Gamma,\pi} = \{P | N \twoheadrightarrow_\eta P\}$ up to β-equality, since $M =_\beta P$.

(iii) Obvious, since if N is a $\beta\eta$-normal form, then by (ii) there is a $\mathcal{V}_{\Gamma,\pi}$ which is a singleton, i.e., $\mathcal{V}_{\Gamma,\pi} = \{N\}$ up to β-equality.

This topology is not yet T_0, since two different unsolvable terms cannot be separated.

The Genericity Lemma is a consequence of the Continuity Theorem, as proved in [1].

Proposition 4. (Genericity Lemma)
Let M and N be lambda terms such that M is unsolvable and N has a normal form. Then, for all lambda terms F,

$$FM = N \Rightarrow \forall L \in \Lambda\,(FL =_{\beta\eta} N).$$

Proof. Let $N_{\beta\eta}$ be the $\beta\eta$-normal form of N, i.e., $N \twoheadrightarrow_{\beta\eta} N_{\beta\eta}$. If $\Gamma \vdash N_{\beta\eta} : \pi$ is a principal typing of $N_{\beta\eta}$, then as shown in Proposition 3(iii), there is $\mathcal{V}_{\Gamma,\pi} = \{N_{\beta\eta}\}$ which is a singleton up to $=_\beta$. Again, by Proposition 3(i) $\mathcal{V}_{\Gamma,\omega} = \Lambda$ is the only open set containing the unsolvable term M. By the Continuity Theorem 1

$$(\forall \mathcal{V}_{\Gamma,\varepsilon} \ni N)(\exists \mathcal{V}_{\Gamma,\delta} \ni M)(L \in \mathcal{V}_{\Gamma,\delta} \Rightarrow FL \in \mathcal{V}_{\Gamma,\varepsilon}).$$

Let us choose $\mathcal{V}_{\Gamma,\varepsilon} = \mathcal{V}_{\Gamma,\pi} = \{N_{\beta\eta}\}$. Namely, $\mathcal{V}_{\Gamma,\delta} = \mathcal{V}_{\Gamma,\omega} = \Lambda$, and so if $L \in \Lambda$, then $FL = N_{\beta\eta}$. By Proposition 1(ii) this means that $FL \twoheadrightarrow_\beta N_{\beta\eta}$, and so $FL =_{\beta\eta} N$.

The *filter topology* on lambda terms is introduced using the intersection type system, as well.

Let $\mathcal{F} \subseteq \mathcal{P}(T)$ be a *filter model*. The *valuation of lambda terms* $\|\ \| : \Lambda \to \mathcal{F}$ is given by the following mapping

$$\|M\| = \{\sigma | \Gamma \vdash M : \sigma\} \in \mathcal{F}.$$

The Scott topology is defined on $\mathcal{F}$ in the following way:

Definition 4. *A set $\mathcal{O} \subseteq \mathcal{F}$ is open if:*

(i) $d \in \mathcal{O}$ and $d \subseteq e$, then $e \in \mathcal{O}$;

(ii) $\cup d_i \in \mathcal{O}$, then there is an i_0 such that $d_{i_0} \in \mathcal{O}$.

This topology induces the so-called *filter topology* on Λ in the sense that open sets in Λ are $\|\mathcal{O}\|^{-1} \subseteq \Lambda$, where $\mathcal{O} \subseteq \mathcal{F}$ is open w.r.t. the Scott topology on filters.

The type topology and the filter topology are equivalent on the set of closed lambda terms, in the sense that open sets w.r.t. the type topology are open w.r.t. the filter topology and vice versa, as shown in [4]. The advantage of the type topology is that it is defined on lambda terms directly, while the filter topology is actually induced by the topology defined on filters.

Each lambda term $M \in \Lambda$ is associated with a certain tree $BT(M)$, the so-called Böhm tree. The *tree topolgy* is based on the notion of Böhm trees. We will recall here only some elementary notions, since the Böhm tree technique is rather involved, so for more details we would refer the reader to [1] (Chapter 10).

A Σ-*labelled tree* is a tree where an element of Σ is written at each node. Let

$$\Sigma = \{\bot\} \cup \{\lambda x_1 \ldots x_n.y | n \in \mathsf{N}, x_1, \ldots, x_n, y \ \ variables\}$$

Then $BT(M)$ is a Σ-labelled tree defined as follows.

$BT(M) = \bot$ if M is unsovable

$$BT(M) = \begin{array}{ccc} & \lambda x_1 \ldots x_n.y & \\ & \swarrow \quad \searrow & \\ BT(M_1) & & BT(M_m) \end{array} \quad \text{if } M \text{ has a hnf } \lambda x_1 \ldots x_n.y M_1 \ldots M_m$$

Let $\mathcal{B}$ denote the set of all Böhm trees. Then $\mathcal{B} = (\mathcal{B} \subseteq)$ is a complete partial ordering. Consider the cpo $\mathcal{B}$ with the Scott topology, as in Definition 4.

The *tree topolgy* on the set of all lambda terms Λ is the smallest one that makes the map

$$BT : \Lambda \to \mathcal{B}$$

continuous, meaning that open sets of Λ are of the form $BT^{-1}(\mathcal{O})$, with $\mathcal{O}$ Scott open in $\mathcal{B}$.

β-equal lambda terms cannot be separated neither in the tree topology, since they have equal head normal forms, and hence the same Böhm tree, nor in the type topology, since they are typable by the same types, as shown in Lemma 2(i). On the other hand η-equal terms can be separated in both of these topologies, since they do not necessarily have the same head normal forms and they are not typed by the same types. For example η-equal terms $1 \equiv \lambda xy.xy \to_\eta \lambda x.x \equiv \mathsf{I}$ are not typable by the same types as we saw in the proof of Lemma 2(ii). Both of them are head normal forms, even normal forms, but not equal. This means that their Böhm trees are different.

REFERENCES

[1] Barendregt, H.P., *The Lambda Calculus - Its Syntax and Semantics*, 2nd edition, North-Holland, Amsterdam, 1984.

[2] Barendregt, H.P., *Lambda calculi with types*, in: S. Abramsky, D. M. Gabbay, T. S. E. Maibaum (eds.), *Handbook of Logic in Computer Science*, Oxford University Press, Oxford, 1992, 117–309.

[3] Barendregt, H.P., Coppo M., and Dezani-Ciancaglini, M., *A filter model and the completeness of type assignment*, The Journal of Symbolic Logic 48 (1983), 931–940.

[4] Ghilezan, S., *Intersection types in lambda calculus and logic*, Ph. D. Thesis, University of Novi Sad, 1993.

[5] Krivine, J.L., *Lambda-calcul types et modèles*, Masson, Paris, 1990.

[6] S. Ronchi della Roca and B. Venneri, *Principal type schemes for an extended type theory*, Theoretical Computer Science 28 (1984), 151–171.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 75-81

GENERALIZING LOGIC PROGRAMMING TO ARBITRARY SETS OF CLAUSES

Slaviša B. PREŠIĆ
Faculty of Mathematics, University of Belgrade,
Studentski trg 16, 11000 Belgrade, Yugoslavia

Abstract. In this paper, which is a brief version of [3], we state how one can extend Logic Programming to any set of clauses.

Keywords: Logic Programming, deduction, completeness

The basic part of Logic Programming, particularly Prolog, in fact deals with the following two inference rules:

(1) $\mathcal{F}, p \vdash p$

(2) $\mathcal{F}, p \vee \neg q_1 \vee ... \vee \neg q_k \vdash p \longleftarrow \mathcal{F} \vdash q_1, ..., q_k$

(where $\mathcal{F}$ is a set of (positive) Horn formulas and p is any atom, i.e. a propositional letter)

Indeed, the informal meaning of rule (1) is:

An atom p is a consequence of a set of clauses if p is an element of that set.

Similarly for rule (2) we have this meaning:

An atom p is a consequence of a set $\mathcal{F}, p \vee \neg q_1 \vee ... \vee \neg q_k$ *(i.e of the set* $\mathcal{F}, q_1 \wedge ... \wedge q_k \Rightarrow p$*), if*$q_1, ..., q_k$ *are consequences of the set* $\mathcal{F}$*.*

In the sequel we use the following facts from mathematical logic (see [2]):

(3) *The notion of formal proof in the case of propositional logic (assuming we have chosen some tautologies as axioms, and that* **modus ponens** *is the only inference rule).*

(4) **The Deduction theorem**[1] : *$\mathcal{F}, A \vdash B \longleftrightarrow \mathcal{F} \vdash A \Rightarrow B$ where $\mathcal{F}$ is a set of propositional formulas and A, B are some such formulas.*

(5) **Completeness Theorem:** *Any propositional formula is a logical theorem if and only if it is a tautology.*

We also use the symbols $\perp, \top$ which can be introduced by the following definitions

$\perp$ stands for $a \wedge \neg a$; $\top$ stands for $a \vee \neg a$

where a is an atom (chosen arbitrarily). Further, let $\mathcal{F}$ be any set of propositional formulas and ψ a formula or one of the symbols $\perp, \top$. Then a **sequent** is any expession of the form $\mathcal{F} \vdash \psi$, with the meaning:

ψ is a logical consequence of $\mathcal{F}$

Lemma 1. *Let $\mathcal{F}$ be any set of propositional formulas not containing the atom p, and let $\phi_1(p), \phi_2(p), \ldots$ be propositional formulas containing p. Then we have the following equivalences*

$$(6)\ (i) \qquad \mathcal{F}, \phi_1(p), \phi_2(p), \ldots \vdash p \longleftrightarrow \mathcal{F}, \phi_1(\perp), \phi_2(\perp), \ldots \vdash \perp$$

$$(ii) \qquad \mathcal{F}, \phi_1(p), \phi_2(p), \ldots \vdash \neg p \longleftrightarrow \mathcal{F}, \phi_1(\top), \phi_2(\top), \ldots \vdash \perp$$

Proof. First we give proof of the $\longrightarrow$ part of (i). Then, we have the following 'implication-chain':

$\mathcal{F}, \phi_1(p), \phi_2(p), \ldots \vdash p$

$\longrightarrow$ For some formulas $f_1, \ldots, f_r$ of $\mathcal{F}$ and some formulas $\phi_{i1}(p), \ldots, \phi_{is}(p)$ we have: $f_1, \ldots, f_r, \phi_{i1}(p), \ldots, \phi_{is}(p), \ldots \vdash p$

(Finiteness of the propositional proof)

$\longrightarrow$ $\vdash f_1 \Rightarrow \ldots \Rightarrow f_r \Rightarrow \phi_{i1}(p) \Rightarrow \ldots \Rightarrow \phi_{is}(p) \Rightarrow p$

(By (4))

$\longrightarrow$ Formula

$$f_1 \Rightarrow \ldots \Rightarrow f_r \Rightarrow \phi_{i1}(p) \Rightarrow \ldots \Rightarrow \phi_{is}(p) \Rightarrow p$$

is a tautology

(By (5))

$\longrightarrow$ Formula $f_1 \Rightarrow \ldots \Rightarrow f_r \Rightarrow \phi_{i1}(\perp) \Rightarrow \ldots \Rightarrow \phi_{is}(\perp) \Rightarrow \perp$

is a tautology

[1] In fact, only the $\longrightarrow$-part is the deduction theorem. But, the $\longleftarrow$-part is almost trivial.

$\longrightarrow$ Formula

$$f_1 \Rightarrow ... \Rightarrow f_r \Rightarrow \phi_{i1}(\perp) \Rightarrow ... \Rightarrow \phi_{is}(\perp) \Rightarrow \perp$$

is a logical theorem

(By (5))

$\longrightarrow$ Formula

$$f_1, ..., f_r, \phi_{i1}(\perp), ..., \phi_{is}(\perp) \vdash \perp$$

holds.

(By (4))

$\longrightarrow$ $\mathcal{F}, \phi_1(\perp), \phi_2(\perp), ... \vdash \perp$

which completes the proof. Proof of the $\longleftarrow$ part of (i) reads:
$\mathcal{F}, \phi_1(\perp), \phi_2(\perp), ... \vdash \perp$

$\longrightarrow$ For some formulas $f_1, ..., f_r$ of $\mathcal{F}$ and some formulas $\phi_{i1}(\perp), ..., \phi_{is}(\perp)$ we have: $f_1, ..., f_r, \phi_{i1}(\perp), ..., \phi_{is}(\perp), ... \vdash \perp$

(Finiteness of every formal proof)

$\longrightarrow$ $\vdash f_1 \Rightarrow ... \Rightarrow f_r \Rightarrow \phi_{i1}(\perp) \Rightarrow ... \Rightarrow \phi_{is}(\perp) \Rightarrow \perp$

(By (4))

$\longrightarrow$ Formula

$$f_1 \Rightarrow ... \Rightarrow f_r \Rightarrow \phi_{i1}(\perp) \Rightarrow ... \Rightarrow \phi_{is}(\perp) \Rightarrow \perp$$

is a tautology

(By (5))

$\longrightarrow$ Formula

$$f_1 \Rightarrow ... \Rightarrow f_r \Rightarrow \phi_{i1}(p) \Rightarrow ... \Rightarrow \phi_{is}(p) \Rightarrow p$$

is a tautology

$\longrightarrow$ Formula

$$f_1 \Rightarrow ... \Rightarrow f_r \Rightarrow \phi_{i1}(p) \Rightarrow ... \Rightarrow \phi_{is}(p) \Rightarrow p$$

is a logical theorem

(By (5))

$\longrightarrow$ Formula

$$f_1, \dots, f_r, \phi_{i1}(p), \dots, \phi_{is}(p) \vdash p$$

holds.

(By (4))

$\longrightarrow \mathcal{F}, \phi_1(p), \phi_2(p), \dots \vdash p$

which completes the proof of (i).
We have omitted a proof of (ii) because (ii) can be proved in a similar way as (i).
Notice that Lemma 1 can be expressed by the following words:

> *A literal*[2] ψ *is a logical consequence of the given set if and only if the corresponding*[3] *set is inconsistent.*

Now we prove the following lemma.

Lemma 2. *The equivalence*

$$\mathcal{F}, p_1 \vee \dots \vee p_k \vdash \bot \longleftrightarrow \mathcal{F} \vdash \neg p_1, \dots, \mathcal{F} \vdash \neg p_k \tag{7}$$

(where p_i is any literal)

is true.

Proof. We have the following 'equivalence-chain':
$\mathcal{F}, p_1 \vee \dots \vee p_k \vdash \bot$

$\longleftrightarrow \mathcal{F} \vdash (p_1 \vee \dots \vee p_k \Longrightarrow \bot)$

(By (4))

$\longleftrightarrow \mathcal{F} \vdash (\neg p_1 \wedge \dots \wedge \neg p_k)$

(Using a well-known tautology)

$\longleftrightarrow \mathcal{F} \vdash \neg p_1, \dots, \mathcal{F} \vdash \neg p_k$

which completes the proof.

Besides (6) and (7) we emphasize the following obvious equivalences

$$\vdash \top \longleftrightarrow \mathcal{F}, \bot \vdash \bot \tag{8}$$

$$\mathcal{F}, \top \vdash A \longleftrightarrow \mathcal{F} \vdash A \tag{9}$$

(A is a literal or the symbol $\bot$)

[2] A literal is an atom or the negation of an atom

[3] i.e. one of the sets $\mathcal{F}, \phi_1(\bot), \phi_2(\bot), \dots$ or $\mathcal{F}, \phi_1(\top), \phi_2(\top), \dots$

Suppose now that $\mathcal{F}$ is a given set of clauses and ψ is a literal or $\perp$. Is it possible that using the equivalences (6), (7), (8), (9) one can establish whether or not ψ is a logical consequence of $\mathcal{F}$? In order to answer this we introduce the following inference rules[4]

(R1) $\mathcal{F}, \perp \vdash \perp \longleftarrow \vdash \top$

(R2) $\mathcal{F}, \phi_1(p), \phi_2(p), \ldots \vdash p \;\longleftarrow\; \mathcal{F}, \phi_1(\perp), \phi_2(\perp), \ldots \vdash \perp$

$\mathcal{F}, \phi_1(p), \phi_2(p), \ldots \vdash \neg p \;\longleftarrow\; \mathcal{F}, \phi_1(\top), \phi_2(\top), \ldots \vdash \perp$

($\phi_i(p)$ is any clause containing p)

(R3) $\mathcal{F}, p_1 \vee \ldots \vee p_k \vdash \perp \longleftarrow\; \mathcal{F} \vdash \neg p_1, \ldots, \mathcal{F} \vdash \neg p_k$

(where p_i is any literal)

(R4) $\mathcal{F}, \top \vdash A \longleftarrow \mathcal{F} \vdash A$

(A is a literal or the symbol $\perp$)

We emphasize that in the sequel for the set $\mathcal{F}$ we suppose that it does not contain a clause of the form $\ldots q \vee \neg q \ldots$,where q is any atom. Namely, such a formula is equivalent to $\top$, consequently it should be omitted[5]. Similarly, if it happens that by applying rule (R2) some clause becomes equivalent to $\top$ then we will also omit it.

Roughly speaking rules (R1),(R2),(R3),(R4) are used as follows:

> *We start with a question (a sequent) of the form $\mathcal{F} \vdash \psi$ and apply rules (R2),(R3),(R4) several times. If at some step we can apply rule (R1), the procedure stops with the conclusion that ψ is a logical consequence of $\mathcal{F}$. However, if at some step we obtain the sequent $\vdash \perp$ (then $\mathcal{F}$ is an empty set) the procedure stops with the conclusion that ψ is not a logical consequence of $\mathcal{F}$.*

Example 1. Answer the following questions:

1) $p \vdash p$? 2) $p, q \vdash p$? 3) $\vdash p$? 4) $q \vdash p$?

5) $\neg q \vee p, q \vee p \vdash p$? 6) $p, \neg p \vee q \vee \neg r, p \vee \neg q \vee s, p \vee s \vee \neg t \vdash \perp$?

where p, q, r, s, t are atoms.

Answer.

1) Applying (R2) we obtain the sequent $\perp \vdash \perp$ and by (R1) we get the sequent $\vdash \top$ so the answer is: *Yes.*

[4] We point out that the set $\mathcal{F}$ may be also an empty set.

[5] This is compatible with rule (R4)

2) Applying (R2) we obtain a new question, i.e the sequent $\perp, q \vdash \perp$, and now applying (R1) we obtain the sequent $\vdash \top$ so the answer is: *Yes.*
3) Applying (R2) we obtain the sequent $\vdash \perp$ so the asnwer is: *No.*
4) By (R2) we obtain the sequent $q \vdash \perp$ and after that by (R3) we obtain the sequent $\vdash \neg q$. Finally, by (R2) we obtain the sequent $\vdash \perp$ such that the answer is : *No.*
5) By (R2) we obtain the sequent $\neg q, q \vdash \perp$. Now by (R3) applied to the literal $\neg q$ we obtain the sequent $q \vdash q$, further by (R2) we obtain the sequent $\perp \vdash \perp$ and finally by (R1) we obtain the sequent $\vdash \top$ so the answer is : *Yes.*
6) Now by (R3) applied to clause p we obtain the sequent

$\neg p \vee q \vee \neg r, p \vee \neg q \vee s, p \vee s \vee \neg t \vdash \neg p$

By (R2) (and (R4) applied twice) we obtain the sequent

$q \vee \neg r \vdash \perp$

At this step applying (R3) we obtain two new sequents, i.e. questions
$\vdash \neg q$? and $\vdash r$?
The answer to the first question is *No*, so the final answer is also: *No.*

Concerning rules (R1)-(R4) we have this lemma.

Lemma 3. (Soundness of rules (R1)-(R4)). *Let $\mathcal{F}$ be any set of clauses. Suppose that we start with a sequent $\mathcal{F} \vdash \psi$, where ψ is a literal or the symbol $\perp$. If using rules* (R1)-(R4) *we obtain the sequent $\vdash \top$ or the sequent $\vdash \perp$, then ψ is / is not a logical consequence of set $\mathcal{F}$, respectively.*

Proof follows immediately from the fact that rules (R1)-(R4) are based on logical equivalences (6)-(9).

Let now $\mathcal{F} \vdash \psi$ be any sequent. By *Val($\mathcal{F} \vdash \psi$)* we denote its *truth value*, defined by:

If ψ is a logical consequence of set $\mathcal{F}$ then *Val($\mathcal{F} \vdash \psi$)* is *true*
otherwise *Val($\mathcal{F} \vdash \psi$)* is *false.*

According to this definition and to rules (R1)-(R4), i.e. to equivalences (6)-(9) we have the following equalities

(10) $Val(\vdash \top) = true$

$Val(\vdash \perp) = false$

$Val(\mathcal{F}, \perp \vdash \perp) = true$

$Val(\mathcal{F}, \top \vdash \psi) = Val(\mathcal{F} \vdash \psi)$

$Val(\mathcal{F}, \phi_1(p), \phi_2(p), ... \vdash p) = Val(\mathcal{F}, \phi_1(\perp), \phi_2(\perp), ... \vdash \perp)$

$$Val(\mathcal{F}, \phi_1(p), \phi_2(p), \ldots \vdash \neg p) = Val(\mathcal{F}, \phi_1(\top), \phi_2(\top), \ldots \vdash \bot)$$

($\phi_i(p)$ is any clause containing p)

$$Val(\mathcal{F}, p_1 \vee \ldots \vee p_k \vdash \bot)$$

$$= Val(\mathcal{F} \vdash \neg p_1) \ \textbf{and} \ \ldots \ \textbf{and} \ Val(\mathcal{F} \vdash \neg p_k)$$

(where p_i is any literal, i.e. an atom or the negation of an atom)

Suppose that $\mathcal{F}$ is a finite set. Then, in fact, these equalities define the function *Val* recursively on the number of all member of set $\mathcal{F}$. Consequently, these equalities suggest how to calculate $Val(\mathcal{F} \vdash \psi)$. In other words we have the following assertion:

(11) *If $\mathcal{F}$ is a finite set then one can effectively calculate $Val(\vdash \psi)$, i.e. establish whether or not ψ is a logical consequence of set $\mathcal{F}$.*

Next we will prove the following basic theorem.

Theorem 1. (Completeness) *Let $\mathcal{F}$ be a set of some clauses and ψ a literal or the symbol $\bot$. Then:*

> *ψ is a logical consequence of set $\mathcal{F}$ if and only if starting with $\mathcal{F} \vdash \psi$ and applying rules (R1)-(R4) a finite number of times one can obtain the sequent $\vdash \top$.*

Proof. The *if* - part follows immediately from Lemma 3. To prove the *only if* - part suppose now that ψ is a logical consequence of set $\mathcal{F}$. Then ψ is a logical consequence of some f i n i t e subset $\mathcal{A}$ of set $\mathcal{F}$ (for: every formal proof is finite). Next, by (11) we conclude that starting with the sequent $\mathcal{A} \vdash \psi$ and applying rules (R1)-(R4) a finite number of times one can obtain the sequent $\vdash \top$. Consequently, also starting with the sequent $\mathcal{F} \vdash \psi$ and applying rules (R1)-(R4) a finite number of times one can obtain the sequent $\vdash \top$. The proof is complete.

REFERENCES

[1] Lloyd J.W., *Foundations of Logic Programming*, Springer-Verlag, 1984

[2] Mendelson E., *Introduction to Mathematical Logic*, Van Nostrand, Princeton, 1979.

[3] Prešić S. B., *How to generalize Logic Programming to arbitrary set of clauses* (forthcoming).

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 83-90

SOME PROPOSITIONAL PROBABILISTIC LOGICS †

Miodrag RAŠKOVIĆ,
PMF, R.Domanovića 12, 34000 Kragujevac, Yugoslavia
Zoran OGNJANOVIĆ
Institute of Mathematics,
Kneza Mihaila 35, 11000 Belgrade, Yugoslavia

Abstract. We discuss a number of propositional probabilistic logics that allow us to express statements such as "If α holds with probability s, and β follows from α with probability t, then the probability of β is r.

Keywords: logic, probability, completeness, decidability.

1. INTRODUCTION

We will give a survey of a number of propositional probabilistic logics. They are conservative extensions of the classical propositional logic that allow statements such as "if α holds with probability s, and β follows from α with probability t, then the probability of β is r". This statement speaks about the probability of α, and β, but in our approach the formula itself is true or false, and it does not have any numerical value as in fuzzy logic. We will describe the syntax and the semantics of the logics and present results about their finiteness, completeness, compactness, and decidability. In this paper we will consider only propositional logics, since the crucial questions about probabilistic logics already arise at that level. We note that a similar approach can also be applied in the first order case.

†This work was supported by the Serbian Ministry of Science and Technology, through the Mathematical Institute, Belgrade.

2. SYNTAX

In order to express statements about the probability of some event, the classical language is augmented by a probabilistic operator $P_{\geq s}$ for every s from a presupposed set called the set of indices and denoted by I. The set I is a subset of $[0, 1]$.

Starting from the set of primitive propositions $\phi = \{p, q, r, \ldots\}$ and using classical operators and a probabilistic operator $P_{\geq s}$ for every $s \in I$, we make the set of all propositional probabilistic formulas. For example, $P_{\geq s}\alpha \wedge P_{\geq t}(\alpha \rightarrow \beta) \rightarrow P_{\geq r}\beta$ is a probabilistic formula.

The intuitive meaning of $P_{\geq s}\alpha$ is that the probability of α is greater or equal to s.

The other probabilistic operators $P_{<s}$, $P_{\leq s}$, $P_{>s}$, and $P_{=s}$ can be defined as $\neg P_{\geq s}$, $P_{\geq 1-s}\neg$, $\neg P_{\leq}$ and $P_{\geq s} \wedge P_{\leq s}$.

3. SEMANTICS

We use a possible-world approach to give semantics to probabilistic formulas. The probabilistic models are similar to the well-known Kripke models. The main difference is that our models use finite additive probabilistic measures to attach the truth-values to the formulas. This means that formula $P_{\geq s}\alpha$ is satisfied if the measure of the set of worlds that satisfy α is greater or equal to s. We assume the so-called measurable case, i.e. that to every formula in every model there corresponds a well-defined probability. The structures of the presented models depend on the corresponding logics, so we will give the precise definitions later on.

4. PROBABILISTIC LOGICS

Let us consider some combinations of the following parameters:

- the length of formulas,
- the set of indices,
- the iteration of probabilistic operators and
- the ranges of measures

that lead to a number of probabilistic logics.

In this paper we will use both finite formulas, and formulas with countable conjunction and disjunction. The set of indices $I = \{s : P_{\geq s}$ is an operator $\}$ can be: a finite subset of $[0, 1]$, the set of rational numbers from $[0, 1]$, or the whole set of real numbers from $[0, 1]$. Obviously, different choices of these parameters produce different levels of expressiveness of the resulting logics.

We do allow (or we do not allow) the iteration of probabilistic operators, so the formula $P_{\geq s}P_{\geq r}\alpha$ is (is not) a formula of the logical language. In the latter case we have a simpler decision procedure, while in the former we can speak about higher order probabilities.

When the models are considered, we distinguish two basic cases: in the first one we allow only probabilistic measures with finite ranges, while in the second one arbitrary probabilistic measures are permitted. The choice of this parameter implies an interesting result: compactness holds only for probabilistic logics whose models have a fixed finite range of measures. This means that in the general case the logics do not have any extended complete finite axiomatization. This is not surprising, because it is well known that compactness trivially follows from the extended completeness and finiteness of proofs. We can only obtain the finiteness property as a consequence of the Archimedean axiom for real numbers when we use a fixed finite range of measures. So, some of our logics are essentially infinite. We use in these logics ordinary finite formulas, while proofs can be infinite.

5. PROBABILISTIC LOGICS WITHOUT THE ITERATION OF PROBABILISTIC OPERATORS

In this section we present probabilistic logics whose languages do not allow the iteration of probabilistic operators, or the mixing of classical and probabilistic formulas. These logics possess similar semantics and axiomatizations.

5.1. The logic $LPP_{S,S}$

In the simplest logic $LPP_{S,S}$ (L means logic, the first P - propositional, and the second one - probabilistic, the first S corresponds to the set of indices, while the second one corresponds to ranges of measures) we use only finite formulas and consider a fixed, finite set $S = \{0 = s_1, s_2, \ldots, s_n = 1\}$, a subset of $[0,1]$, while the set I of indices and ranges of measure are equal to S. Also, we do not allow the iteration of probabilistic operators, or the mixing of pure classical propositional formulas and formulas that contain probabilistic operators. For example, if α is a classical propositional formula, then $\alpha \vee P_{\geq s}(\beta)$ is not a propositional formula. $LPP_{S,S}$ is fully described in [3] and [4].

Let (in this section) α, β and γ denote a classical propositional formula, while A and B denote probabilistic formulas.

The models for $LPP_{S,S}$ are tuples

$$M = (W, \pi, \{[\alpha]_W\}, \mu)$$

where W is a set of worlds, π is a propositional valuation, $[\alpha]_W$ is the set of worlds that classically satisfy the classical propositional formula α, $\{[\alpha]_W\}$ is an algebra of subsets of W, and μ is a finite additive probabilistic measure with finite range S defined over $\{[\alpha]_W\}$.

As can be seen, there is only one measure in each of the models. It follows that satisfiability is a property of the whole model, and not of worlds in the models.

Let M be a probabilistic model of the class described above. The following properties hold for the satisfaction relation $|\vdash$:

- $M |\vdash \alpha$ iff $(\forall w in W) w \models \alpha$,
- $M |\vdash P_{\geq s}\alpha$ iff $\mu([\alpha]_W) \geq s$,
- $M |\vdash \neg B$ iff it is not $M |\vdash B$ and
- $M |\vdash A \wedge B$ iff $M |\vdash A$ and $M |\vdash B$.

A complete axiom system for $LPP_{S,S}$ include the following eight axioms:

A1. $\alpha \to (\beta \to \alpha)$
A2. $(\alpha \to (\beta \to \gamma)) \to ((\alpha \to \beta) \to (\alpha \to \gamma))$
A3. $(\neg\beta \to \neg\alpha) \to (\alpha \to \beta)$
A4. $P_{\geq 0}\alpha$
A5. $P_{\geq s}\alpha \to P_{\geq r}\alpha$, $s > r$
A6. $(P_{\geq s}\alpha \wedge P_{\geq r}\beta \wedge P \geq 1(\neg\alpha \vee \neg\beta)) \to P_{\geq min(1,s+r)}(\alpha \vee \beta)$
A7. $(P_{\leq s}\alpha \wedge P_{<r}\beta) \to P_{<min(1,(s+r))}(\alpha \vee \beta)$
A8. $P_{>s_i}\alpha \to P_{\geq s_{i+1}}\alpha$

and two rules of inference ($\vdash$ denotes provability):

R1. From $\vdash A$ and $\vdash A \to B$ infer $\vdash B$.
R2. From $\vdash \alpha$ infer $\vdash P_1\alpha$.

The first three axioms with the modus ponens form an axiomatization of the classical propositional logic. The other axioms are about probabilistic reasoning: A4 is about the nonnegativity of probabilistic measure, A5 is about the monotonicity of probabilistic measure, A6 and A7 are about finite additivity and A8 concerns the finite range of measures. The R2 rule is so-called probabilistic generalization.

The extended completeness, compactness and decidability theorems hold for $LPP_{S,S}$. In the proof of the extended completeness theorem we construct a maximal extension F of a consistent set G of formulas in the following way: let $\overline{G}$ be the set of all propositional consequences of G, let $A_1, A_2, \ldots$ be an enumeration of all probabilistic formulas, and $F_0 = G \cup \overline{G} \cup \{P_{\geq 1}\alpha : \alpha \in \overline{G}\}$. Then $F_i = F_{i-1} \cup \{\alpha_i\}$, if $F_{i-1} \cup \{\alpha_i\}$, is consistent, otherwise $F_i = F_{i-1}$, and finally, $F = \cup_i F_i$. Now, we define worlds of the canonical model to be classical propositional interpretations that satisfy all classical formulas from F, and the measure μ so that $\mu([\alpha]) = \max\{s : P_{\geq s}\alpha \in F\}$. The axioms guarantee that we obtain a probabilistic model that satisfies G. Since we use a finite axiomatization, compactness is an easy consequence of the extended completeness. Decidability follows from the fact that $LPP_{S,S}$ possesses the final model property (every logic mentioned in this paper posseses this property) and that measures have a fixed finite range. So, we have to examine only a finite number of finite models to check whether a formula is satisfiable.

5.2. The logic $LPP_{[0,1],S}$

As a generalization of $LPP_{S,S}$ we can consider the logic denoted $LPP_{[0,1],S}$ in which the set I of indices is allowed to be the whole set of $[0,1]$. On the other hand, the corresponding models are as above, and the ranges of their measures remain equal to S. Now, there are uncountable many formulas. But, by adding the following axioms to the axiom system for $LPP_{S,S}$:

A9. $P_{<s}\alpha \rightarrow P_{\leq s}\alpha$

A10. $P_{\geq s}\alpha \rightarrow P_{>r}\alpha$, $s > r$

we can prove the same theorems as above. The main difference in the proofs is that in the extended completeness proof we first define a fragment as a suitable countable subset of the set of all formulas, and then use only formulas from the fragment in the construction of the maximal consistent set.

Note that we cannot perform the opposite. Namely, we cannot describe the models whose ranges of measures are $[0,1]$ using a set of formulas that contain probabilistic operators whose indices belong only to the finite subset of $[0,1]$.

5.3. The logic $LPP_{[0,1],[0.1]}$

If we allow that the set I of indices and ranges of measures are $[0,1]$, we have a quite different situation: the compactness does not hold. For example, although every finite subset of the following set of formulas

$$F = \{\neg P_{=0}\alpha\} \cup \{P_{<1/n}\alpha : n = 1, 2, \ldots\}$$

is satisfiable, the set F itself is not satisfiable.

To achieve an extended complete axiomatization of $LPP_{[0,1],[0,1]}$ we have to avoid the finiteness of the previous logics. We add the following rule:

R3. From $\vdash B \to P_{\geq s-1/n}\alpha$, for every positive integer, infer $\vdash B \to P_{\geq s}\alpha$ to the axiom system for $LPP_{[0,1],S}$. Also, we exclude axiom A8, because it is easy to see that it is not valid with respect to the considered class of models. In the proof of the extended compactness theorem we use a fragment as above, while in the construction of an maximal consistent extension of a consistent set of formulas there is also the following rule: if $A_i = B \to P_{\geq s}\alpha$, and if $F_{i-1} \cup \{A_i\}$ is not consistent, then $F_i = F_{i-1} \cup \{B \to \neg P_{\geq s-1/n}\alpha\}$, for some n. And this is enough to prove the theorem.

Although $LPP_{[0,1],[0,1]}$ possesses the finite model property, this does not mean that we have to consider only a finite number of models in the satisfiability checking procedure, as can be done for $LPP_{S,S}$ and $LPP_{[0,1],S}$. The cause is that ranges of measures are not fixed and finite. Nevertheless, $LPP_{[0,1],[0,1]}$ is decidable.

Let $p_1, \ldots, p_n$ be all the primitive propositions in a formula A and let an atom be $At = \pm p_1 \wedge \ldots \pm p_n$, where where $\pm p$ denotes either p, or $\neg p$. By the propositional reasoning, we transform a formula A into its disjunctive normal form $DNF(A)$. A is satisfiable if and only if at least one of the members of $DNF(A)$ is satisfiable. These members look like $D_i = \pm P_{\geq s_1}\alpha_1 \wedge \ldots \pm P_{ges_m}\alpha_n$. We now translate every member of $DNF(A)$ into a system of linear equalities and inequalities. Every system contains the following equalities :

$$\sum_{At} \mu(At) = 1$$
$$\mu(At) \geq 0$$

as well as a linear inequality for every member $\pm P_{\geq s_j}\alpha_j$ of D_i:

$$\sum_{At \in DNF(\alpha_j)} \mu(At)\ \rho\ s$$

where ρ denotes $<$ if $\pm$ means $\neg$, and $\geq$ otherwise. So, we reduce the satisfiability checking problem to a decidable linear system solving problem.

6. THE LOGIC $LPP_{[0,1],c}$

Let us consider a class of probabilistic models that satisfy the constraint: there is a real number c from $(0,1)$ such that for every model M, its measure μ, and for every formula α if $\mu([\alpha]_W) > 0$, then $\mu([\alpha]_W) > c$. Using a lemma from the measure theory, we can show that every model in the class has a measure with a finite range. Again, compactness does not hold: an example is the set $F = \{P_{>c}(\alpha) \wedge P_{<c+1/n}(\alpha)\}$. We obtain a complete axiomatization by adding

A11. $P_{>0}(\alpha) \to P_{>c}(\alpha)$

to the axiom system of $LPP_{[0,1],[0,1]}$. The decidability of the logic can be proved using the same arguments as above.

6.1. The logic $LPP_{\mathcal{A},\omega_1,Fin}$

The last logic in this section concerns the class of all probabilistic models whose measures have finite ranges. Let $\mathcal{A}$ be a countable admissible set that contains ω and the propositional probabilistic language. In $LPP_{\mathcal{A},\omega_1,Fin}$ the conjunction symbol and the disjunction symbol may be applied to finite or countable sets of formulas. We extend the axiom system for $LPP_{[0,1],[0,1]}$ with the axioms of the propositional part of the $L_{\mathcal{A},P_1,P_2}$ logic [2], as well as with a generalization of the axiom A11:

A12. $\bigvee_{c>0} \bigwedge_{\alpha\in F} (P_{>0}\alpha \rightarrow P_{>c}\alpha)$

where $F \in \mathcal{A}$. Hence, in this logic we have both infinite formulas and proofs. The completeness now follows from the above explanation and the middle model construction procedure [2].

7. PROBABILISTIC LOGICS WITH THE ITERATION OF PROBABILISTIC OPERATORS

Let us shortly consider the class of models with the following structure:

$$M = (W, \pi, Prob)$$

where W is a set of worlds, and π is a classical valuation as above. $Prob$ is a function which assigns to every world $u \in W$ a probability space $(W(u), V(u), \mu(u))$, where $W(u) \subset W$, $V(u)$ is an algebra of subsets of $W(u)$ and $\mu(u)$ is a finite additive measure over $V(u)$. To describe such models we allow formulas with the iteration of probabilistic operators. So, $\beta \rightarrow P_{\geq s}P_{\geq r}\alpha$ is a formula of our language. The main semantical consequence of the above choice is that the satisfaction relation $\Vdash$ is a relation between worlds and formulas:

$w \Vdash P_{\geq s}\alpha$ iff $\mu(w)([\alpha]_{W(u)}) \geq s$

Now, we can combine the other parameters from Section 4, and produce the corresponding logics. The new logics, and the logics from Section 5, possess similar properties. For more details see [1] where a logic that is an extension of $LPP_{S,S}$ is described.

8. SOME APPLICATION OF PROBABILISTIC LOGICS

Reasoning about uncertainty is required by many artificial intelligence applications (expert systems, decision making systems, fault tree analysis, ...). Very often data and rules used in deduction are not known with certainty. Since classical logic is useful when the knowledge is crisp, it is desirable to have a logic that enables direct reasoning about uncertainty.

Using the described probabilistic logics we can deduce whether some conclusions follow from premises. Similarly, we can examine the probability that conclusions follow from suppositions, or we can look for the most probable event in a set. Probabilistic logics are also applicable in situations where it is difficult to attach precise probabilities to events, but it is possible to give upper and lower bounds for probabilities, or, at least, it is possible to compare (unknown) probabilities of events.

REFERENCES

[1] Ognjanović Z., Rašković M., *A logic with higher order probabilities*, to appear in Publications de L'Institut Mathematique.

[2] Rašković M., *Completeness theorem for biprobability models*, The Journal of symbolic logic 51 (1986), 586 – 590.

[3] Rašković M., *Classical logic with some probability operators*, Publications de L'Institut Mathematique 53(67) (1993), 1 – 3.

[4] Rašković M., Ognjanović Z., *A logic with a probabilistic operator*, International Congres of Industrial and Applied Mathematics, Hamburg (1995).

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 91-96

NOTE ON TWO PROBLEMS IN CONSTRUCTIVE COMMUTATIVE COIDEAL THEORY

Daniel A. ROMANO
University of Banja Luka, Faculty of Phylosophy,
Department of Mathematics and Informatics,
78000 Banja Luka, Bana Lazarevića 1, Republic of Srpska

Abstract. Among the most important concept's of Bishop's constructive mathematics are coideals in the constructive commutative ring theory. In this short note we shall give two results about cocongruence on a commutative ring with apartness which are preceded by some of the author's results on the existence of coideals. After that, we shall define two new questions about relations between the ideals and coideals of a commutative ring with apartness which come from the above results.

Keywords: constructive mathematics, coequality relation, cocongruence, coideal.

AMS Subject Classification (1991): 03F55, 13A99

Let $(R, =, \neq, +, \cdot)$ be a commutative ring with an identitiy where the diversity relation $\neq$ is an apartness in the sense of books [1], [3], [8], [9] and papers [2], [4], [5]. A subset S of R is **a coideal** ([4], [5], [8], [9]) of R if and only if

$$o \# S,$$
$$-x \in S \Rightarrow x \in S,$$
$$x + y \in S \Rightarrow x \in S \vee y \in S,$$
$$xy \in S \Rightarrow x \in S \wedge y \in S.$$

The coideals of a commutative ring with apartness were first defined and studied by W. Ruitenburg 1982 ([8]). After that, coideals (anti-ideals) were studied by A. S. Troelstra and D. van Dalen in their monograph [9]. The author proved in his paper [4] that if S is a coideal of ring R, then the relation q on R, defined by $(x, y) \in q$ if $x - y \in S$, satisfies the following properties:

(1) $(\forall x \in R)((x, x)\# q)$ consistent ([3])

(2) $(\forall x, y \in R)((x, y) \in q \Rightarrow (y, x) \in q)$ symmetric

(3) $(\forall x, z \in R)((x, z) \in q \Rightarrow (\forall y \in R)((x, y) \in q \vee (y, z) \in q)$ cotransitive ([3])

(4) $(\forall x, y, u, v \in R)((x + u, y + v) \in q \Rightarrow (x, y) \in q \vee (u, v) \in q)$

(5) $(\forall x, y, u, v \in S)((xu, yv) \in q \Rightarrow (x, y) \in q \vee (u, v) \in q)$

A relation q on R, which satisfies properties (1)-(5), is called **cocongruence** on R ([4],[5]). A relation q on a set $(R, =, \neq)$, which satisfies properties (1)-(3), i.e. which is a consisitent, symmetric and cotransitive relation on R is called a **coequality relation** on R ([2]). If q is a cocongruence on ring R then the set $S = \{x \in R \mid (x, o) \in q\}$ is a coideal of R ([4]). Let J be an ideal and S be a coideal of ring R. Ruitenburg, in his dissertation ([8], page 33) first stated the requirement that $J \subseteq \neg S$. This condition is equivalent with the following condition:

$$(\forall x, y \in R)(x \in J \wedge y \in S \Rightarrow x + y \in S).$$

In this case we say that J and S are **compatible**. W. Ruitenburg [8] **first** raised the question about the existence of a coideal S compatible with a given ideal J and the question about the existence of an ideal J compatible with a given coideal S. If e is a congruence on R, which is determined by S, then J and S are compatible if and only if:

$$(6) \qquad (\forall x, y, z \in R)((x, y) \in e \wedge (y, z) \in q \Rightarrow (x, z) \in q).$$

In the general case, if e is an equivalence and if q is coequivalence on a set $(R, =, \neq)$ we say that they are **compatible** if and only if they satisfy the condition (6). So, we have the following questions:

(i) If S is a coideal of R, is there an ideal J of R compatible with S?

(ii) If J is an ideal of R is there a coideal S of R compatible with J?

In this paper we shall answer these questions by citing some previous results. Most of these results are obtained by author. For other considerations on the notions and notations of constructive mathematics we refer to books [1], [3], [8], [9] and on the constructive commutative ring theory we refer to books [3], [8], [9] and papers [2], [4], [5], [6], [7].

In the further course of this work we need the following result which answers the first question. Let Y be a subset of R. By $x\#Y$ we denote $(\forall y \in Y)(x \neq y)$ and by $\overline{Y}$ we denote the set $\{x \in R \mid x\#Y\}$. The following result answers the question about the existence of an ideal of R compatible with the given coideal S of R.

Theorem 1. ([5], Proposition 3.5) *Let S be a coideal of ring R. Then, $\overline{S}$ is an ideal of R compatible with S.*

To answer the main question about the existence of a coideal S compatible with given ideal J we need the following notion. Let a and b be relations on a set $(R, =, \neq)$. By $b * a$ we denote the **filled** product of relation a and relation b defined by

$$b * a = \{(x, z) \in R \times R \mid (\forall y \in R)((x, y) \in a \vee (y, z) \in b)\}.$$

The filled product is associative and $(b * a)^{-1} = a^{-1} * b^{-1}$. For $n \geq 2$, ${}^n a$ means $a * a * \cdots a$ (n factors). Put ${}^1 a = a$. The next theorem gives a very important construction.

Theorem 2. *Let a be a relation on a set $(R, =, \neq)$. Then the relation $c(a) = \bigcap_{n \in N} {}^n a$ is a cotransitive relation on R.*

Proof. For cotransitivness we need to prove that

$$(x, z) \in c(a) \Rightarrow (\forall y \in R)((x, y) \in c(a) \vee (y, z) \in c(a)),$$

i.e. we need to prove that

$$(x, z) \in c(a) \Rightarrow (\forall y \in R)((\forall i \in N)(x, y) \in {}^i a) \vee ((\forall j \in N)(y, z) \in {}^j a).$$

First, we have

$$\begin{aligned}(x, z) \in c(a) &\Rightarrow (x, z) \in {}^2 a = a * a \\ &\Rightarrow (\forall y \in R)((x, y) \in a \vee (y, z) \in a).\end{aligned}$$

Second, for $n \geq 2$, suppose that

$$\begin{aligned}(x, z) \in c(a) &\Rightarrow (x, z) \in {}^{2n} a = {}^n a * {}^n a \\ &\Leftrightarrow (\forall y \in R)((x, y) \in {}^n a \vee (y, z) \in {}^n a).\end{aligned}$$

Thus, we have

$$\begin{aligned}(x, z) \in c(a) &\Rightarrow (x, z) \in {}^{2(n+1)} a = {}^{n+1} a * {}^{n+1} a \\ &\Leftrightarrow (\forall y \in R)((x, y) \in {}^{n+1} a \vee (y, z) \in {}^{n+1} a).\end{aligned}$$

Therefore, for each natural number n, we have

$$(x,z) \in c(a) \Rightarrow (\forall y \in R)((\forall i \leq n)((x,y) \in {}^{i}a) \vee (\forall j \leq n)((y,z) \in {}^{j}a).$$

Finally it means

$$(x,z) \in c(a) \Rightarrow (\forall y \in R)((x,y) \in c(a) \vee (y,z) \in c(a)).$$

If a is relation on R, then the relation $c(a)$ is called **cotransitive closure** of a. For this notions we have the following result:

Corollary 2.1. ([6]) *Let e be an equivalence on a set $(R, =, \neq)$. Then the cotransitive closure $c(\overline{e})$ is a coequality relation on R compatible with e.*

The following theorem gives a construction of cocongruence on the ring $(R, =, \neq, +, \cdot)$ on the basis of the given coequality relation on R.

Theorem 3. *Let q be the coequality relation on ring R. Then the relation*

$$q^* = \{(x,y) \in R^2 \mid (\exists s,t \in R)(xt+s, yt+s) \in q)\}$$

is a cocongruence on R.

Proof.

$$\begin{aligned}
(i)\ (u,v) \in q^* &\Leftrightarrow (\exists s,t \in R)((ut+s, vt+s) \in q) \\
&\Rightarrow (\exists s,t \in R)(\forall x \in R)(\exists s,t \in R)((ut+s, vt+s) \\
&\quad \neq (xt+s, xy+s)) \\
&\Leftrightarrow (\exists s,t \in R)(\forall x \in R)(ut+s \neq xt+s \\
&\quad \vee vt+s \neq xt+s) \\
&\Rightarrow (\forall x \in R)(u \neq x \vee v \neq x) \\
&\Leftrightarrow (\forall x \in R)((x,x) \# q); \\
(ii)\ (u,v) \in q^* &\Leftrightarrow (\exists s,t \in R)((ut+s, vt+s) \in q) \\
&\Leftrightarrow (\exists s,t \in R)((vt+s, vt+s) \in q) \\
&\Leftrightarrow (v,u) \in q; \\
(iii)\ (u,w) \in q^* &\Leftrightarrow (\exists s,t \in R)((ut+s, wt+s) \in q) \\
&\Rightarrow (\exists s,t \in R)(\forall v \in R)((ut+s, vt+s) \in q \\
&\quad \vee (vt+s, wt+s) \in q) \\
&\Rightarrow (\forall v \in R)((u,v) \in q^* \vee (v,w) \in q^*); \\
(iv)\ (u+x, v+y) \in q^* &\Leftrightarrow (\exists s,t \in R)(((u+x)t+s, (v+y)t+s) \in q) \\
&\Rightarrow (\exists xt+s, t \in R)((ut+xt+s, vt+xt+s) \in q
\end{aligned}$$

$$
\begin{aligned}
&\vee(\exists s+vt, t \in R)((vt+xt+s, vt+yt+s) \in q) \\
&\Rightarrow (u,v) \in q^* \vee (x,y) \in q^*;
\end{aligned}
$$

$$
\begin{aligned}
(v)\ (ux, vy) \in q^* &\Leftrightarrow (\exists s, t \in R)((uxt+s, vyt+s) \in q) \\
&\Rightarrow (\exists s, t \in R)((uxt+s, uyt+s) \in q \\
&\quad \vee (uyt+s, vyt+s) \in q) \\
&\Rightarrow (\exists s, ut \in R)((x(ut)+s, y(ut)+s) \in q \vee \\
&\quad (\exists s, yt \in R)(u(yt)+s, v(yt)+s) \in q); \\
&\Rightarrow (x,y) \in q^* \vee (u,v) \in q^*.
\end{aligned}
$$

Corollary 3.1. ([7]) *Let J be an ideal of ring R. Then there is coideal $S(J)$ compatible with J.*

Proof. Let J be an ideal of ring R. Then $e = \{(x,y) \in R^2 \mid x - y \in J\}$ is a congruence on R. Thus, by Corollary 2.1, the relation $q = c(\overline{e})$ is a coequality relation on ring R, and, by Theorem 3, the relation q^* is a cocongruence on R. By one of the theorems of paper [4], the set $S(J) = \{x \in R \mid (x, o) \in q^*\}$ is a coideal of R. Further, we have

$$
\begin{aligned}
&x \in J \wedge y \in S \Rightarrow x \in J \wedge y + x - x \in S \\
&\Leftrightarrow x \in J \wedge (y + x \in S \vee -x \in S) \\
&\Rightarrow (-x \in J \wedge -x \in S) \vee y + x \in S \\
&\Leftrightarrow ((-x, o) \in e \wedge (-x, o) \in q^*) \vee y + x \in S \\
&\Rightarrow (\exists t, s \in R)((-xt+s, s) \in e \wedge (-xt+s, s) \in q) \vee y + x \in S \\
&\Rightarrow (\exists t, s \in R)((-xt+s, s) \in e \wedge (-xt+s, s) \in \overline{e}) \vee y + x \in S \\
&\Rightarrow (\exists t, s \in R)((-xt+s, s) \in e \cap \overline{e} = \emptyset) \vee y + x \in S \\
&\Leftrightarrow y + x \in S.
\end{aligned}
$$

We can now formulate two questions about the connections between ideals and coideals.

Question 1. *Let Q be a coideal of a commutative ring R. We can construct, by Theorem 1, the ideal $\overline{Q}$ compatible with J, and we can construct, by Corollary 3.1, the coideal $S(\overline{Q})$ compatible with $\overline{Q}$. What kind of connection exists between coideal Q and $S(\overline{Q})$?*

Question 2. *Let J be an ideal of a commutative ring R. We can construct, by Corollary 3.1, the coideal $S(J)$ of R compatible with J, and we can construct, by Theorem 1, the ideal $\overline{S(J)}$ of R compatible with $S(\overline{J})$. What kind of connection exists between ideal J and $\overline{S(J)}$?*

REFERENCES

[1] Bishop E., *Foundations of Constructive Analysis*, McGraw-Hill, New York, 1967.

[2] Božić M., Romano D.A., *Relations, function's relations and operations in constructive mathematics*, Publ. VTS, Ser. Math. 2(1985), 25–39.

[3] Mines R., Richman F., and Ruitenburg W., *A Course of Constructive Algebra*, Springer, New York, 1988.

[4] Romano D.A., *Rings and fields, a constructive view*, Z. Math. Logik Grundl. Math. 34(1) (1988), 25–40.

[5] Romano D.A., *Rings and modules, a constructive view*, in: S. Crvenković (ed.), *Proceedings of the 6th Conference "Algebra and Logic", Sarajevo 1987*, University of Novi Sad, Institute of Mathematics, Novi Sad 1989, 113-140.

[6] Romano D.A., *A construction of the coequality relation*, Publ. Inst. Math. (Belgrade) (to appear).

[7] Romano D.A., *A construction of coideal compatible with given ideal of commutative ring with apartness*, Filomat (Niš), (to appear).

[8] Ruitenburg W., *Intuitionistic algebra*, Ph. D. Thesis, University of Utrecht, Utrecht, 1982.

[9] Troelstra A. S., van Dalen D., *Constructivism in Mathematics, an Introduction*, North-Holland, Amsterdam 1988.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 97-104

ON p-ADIC SERIES WITH RATIONAL SUMS

Branko DRAGOVIĆ
Institute of Physics, P. O. Box 57,
11001 Belgrade, Yugoslavia

Abstract. For a class of power series with coefficients which contain factorials, the domain of p-adic convergence and the summation formula are derived. The problem of rational points is studied in detail. The so-called adelic summation of divergent counterparts in the real case is also considered.

Keywords: p-adic series, summation formula, adelic summation

1. INTRODUCTION

Prof. Djuro Kurepa was a mathematician with wide interests in mathematics, and not only in mathematics; among the many problems he was interested in p-adic spaces [12]. This paper on p-adic series is devoted with respect to his scientific personality and to him as a great Serbian mathematician.

P-adic numbers were introduced in mathematics about one hundred years ago by K. Hensel. In modern mathematics, p-adic analysis (for an excellent exposition see [13]) is a rapidly developing subject. Different aspects of p-adic numbers and analysis have been employed in many fields of mathematics, e. g. in number theory and the arithmetic theory of algebraic groups. Since 1987, there has been significant activity in the application of p-adic mathematics in various parts of theoretical and mathematical physics (for a review, see e.g.[2], [11], [14], and for new achievements see [10]).

According to the Ostrowski theorem any non-trivial norm on the field of rational numbers Q is equivalent to the usual absolute value or to some p-adic norm. By definition, the p-adic norm of a rational number $0 \neq x = p^\nu r/s$ (where integers r and s are not divisible by given prime number p) is

$|x|_p = p^{-\nu}$, and $|0|_p = 0$. This norm is a mapping from Q into non-negative real numbers and has the following properties:

(i) $|x|_p \geq 0$, $|x|_p = 0$ if and only if $x = 0$,

(ii) $|xy|_p = |x|_p |y|_p$,

(iii) $|x+y|_p \leq \max(|x|_p, |y|_p) \leq |x|_p + |y|_p$, for all $x, y \in Q$.

Because of the strong triangle inequality $|x+y|_p \leq \max(|x|_p, |y|_p)$ it is called non-archimedean (or ultrametric) norm (valuation).

It is well known that the field of real numbers R can be obtained by completion of R with respect to the absolute value. The field of p-adic numbers Q_p may be regarded as a completion of Q with respect to the p-adic norm. Note that for each prime number p there is one Q_p. Q_p is neither isomorphic to R nor Q_q, where q is a prime distinct of p. Any $x \in Q_p$ has the expansion

$$(1) \qquad x = \sum_{n=k}^{+\infty} a_n p^n , \qquad a_n \in \{0, 1, \cdots, p-1\} , \quad k \in \mathbf{Z}$$

which is convergent with respect to the p-adic norm, and it is in an opposite direction to that one of the real case.

Many concepts of p-adic analysis are introduced in analogy to classical (real) analysis, but convergence is treated with respect to the p-adic norm. One of the basic objects which we often encounter in pure p-adic analysis as well as in its applications to physics, is the infinite power series.

My interest in p-adic series was initiated in 1987 [1] by the observation that divergent perturbative series which we usually deal with in theoretical physics, are p-adic convergent. In fact, such power series have the form

$$(2) \qquad \sum_{n=0}^{\infty} A_n x^n , \qquad A_n \in Q ,$$

and may be treated in R ($x \in R$) as well as in any Q_p ($x \in Q_p$). Loosely speaking, the less convergence there is in the real case, the more convergence there is in the p-adic one, and vice versa.

This series property (2) that it may be divergent in R and convergent in all Q_p (for all but a finite number of p), leads to a question regarding to possible connection between convergence in some p-adic number fields and the summation of the series divergent counterpart in R. An answer can be found within Q, because Q is a subfield of R and Q_p for all p. Namely, the sum of the divergent series depends on the way of summation. When a series is p-adic convergent and has a definite rational sum for all or almost all Q_p, then it seems natural to attach this rational sum to the divergent counterpart in R. This method of summation of divergent series is called adelic summation [3].

The power series which we usually encounter have only a trivial rational sum *i.e.* there exists a rational sum only for the argument $x = 0$. However, one can find power series (2) which are p-adic convergent and have (usually one) non-trivial rational point. Such series contain factorials in the coefficients and they are similar to divergent perturbative expansions in quantum field theory and string theory. Since in such cases x plays the role of a coupling constant, the rational value (which gives a rational sum) could be taken as its physical value.

In this article we mainly consider summation in the rational points of the series

$$\sum_{n=0}^{\infty} \varepsilon^n (\alpha n + \beta)! P_k(n) x^{\mu n + \nu} \tag{3}$$

from the p-adic point of view.

Some other aspects of p-adic series and their possible connection with the real case can be found in the author's papers [3]-[9].

2. CONVERGENCE

Power series (2) is p-adic convergent for some $x \in Q_p$ if and only if [13]

$$| A_n x^n |_p \to 0 \text{ as } n \to \infty \ . \tag{4}$$

Unlike the real case, (4) is not only a necessary but also a sufficient condition, and this is a direct consequence of the strong triangle inequality of the p-adic norm.

To find the domain of convergence of (3) we shall need an equation of the form

$$| m! |_p = p^{-\frac{m - S_m}{p-1}} \ , \tag{5}$$

where S_m is the sum of the digits in the canonical expansion of positive integer m over p. For example, if $m = m_0 + m_1 p + \cdots + m_r p^r$ then $S_m = \sum_{i=0}^{r} m_i$. In (5) the exponent $(m - S_m)/(p-1)$ is the number of factors p in $m!$. For a derivation see [13].

Let

$$P_k(n) = C_k n^k + \cdots + C_1 n + C_0 \tag{6}$$

be a polynomial of degree k over n with the coefficients $C_k, \cdots, C_0 \in Q$.

Theorem 1. *The power series*

$$\sum_{n=0}^{\infty} \varepsilon^n (\alpha n+\beta)! P_k(n) x^{\mu n+\nu} \ , \ \varepsilon = \pm 1 \ , \tag{7}$$

where $\alpha, \mu \in N$, $\beta, \nu \in N_0 = N \cup \{0\}$, *and* $P_k(n)$ *is a polynomial (6), has the domain of p-adic convergence given by*

$$| x |_p < p^{\frac{\alpha}{(p-1)\mu}} \ . \tag{8}$$

Proof. The p-adic norm of the general term of (7) is

$$| (\alpha n+\beta)! |_p | P_k(n) |_p | x |_p^{\mu n+\nu} \ , \tag{9}$$

where $| P_k(n) |_p \leq \max_{0\leq i\leq k} | C_i |_p$. According to (5) one has

$$p^{-\frac{\alpha n+\beta - S_{\alpha n+\beta}}{p-1}} | P_k(n) |_p | x |_p^{\mu n+\nu}$$

which for a large enough n behaves like

$$\left(p^{-\frac{\alpha}{(p-1)\mu}} | x |_p \right)^{\mu n} . \tag{10}$$

Expression (10) tends to zero as $n \to \infty$ if and only if $| x |_p$ satisfies (8).

From (8) it follows that the region of convergence of (7) depends on parameters α and μ, but it cannot be smaller than

$$| x |_p \leq 1 \ . \tag{11}$$

Recall that

$$| k |_p \leq 1 \ , \tag{12}$$

where k is an integer. Hence, the series (7) is p-adic convergent for any $x \in Z$ and it holds in Q_p for every p. There are also possibilities for (7) to be convergent for some $x \in Q \backslash Z$ if $\alpha \geq 2$.

3. SUMMATION

We want to give a formula which is suitable for summation of the series (7) for all cases which allow rational sums.

Let us denote $(m+1)_\alpha = (m+1)\cdots(m+\alpha)$ and let $A_\eta(n)$ be a polynomial of the form (6).

Theorem 2. *The summation formula*

$$\sum_{n=0}^{\infty} \varepsilon^n (\alpha n+\beta)!\{(\alpha n+\beta+1)_\alpha A_\eta(n+1)x^\mu - \varepsilon A_\eta(n)\}x^{\mu n+\nu} = \tag{13}$$
$$= -\varepsilon\beta! A_\eta(0)x^\nu$$

holds under those conditions which make the series (7) convergent.

Proof. The left-hand side of (13) may be rewritten in the form

$$\sum_{n=1}^{\infty} \varepsilon^{n+1}(\alpha n+\beta)! A_\eta(n)x^{\mu n+\nu} - \sum_{n=0}^{\infty} \varepsilon^{n+1}(\alpha n+\beta)! A_\eta(n)x^{\mu n+\nu}$$

which after mutual cancelation of terms for $n \geq 1$ gives

$$-\varepsilon\beta! A_\eta(0)x^\nu \ . \tag{14}$$

Although Eq. (13) is based on a rather trivial structure it gives highly non-trivial results.

Note that one can take any x which belongs to the domain of convergence (8) and put it into formula (13). In such case one gets a series of the form (7) with the corresponding sum. If $x \in Q$ and satisfies (8) then the sum of that series is the rational number (14). The summation formula (13) does not depend on the number field Q_p and for $x \in Q$ from the domain of convergence yields the same sum (14) valid in all Q_p.

Proposition 1. *The series (7) has a rational sum for some $x \in Q$ which satisfies (8), if there exists an auxiliary polynomial $A_\eta(n)$ such that*

$$P_k(n) = (\alpha n+\beta+1)_\alpha A_\eta(n+1)x^\mu - \varepsilon A_\eta(n) \ , \tag{15}$$

where $\eta = k - \alpha$.

Proof. Let there exists a polynomial $A_\eta(n) = a_\eta n^\eta + \cdots + a_1 n + a_0$, where $a_\eta, \cdots a_0 \in Q$, which satisfies (15). Then according to the summation formula (13) the rational sum of (7) does exist and the sum is given by (14).

Proposition 2. *For a given $x \in Q$ also belonging to the domain of convergence one can always find one polynomial $P_\alpha(n)$, and countably many of them if $k > \alpha$, such that the series (7) has a sum which is a rational number.*

Proof. According to (15) for $\eta = 0$ one obtains $P_\alpha(n) = (\alpha n+\beta+1)_\alpha x^\mu - \varepsilon$ as the polynomial of the lowest degree. If $\eta \geq 1$ then the polynomial $P_k(n) = P_{\alpha+\eta}(n)$ and it is determined by the coefficients of $A_\eta(n)$.

All possible polynomials $A_\eta(n)$ generate all polynomials $P_k(n)$, $k = \alpha+\eta \geq \alpha$, which yield the rational summation of (7). The problem of finding all possible $P_k(n)$, which are related to rational sums for some $x_0 \in Q$ may be reduced to

$$\sum_{n=0}^{\infty} \varepsilon^n (\alpha n + \beta)!(n^k + u_k) x_0^{\mu n+\nu} = v_k \ , \tag{16}$$

where

$$(\alpha n + \beta + 1)_\alpha A_\eta(n+1) x_0^\mu - \varepsilon A_\eta(n) = n^k + u_k \ , \tag{17}$$

$$-\varepsilon \beta! A_\eta(0) x_0^\nu = v_k \ . \tag{18}$$

Then the general form of desired $P_k(n)$ is

$$P_k(n) = C_k n^k + \cdots + C_\alpha n^\alpha + \sum_{i=\alpha}^{k} C_i u_i \ . \tag{19}$$

Now we can write the subclass of (7) which has a rational sum for some $x_0 \in Q$ from the common p-adic domain of convergence, *i.e.*

$$\sum_{n=0}^{\infty} \varepsilon^n (\alpha n + \beta)! [\sum_{i=\alpha}^{k} C_i n^i + \sum_{i=\alpha}^{k} C_i u_i] x^{\mu n+\nu} \ , \tag{20}$$

where u_i are defined by (17) and $C_\alpha, \cdots C_k \in Q$. When we put $x = x_0$ then (20) has the rational sum

$$\sum_{i=\alpha}^{k} C_i v_i = -\varepsilon \beta! x_0^\nu \sum_{\eta=0}^{k-\alpha} A_\eta(0) \ . \tag{21}$$

To illustrate the above general consideration we shall give two simple examples, where the polynomial $P_k(n)$ is in the reduced form.

Example 1. $(\alpha = x = 1 \ , \beta \in N_0 \ , A_0(n) = 1)$.

$$\sum_{n=0}^{\infty} \varepsilon^n (n + \beta)!(n + \beta - \varepsilon + 1) = -\varepsilon \beta! \ . \tag{22}$$

Example 2. $(\alpha = x = 1 \ , \beta \in N_0 \ , A_1(n) = n + a_0)$.

$$\sum_{n=0}^{\infty} \varepsilon^n (n + \beta)! [n^2 - (\beta + 1)^2 + (2\beta + 3)\varepsilon - 1] = \varepsilon \beta! (\beta - \varepsilon + 2) \ . \tag{23}$$

4. CONCLUDING REMARKS

From the above consideration it follows that series (7) may have rational sums only for some polynomials $P_k(n)$ of the form (15) with the degree $k \geq \alpha$. In other words, only series of the form (20) have rational sums which are given by (21).

In the Introduction, a possible connection between the divergent series in R and its convergence in all Q_p was mentioned as a motivation for the investigation of p-adic series. Now we can introduce the following method of summation of divergent series at rational points.

Definition. (of adelic summation). *Let a series be divergent in the real case and convergent in Q_p for all but a finite number of p. Let also such series allows a number field invariant summation with a rational sum for some variable $x \in Q$. We call adelic summation of the divergent series extrapolation of the number field invariant summation for convergent counterparts to this divergent series at the same rational points.*

Thus if the series (20), (22) and (23) in the real case make sense at all, then it seems the most natural to assign to them the rational sum (21).

There is no signature that there exists some other summation formula except (13), which is number field invariant and leads to rational summation. Thus there is a reason to introduce the following

Hypothesis. Equation (13) is a unique summation formula which includes all rational sums of the series (7).

As a consequence of the above Hypothesis it follows that we have thus presented all possible series of the form (7) which have rational sums for rational values of x also belonging to the region of convergence. In particular, the most simple examples

$$\sum_{n=0}^{\infty} n!\,, \qquad \sum_{n=0}^{\infty} (-1)^n n!$$

do not have a rational sum. A similar result has been conjectured earlier [6], [9], [13].

REFERENCES

[1] Aref'eva I. Ya., Dragovich B. and Volovich I. V., *On the p-adic summability of the anharmonic oscillator*, Physics Letters B 200 (1988), 512-514.

[2] Brekke L. and Freund P. G. O., *P-adic numbers in physics*, Physics Reports 233 (1993), 1-66.

[3] Dragovich B., *P-adic perturbation series and adelic summability*, Physics Letters B 256 (1991), 392-396.

[4] Dragovich B., *On p-adic aspects of some perturbation series*, Teoreticheskaya i Matematicheskaya Fizika 93 (1992), 211-218.

[5] Dragovich B., *Power series everywhere convergent on R and all Q_p*, Journal of Mathematical Physics 34 (1993), 1143-1148.

[6] Dragovich B., *Rational summation of p-adic series*, Teoreticheskaya i Matematicheskaya Fizika 100 (1994), 342-353.

[7] Dragovich B., *On p-adic series in mathematical physics*, Proceedings of V. A. Steklov Institute of Mathematics 203 (1994), 255-270.

[8] Dragović B., *Adelično sumiranje divergentnih redova*, IX Kongres fizičara Jugoslavije Zbornik radova 17-19. Petrovac na moru, 29-31. maj 1995.

[9] Dragovich B., *On some p-adic series with factorials*, to be published in the Proceedings of the Fourth International Conference on P-adic Functional Analysis, Nijmegen, 1996.

[10] Dragovich B., *Adelic harmonic oscillator*, International Journal of Modern Physics A 10 (1995), 2349-2365.

[11] Khrennikov A., *P-adic Valued Distributions in Mathematical Physics*, Kluwer Academic Publishers, Dordrecht, 1994.

[12] Kurepa G., *On p-adic spaces of Hensel*, Publications de l'institut matematique, Srpska Akad. Nauka i Umet., Beograd 19 (1965), 133-135.

[13] Schikhof W. H., *Ultrametric Calculus. An Introduction to P-adic Analysis*, Cambridge University Press, Cambridge, 1984.

[14] Vladimirov V. S., Volovich I. V. and Zelenov E. I., *P-adic Analysis and Mathematical Physics*, World Scientific, Singapore, 1994.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 105-113

A SEQUENCE $u_{n,m}$ AND KUREPA'S HYPOTHESIS ON LEFT FACTORIAL

Zoran N. ŠAMI
Faculty of Transport and Traffic Engineering University of Belgrade, Vojvode Stepe 305, 11000 Belgrade, Yugoslavia

Abstract. In this work, one can find a definition of a sequence $u_{n,m}$, $n \in N \cup \{0\}$, $m \in Z$ which is a generalisation of the functions $n!$ and $!n$. The author discusses Kurepa's hypothesis and, using a sequence $u_{n,m}$, deduces several hypotheses which are equivalent to Kurepa's.

Keywords: left factorial, Kurepa's hypothesis.

1. INTRODUCTION

In [5] Đ. Kurepa defined the left factorial by:

$$!n = \sum_{k=0}^{n-1} k!,\ n \in N. \tag{1.1}$$

In the same paper a hypothesis was stated

$$(!n, n!) = 2,\ n \in N \wedge n > 1, \tag{KH}$$

where (a, b) denotes the greatest common divisor of integers a and b . It was proved in [5] that KH is equivalent to the assertion that for all prime integers p, $p > 2$,

$$!p \not\equiv 0(\bmod p) \tag{1.2}$$

and this is the most frequent form of considering KH. In [7] - [11] there are several statements equivalent to KH, which are all shown in [4]. KH is verified in [2] for $n < 10^6$, and in [14] for $n < 2^{23}$. In [13], a generalisation of the functions n! and !n is given. This paper is the natural continuation of paper [13].

2. SEQUENCES y_n AND $u_{n,m}$

A sequence y_n is defined in [13] as

$$(2.1) \qquad y_n = \left[\frac{e^{-x}}{1-x}\right]^{(n)}_{x=0} = \sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}k!, \; n \in N \cup \{0\}.$$

We can give a combinatorial meaning to a sequence y_n, where y_n, $n > 0$, represents the number of derangements of the set of n elements, i.e. the number of those mappings of the set of n elements 1-to-1 and onto itself, that do not leave any element invariant, namely, do not map any element onto itself. In [13], beside other things, it is proved that for every $n \in N$, the following equalities hold

$$(2.2) \qquad y_n = ny_{n-1} + (-1)^n$$

$$(2.3) \qquad \sum_{k=0}^{n}\binom{n}{k}y_k = n!$$

$$(2.4) \qquad \sum_{k=1}^{n}\binom{n}{k}y_{k-1} = !n$$

$$(2.5) \qquad \sum_{k=1}^{n}\binom{n}{k}(-1)^{n-k}(!k) = y_{n-1}$$

$$(2.6) \qquad y_n = \left[\frac{n!}{e}\right] + \frac{1+(-1)^n}{2}$$

where $[x]$ denotes the integer part of real number x. From the cited equalities a number of congruences by module p, $p \in P$ follow. Among others, the following holds

$$(2.7) \qquad y_p \equiv -1 \pmod{p},$$

$$(2.8) \qquad y_{p-1} \equiv !p \pmod{p},$$

$$(2.9) \qquad y_m + y_{m+p} \equiv 0 \pmod{p}, \; m \in N \cup \{0\}.$$

A sequence $u_{n,m}$ is defined in [13] by

$$u_{n,m} = u_m^{(n)}(0). \tag{2.10}$$

where

$$u_m(x) = \begin{cases} e^x \int\limits_0^x dt_1 \int\limits_0^{t_1} dt_2 \cdots \int\limits_0^{t_{m-1}} f(t_m)dt_m, \ m > 0 \\ e^x f^{(-m)}(x), m \leq 0 \end{cases} \tag{2.11}$$

and where $f(x) = e^{-x}(1-x)^{-1}$. In [13], it is firstly proved that a sequence $u_{n,m}$ in special cases produces functions $n!$, $!n$ and y_n, i.e. it is shown that for every $n \in N \cup \{0\}$, the following equalities hold

$$u_{n,0} = n!, \tag{2.12}$$

$$u_{n,1} = !n, \ (!0 = 0), \tag{2.13}$$

$$u_{0,-n} = y_n. \tag{2.14}$$

A number of equalities that are satisfied by a sequence $u_{n,m}$ were also proved in [13]. Thus, among other things, it was proved that:

$$u_{n,m} + u_{n,m+1} = u_{n+1,m+1}, \ \ n \in N \cup \{0\}, m \in Z, \tag{2.15}$$

$$m > n \Rightarrow u_{n,m} = 0, \ \ n \in N \cup \{0\}, m \in N, \tag{2.16}$$

$$\sum_{k=m}^{n} (-1)^{k-m} u_{n,k} = u_{n-1,m-1}, \ \ n \in N \cup \{0\}, \ m \in Z, \ m \leq n, \tag{2.17}$$

$$\sum_{k=0}^{n} u_{k,m} = u_{n,m+1} - u_{0,m+1}, \ \ n \in N \cup \{0\}, \ m \in Z, \ m \leq n, \tag{2.18}$$

$$\sum_{i=0}^{n} (-1)^{n-i} \binom{n}{i} u_{k+i,m} = u_{k,m-n}, \ \ k, n \in N \cup \{0\}, \ m \in Z, \tag{2.19}$$

$$\sum_{i=0}^{n} \binom{n}{i} u_{k,m-i} = u_{n+k,m}, \ \ k, n \in N \cup \{0\}, \ m \in Z, \tag{2.20}$$

$$(2.21)\ u_{n,-k}=\sum_{i=0}^{n}\binom{n}{i}y_{i+k}=n!\sum_{i=0}^{k}(-1)^{k-i}\binom{k}{i}\binom{n+i}{i}i!,\ k,n\in N\cup\{0\},$$

$$(2.22)\ (m-1)u_{n,m}=(n-m+1)u_{n,m-1}-u_{n,m-2}+a(n,m),\ n\in N,\ m\in Z,$$

$$(2.23)\ u_{n+2,m}=(n-m+3)u_{n+1,m}-(n+1)u_{n,m}+a(n,m)\ n\in N,\ m\in Z,$$

where $a(n,m)=\begin{cases}\binom{n}{m-2}, & m\geq 2\\ 0, & m<2\end{cases}$.

From the cited equalities, a number of congruences by module p, $p\in P$ follow. Thus, among others, for every $p\in P$ the following holds:

$$(2.24)\qquad u_{k,m}+u_{k,m-p}\equiv u_{p+k,m}\ (\operatorname{mod} p),\ k\in N\cup\{0\},\ m\in Z,$$

$$(2.25)\qquad u_{k,m}\equiv u_{p+k,m}\ (\operatorname{mod} p),\ k\in N\cup\{0\},\ m\in Z,$$

$$(2.26)\qquad u_{p+1,2}\equiv 1\ (\operatorname{mod} p).$$

Let us prove some more properties of a sequence $u_{n,m}$, i.e. especially of sequence y_n.

Theorem 2.1. *For every $n,m\in N$, $m\leq n$, the following holds*

$$(2.27)\qquad y_n=\binom{n}{m}m!y_{n-m}+\sum_{k=0}^{m-1}(-1)^{n-k}\binom{n}{k}k!.$$

Proof. Bearing in mind (2.1), we obtain

$$\begin{aligned}\left[x^m\frac{e^{-x}}{1-x}\right]^{(n)}_{x=0}&=\binom{n}{m}m!y_{n-m}=\\&=\left[e^{-x}\frac{x^m}{1-x}\right]^{(n)}_{x=0}=\\&=\sum_{k=0}^{n}\binom{n}{k}(-1)^{n-k}\sum_{i=0}^{k}\binom{k}{i}i!(x^m)^{(k-i)}_{x=0}=\\&=\sum_{k=m}^{n}\binom{n}{k}(-1)^{n-k}\binom{k}{m}m!(k-m)!=\\&=\sum_{k=m}^{n}\binom{n}{k}(-1)^{n-k}k!=\\&=y_n-\sum_{k=0}^{m-1}(-1)^{n-k}\binom{n}{k}k!\end{aligned}$$

and that is what we wanted to prove.

Theorem 2.2. *Let A and B be sets, both containing $n+k$ elements, such that $A \cap B$ contains k elements. The number of those mappings of set A onto set B that do not map any element onto itself is $u_{n,-k}$.*

Proof. Let

$$A = \{1,2,\ldots,k,k+1,\ldots,k+n\} \text{ and}$$
$$B = \{1,2,\ldots,k,1',2',\ldots,n'\}.$$

Under the conditions of the theorem, none of the elements, $1,2,\ldots,k$, maps onto itself.

Bearing in mind the combinatorial meaning of sequence y_n, one can deduce that, if it had been forbidden for element $k+i$ to map onto i' for $i = 1,2,\ldots,n$, the number of mappings would have been y_{k+n}.

But, according to the conditions of the theorem, it is not forbidden for element $k+i$ to map onto i', $i = 1,2,\ldots,n$, and thus the number, $a_{k,n}$, of mappings is

$$a_{k,n} = \sum_{j=0}^{n} \binom{n}{j} y_{j+k},$$

since neither of the elements $k+i$, $i = 1,2,\ldots,n$, or one of them, or two, ... , or all n of them can be mapped onto i'. Bearing in mind (2.21) it follows that

$$a_{k,n} = u_{n,-k},$$

and that is what we wanted to prove.

Theorem 2.3. *For every $m,n \in N$, $m \leq \frac{n}{2}$ the following holds*

$$y_n = \sum_{k=0}^{m} \binom{m}{k} \binom{n-m}{m-k} u_{m-k,-k} u_{m-k,-(n-2m+k)}. \tag{2.28}$$

Proof. The number of derangements of the set $A = \{1,2,\ldots,n\}$ is y_n.

Let B and C be subsets of set A, such that each contains m elements $(m \leq \frac{n}{2})$.

One can count the number of derangements of set A on following way:

We map B onto C, and $A - B$ onto $A - C$, when B is fixed and C is chosen among all choices.

If we assume that $B \cap C$ contains k elements, $k = 0, 1, \ldots, m$, then for fixed set B, we can choose C in $\binom{m}{k}\binom{n-m}{m-k}$ ways.

According to Theorem 2.2, the number of mappings of the set B onto set C (such that there is no invariant element) is $u_{m-k,-k}$, and the number of mappings of the set $A - B$ onto $A - C$, (also with no invariant element) is $u_{m-k,-(n-2m+k)}$, and thus (2.28) directly follows, and that proves the theorem.

Especially, for $m = n$ and by substituting $2n$ instead of n into (2.28) one obtains

$$y_{2n} = \sum_{k=0}^{n} \left[\binom{n}{k} u_{n-k,-k} \right]^2 . \tag{2.29}$$

Also, substitution $m = p$, $n = kp$, $p \in P$, $k \geq 2$, in (2.28) gives

$$\begin{aligned} y_{kp} &= \sum_{i=0}^{p} \binom{p}{i} \binom{(k-1)p}{p-i} u_{p-i,-i} u_{p-i,-((k-2)p+i)} \equiv \\ &\equiv \binom{(k-1)p}{p} u_{p,0} u_{p,-(k-2)p} + u_{0,-p} u_{0,-(k-1)p} \pmod{p^2}. \end{aligned}$$

Since, according to (2.21)

$$u_{p,0} \equiv u_{p,-(k-2)p} \equiv 0 \pmod{p!}$$

and bearing in mind (2.14), it follows that

$$y_{kp} \equiv y_p y_{(k-1)p} \pmod{p^2}. \tag{2.30}$$

Considering (2.30), by induction on k, one can easily establish that

$$y_{kp} \equiv y_p^k \pmod{p^2}. \tag{2.31}$$

3. KUREPA'S HYPOTHESIS

Using the properties of the sequences $u_{n,m}$ and y_n it is not hard to formulate a number of hypotheses equivalent to KH. In [13], among the other things it was marked that KH is equivalent to every one of the following assertions

$$y_{p-1} \not\equiv 0 \pmod{p}, \quad \forall p \in P \wedge p \geq 3, \tag{3.1}$$

$$y_{p-2} \not\equiv 1 \pmod{p}, \quad \forall p \in P \wedge p \geq 3, \tag{3.2}$$

$$(3.3) \qquad y_p \not\equiv -1 \pmod{p^2}, \quad \forall p \in P \wedge p \geq 3,$$

$$(3.4) \qquad u_{p-1,2} \not\equiv 0 \pmod{p}, \quad \forall p \in P \wedge p \geq 3,$$

$$(3.5) \qquad y_{p+1,2} \not\equiv p+1 \pmod{p^2}, \quad \forall p \in P \wedge p \geq 3,$$

$$(3.6) \qquad (\exists k)(k \geq p \wedge u_{k,2} \not\equiv 1 \pmod{p}), \forall p \in P \wedge p \geq 3.$$

Let us prove some more assertions equivalent to KH.

Theorem 3.1. *KH is equivalent to:* $(\forall p \in P \wedge p \geq 3)(\exists m \in \{0, 1, \ldots, p-1\})$ *such that*

$$(3.7) \qquad y_m \not\equiv\, !(p-m-1)(m!) \pmod{p}.$$

Proof. By substituting $n = p-1$ in (2.27), and bearing in mind (1.2) and (2.8) we obtain

$$!p \equiv y_{p-1} = \binom{p-1}{m} m! y_{p-m-1} + \sum_{k=0}^{m-1} (-1)^{p-k-1} \binom{p-1}{k} k! \equiv$$

$$\equiv (-1)^m m! y_{p-m-1} + \sum_{k=0}^{m-1} k! =$$

$$(3.8) \qquad = (-1)^m m! y_{p-m-1} + !m \not\equiv 0 \pmod{p}.$$

Since

$$m!(p-m-1)! = m!(p-(m+1))(p-(m+2))\cdots(p-(p-1)) =$$
$$\equiv (p-1)!(-1)^{p-m-1} \equiv (-1)^{m+1} \pmod{p},$$

it follows that by substituting $p-m-1$ by m, in (3.8) , we obtain (3.7), and that is what we wanted to prove.

Let us note that, especially, for $m = p-1$ and $m = p-2$, in (3.7), we obtain equalities (3.1) and (3.2). Also, for $m = \frac{p-1}{2}$, (3.7) becomes

$$(3.9) \qquad y_{\frac{p-1}{2}} \not\equiv\, !\left(\frac{p-1}{2}\right)\left(\frac{p-1}{2}\right)! \pmod{p}.$$

Theorem 3.2. *KH is equivalent to the assertion:* $(\forall p \in P \wedge p \geq 3)(\exists n \in N \cup \{0\})$ *such that*

$$u_{n+p,3} \not\equiv n-1 \pmod{p}. \tag{3.10}$$

Proof. Since (2.23), bearing in mind (2.13) and (2.15), we obtain

$$\begin{aligned} u_{n,2} &= (n-1)(u_{n-1,2} - u_{n-2,2}) + 1 = (n-1)u_{n-2,1} + 1 = \\ &= (n-1)(!(n-2)) + 1 \quad (n \geq 2) \end{aligned}$$

and thus from (2.22) it follows that

$$\begin{aligned} 2u_{n+p,3} &= (n+p-2)u_{n+p,2} - u_{n+p,1} + n + p = \\ &= (n+p-2)(n+p-1)(!(n+p-2)) - !(n+p) + 2n + 2p - 2 = \\ &= [(n+p-2)(n+p-1) - 1](!(n+p)) - \\ &- (n+p-2)(n+p-1)((n+p-2)! + (n+p-1)!) + 2n + 2p - 2 = \\ &= [(n+p-2)(n+p-1) - 1](!(n+p)) - \\ &- (n+p-2)(n+p)! + 2n + 2p - 2 \equiv \\ &\equiv (n^2 - 3n + 1)(!p) + 2n - 2 \pmod{p}, \end{aligned}$$

i.e.

$$u_{n+p,3} \equiv \frac{1}{2}(n^2 - 3n + 1)(!p) + n - 1 \pmod{p},$$

and according to (1.2) we conclude that the theorem is correct.

Let us note, that according to the properties of sequence $u_{n,m}$, one could formulate many more assertions equivalent to KH. But the nature of these assertions implies that if KH is the theorem, then the whole problem belongs to the category of hard problems of the number theory.

REFERENCES

[1] Carlitz L., *A note on the left factorial function*, Math. Balcan. 5:6 (1975), 37 - 42.

[2] Gogić G., *Parallel algorithms in arithmetic*, Master thesis, Belgrade Univ., 1991.

[3] Guy R., *Unsolved Problems in Number Theory*, Springer - Verlag, 1981.

[4] Ivić A., Mijajlović Ž., *On Kurepa problems in number theory*, Publ. Inst. Math. 57(71) (1995), 19 - 28.

[5] Kurepa Đ., *On the left factorial function* $!n$, Math. Balcan. 1 (1971), 147 - 153.

[6] Kurepa Đ., *Left factorial function in complex domain*, Math. Balcan. 3 (1973), 297 - 307.

[7] Kurepa Đ., *On some new left factorial propositions*, Math. Balcan. 4 (1974), 383 - 386.

[8] Mijajlović Ž., *On some formulas involving* $!n$ *and the verification of the* $!n$ *– hypothesis by use of computers*, Publ. Inst. Math 47 (61) (1990), 24 - 32.

[9] Stanković J., *Über einige Relationen zwischen Fakultäten und den linken Fakultäten*, Math. Balcan. 3 (1973), 488 - 497.

[10] Stanković J., Žižović M., *Noch einige Relationen zwischen den Fakultäten und den linken Fakultäten*, Math. Balcan. 4 (105) (1974), 555 - 559.

[11] Šami Z., *On the M - hypothesis of Đ. Kurepa*, Math. Balcan. 3 (1973), 530 - 532.

[12] Šami Z., *A sequence of numbers* $a_{j,n}^{m}$, Glas. Math. 23 (43) (1988), 3 -13.

[13] Šami Z., *On a generalisation of functions* $n!$ *and* $!n$, Publ. Inst. Math. 60(74) (1996).

[14] Živković M., *On Kurepa left factorial hypothesis*, Kurepa's symposium, Belgrade 1996.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 115-118

BOURGAIN ALGEBRAS OF TOPOLOGICAL ALGEBRAS†

H. ARIZMENDI
National Autonomous University of Mexico
Mexico City, Mexico

T. TONEV
The University of Montana - Missoula
Missoula, Montana, USA

Abstract. In this note a notion of the Bourgain algebra $(A, B)_b$ is extended to the case of commutative topological algebras. It is proved that $(A, B)_b$ is a closed subalgebra of B, if B^* separates the points of B. It is also shown that $(H(\Delta), C(\Delta))_b = C(\Delta)_b$, where $H(\Delta)$, resp. $C(\Delta)$ denotes Frechet algebra of all holomorphic, resp. continous, functions in the open unit disc $\Delta \subset \mathbf{C}$.

Keywords: Bourgain algebras, topological algebras.

Let B be a commutative Banach algebra with norm $\|\cdot\|$ and let $A \subset B$ be a linear subspace (not necessarily closed). We let $c_o^w(A)$ denote the space of weakly null sequences in A. The notion of the Bourgain algebra of A with respect to B was introduced by J. Cima and R. Timoney [3] in their study of the Dunford-Pettis property of uniform algebras and it is based on a construction of J. Bourgain involving operators of the Hankel type.

The Hankel type operator S_f on A generated by the element $f \in B$ is the mapping $S_f : A \to B/A$ defined as $S_f(g) = \pi_A(fg)$, where $\pi_A : B \to B/A$ is the natural projection. The *Bourgain algebra* $(A, B)_b$ of A with respect

†The research of the second author was partially supported by grants from the IREX and from the NSF (MONTS)

to B is the set of all elements $f \in B$ for which S_f is completely continuous.

Thus $f \in (A, B)_b$ when $S_f(c_o^w(A)) \subset c_o^{\|\cdot\|}(B/A)$. In other words $(A, B)_b$ consists of all $f \in B$ such that for every $\{\varphi_n\} \in c_o^w(A)$ there exists a sequence $\{g_n\}$ in A for which

$$\lim_{n\to\infty} \|\varphi_n f - g_n\| = 0. \tag{1}$$

It is known that $(A, B)_b$ is a commutative Banach algebra; moreover, $A \subset (A, B)_b$ if A is an algebra (e.g. [3]). We call B the *enveloping algebra* of the Bourgain algebra $(A, B)_b$. Note that the structure of $(A, B)_b$ is quite sensitive to B.

The argument is actually valid in the more general setting of commutative topological algebras. Let B be a commutative topological algebra and A its subalgebra. We denote by $c_o^{bw}(A)$ the space of bounded weakly null sequences in A.

Definition 1. *The Bourgain algebra $(A, B)_b$ of A relative to B is the set of all elements $f \in B$ for which $S_f(c_o^{bw}(A)) \subset c_o^{\tau}(B/A)$, where τ is the topology on B/A that is inherited from B.*

In other words $(A, B)_b$ consists of all $f \in B$ such that for every $\{\varphi_n\} \in c_o^{bw}(A)$ there exists a sequence $\{g_n\}$ in A for which

$$\lim_{n\to\infty} (\varphi_n f - g_n) = 0. \tag{2}$$

Proposition 2. *Let $A \subset B$ be commutative topological algebras. Every completely continuous Hankel type operator $S_f : A \to \pi_A(fA)$ maps bounded weakly Cauchy sequences in A onto Cauchy sequences in B/A.*

Proof. Suppose that $\{g_n\}_n$ is a bounded weakly Cauchy sequence in A for which the sequence $\{\pi_A(fg_n)\}_n$ is not Cauchy in B/A. Then there is a neighbourhood U of 0 in B/A such that for every natural $M > 0$ one can find naturals $n_M, m_M \geq M$ with $\pi_A(fg_{n_M}) - \pi_A(fg_{m_M}) \notin U$. Hence the sequence $\{\pi_A(f(g_{n_M} - g_{m_M}))\}_{M=1}^{\infty}$ does not tend to 0 in B/A. The complete continuity of S_f implies that the boubnded sequence $\{g_{n_M} - g_{m_M}\}_{M=1}^{\infty}$ is not weakly null in A. Hence $F(g_{n_M} - g_{m_M})$ does not tend to 0 for some $F \in A^*$. Therefore $\{F(g_n)\}_n$ is not Cauchy, i.e. $\{g_n\}_n$ cannot be a weakly Cauchy sequence.

Note that the dual space B^* does not separate the points of B for every commutative topological algebra B. Local convexity of B is one sufficient condition for this.

Theorem 3. *Let B be a commutative topological algebra and A a subalgebra of B. The Bourgain algebra $(A,B)_b$ is a closed commutative topological subalgebra of B.*

Proof. If $f \in (A,B)_b$ then given a bounded weekly null sequence $\{\varphi_n\} \in c_o^w(A)$, $\varphi_n \in A$, there are elements $h_n \in A$, such that $\varphi_n f - h_n \to 0$. Note that $\{h_n\}$ is a bounded weakly null sequence in A, since $\varphi_n f$ is bounded and tends weakly to 0.

Now let $f_1, f_2 \in (A,B)_b$ and suppose that $\{\varphi_n\}$ is a bounded weakly null sequence in A. By the Bourgain algebra property there are $h_n \in A$ such that $\varphi_n f_1 - h_n \to 0$. By the above remark $\{h_n\}$ is a bounded weakly null sequence in A. Therefore there are $k_n \in A$ such that $h_n f_2 - k_n \to 0$. Now

$$f_1 f_2 \varphi_n - k_n = f_2(f_1\varphi_n - h_n) + (f_2 h_n - k_n) \to 0. \tag{3}$$

Consequently $f_1 f_2 \in (A,B)_b$ and hence $(A,B)_b$ is an algebra.

Let $\{\varphi_n\}$ be a bounded weakly null seqence in A, $f \in B$ is the limit of elements $f_k \in (A,B)_b$, and let U be a bounded set in A that containes $\{\varphi_n\}$.

For a given neighborhood W of 0 in B let V be a neighborhood of 0 such that $V + V \subset W$. Take a neighborhood V_1 of 0 with $V_1^2 \subset V$. There is a $t > 0$ such that $tU \subset V_1$. Let k_0 be such that $f - f_k \in tV_1$ for all $k \geq k_0$. Take such a k and choose $h_n^k \in A$ such that $f_k\varphi_n - h_n^k \to 0$ as $n \to 0$. Then $f\varphi_n - h_n^k = (f - f_k)\varphi_n + (f_k\varphi_n - h_n^k) \in tV_1 \cdot U + V = tU \cdot V_1 + V \subset V_1^2 + V \subset V + V \subset W$ for big enough n. Consequently, $(A,B)_b$ is closed in B.

Note that $A \subset (A,B)_b$ if A is an algebra.

Example 4. *Consider the algebra $B = C(\Delta)$ equipped with compact open topology on the open unit disc Δ and let $A = H(\Delta)$ be the subalgebra of holomorphic functions in Δ. Then $(H(\Delta), C(\Delta))_b = C(\Delta)$.*

Note that both $C(\Delta)$ and $H(\Delta)$ are Frechét algebras. Therefore every weakly null sequence in $C(\Delta)$ is bounded by the uniform boundedness principle (cf. [4], Theorem 2.6).

First we show that the function $\bar{z}$ belongs to $(H(\Delta), C(\Delta))_b$. The calculation is similar to the corresponding one for $C(\bar{\Delta})$ (e.g. [2]). Given a weakly null sequence $\{\varphi_n\}$ in $H(\Delta)$, consider the functions $h_n(z) = \frac{\varphi_n(z)-\varphi_n(0)}{z} \in H(\Delta)$. Note that $h_n(z)$ tends weakly to 0 in $H(\Delta)$ since the map $\varphi \longmapsto h = \frac{\varphi - \varphi(0)}{z} : H(\Delta) \to H(\Delta)$ is a continuous linear operator. From $\varphi_n(z) = zh_n(z) + \varphi_n(0)$ we have that $\bar{z}\varphi_n(z) - h_n(z) = (|z|^2 - 1)\ h_n(z) + \bar{z}\varphi_n(0)$. Fix an $r \in (0,1)$. Then $\max\limits_{z\in\Delta_r} |\bar{z}\varphi_n(z) - h_n(z)| \leq \max\limits_{z\in\Delta_r} |h_n(z)| + |\varphi_n(0)| \to 0$

by Montel's theorem. Therefore $\bar{z}\varphi_n(z) - h_n(z)$ tends to 0 in the compact open topology in Δ. Consequently $\bar{z} \in (H(\Delta), C(\Delta))_b$, as claimed.

Thus $(H(\Delta), C(\Delta))_b$ contains the restrictions of all polynomials in z and $\bar{z}$ on Δ; therefore it contains the algebra $C(\bar{\Delta})$. Since $C(\Delta)$ is the closure of $C(\bar{\Delta})$ in the compact open topology in $C(\Delta)$ we conclude that $(H(\Delta), C(\Delta))_b = C(\Delta)$.

REFERENCES

[1] Bourgain J., *The Dunford-Pettis property for the ball algebra, the polydisc-algebra and the Sobolev spaces*, Studia math. 77 (1984), 245-253.

[2] Cima J., Stroethoff K., Yale K., *Bourgain algebras on the unit disk,* Pacific J. Math. 160 (1993), 27-41.

[3] Cima J., Timoney R., *The Dunford-Pettis property for certain planar uniform algebras,* Michigan Math. J. 34 (1987), 99-104.

[4] Rudin W., *Functional Analysis*, McGraw-Hill, Inc., 1991.

[5] Yale K., *Bourgain algebras,* in: K. Jarosz, (ed.), *Proceedings of the Conference on Function Spaces*, Lecture Notes in Pure and Appl. Math., Vol. 136, Marcel Dekker, 1992, 413-422.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 119-125

ON TYPES OF SETS IN TOPOLOGICAL CLASS SPACES

D. ĆIRIĆ
Department of Mathematics, Faculty of Philosophy
University of Niš, Ćirila i Metodija 2, 18000 Niš, Yugoslavia

Abstract. In our previous papers, we introduced the notion of a class space, i.e. topologies on proper classes, and defined and studied the various topological concepts on such spaces. In this paper we introduce and study a class of τ-regular sets, τ-transitive sets, τ-ordinals and τ-inductive sets, an analogues of corresponding notions in the Set Theory and give some of their properties.

Keywords: class topology, transitive set, inductive sets

1. INTRODUCTION

In our previous papers [2] and [3] we introduced the notion of a class space, i.e. topologies on proper classes, and gave a reason for studying these spaces. First we shall review some notations. Concerning class spaces, we shall use the notations and definitions introduced in [2], [3]. For example, capital letters $X, Y, Z, \ldots$ denote classes, and $x, y, z, \ldots$ sets. Greek letters may stand both for classes and for sets. Our metatheory is based on the NBG class theory if not otherwise stated. Further, we shall assume the usual constructions and definitions from set theory and class theory. For example, we would remind the reader that a class X is transitive if it follows from $x \in y \in X$ that $x \in X$. Throughout the paper K will denote a transitive class. For the sake of convenience, following is a review axioms for class spaces as we shall often refer to them.

Let K be a class and τ and σ be classes of subsets of K. We call triple $\mathcal{K} = (K, \tau, \sigma)$ a topological class space if the following axioms are satisfied:

0. $\emptyset \in \tau, \quad \emptyset \in \sigma$
1. $x, y \in \tau \Rightarrow x \cap y \in \tau$

2. For any i, and $\langle x_j | \ j \in i \rangle$, $(\forall j \in i \ x_j \in \tau) \Rightarrow \bigcup_j x_j \in \tau$
3. For any $a \in K$ there is $x \in \tau$ such that $a \in x$.
4. $\forall x \in \tau \forall y \in \sigma \ \ x - y \in \tau$.

1'. $x, y \in \sigma \Rightarrow x \cup y \in \sigma$.
2'. For any i, and $\langle x_j | \ j \in i \rangle$, $(\forall j \in i \ x_j \in \sigma) \Rightarrow \bigcap_j x_j \in \sigma$
3'. For any subset x of K there is $y \in \sigma$ such that $x \subseteq y$.
4'. $\forall x \in \tau \forall y \in \sigma \ \ y - x \in \sigma$.

Elements of τ are called open subsets while elements of σ are closed subsets of $\mathcal{K}$. By a proposition in [2] it follows that σ is uniquely determined by τ, and vice versa. Also, if K is a set, then $\mathcal{K}$ becomes a standard topological space.

Various topological notions for class spaces, such as continuity, compactness, the product of class spaces, etc., were introduced in our previous papers, and some results concerning these notions were proved. The most important result obtained is that the finite product of compact class spaces is also a compact class space.

In this paper we shall discuss and develop the notion of τ-regular sets, and we shall introduce the notion of τ-ordinal numbers and research some of their properties.

In the following, $\bar{x}$ denotes the closure of set $x \subseteq K$ in class space $\mathcal{K}$. If not otherwise stated, N denotes the set of non-negative integers.

2. τ-REGULAR SETS

Let $\mathcal{K} = (K, \tau, \sigma)$ be a topological class space, where K is a transitive class. The set $S \in K$ satisfies the condition (*) in $\mathcal{K}$, if there exists an $x \in S$ such that $\bar{x} \cap S = \emptyset$.

Proposition 1. *If class K, in topological class space $\mathcal{K} = (K, \tau, \sigma)$, is a model of pairing, union and infinity axioms such that every element of K satisfies the condition (*), then there is no infinite sequence $x_0, x_1, ..., x_n, ...$ so that the following holds:*

$$(**) \qquad \bar{x}_0 \ni x_1, \bar{x}_1 \ni x_2, \ldots, \bar{x}_n \ni x_{n+1}, \ldots$$

In particular, there is no set $x \in K$ such that $x \in \bar{x}$ and there are no "cycles":

$$x_0 \in \bar{x}_1, x_1 \in \bar{x}_2, \ldots, x_n \in \bar{x}_0.$$

Proof. If we suppose that there exists an infinite sequence $x_0, x_1, \ldots, x_n, \ldots$ of elements in class K with property (**), we can consider set $S = \{ x_0, x_1,$

$\ldots, x_n, \ldots\}$. Pairing, union and infinity axioms hold in class K, and set S will also be an element of K, hence set S is a nonempty set and this set satisfies the condition (*). So, there exists an element $x \in S$ with $\bar{x} \cap S = \emptyset$, and also, there exists an index $k \in N$, such that $x = x_k \in S$, so, $\bar{x}_k \cap S = \emptyset$, but this is impossible because $x_{k+1} \in \bar{x}_k \cap S$.

Definition 2. *Let $\mathcal{K} = (K, \tau, \sigma)$ be a topological class space, with a transitive class K. The set $S \in K$ is a τ-regular set if $S \notin \bar{S}$.*

Proposition 3. *Let $K = \bigcup_{\alpha \in ORD} V_\alpha$. It is possible to construct a nondiscrete structure of topological class space on class K, $\mathcal{K} = (K, \tau, \sigma)$, such that every element $S \in K$ is a τ-regular set.*

Proof. Let α be a nonlimit ordinal. Put $X = V_\alpha - V_{\alpha-1}$. Let us assume some nondiscrete topology on X. With the help of the family of all open and closed sets in topological space X, we can define two classes of subsets of K such that $\mathcal{K} = (K, \tau, \sigma)$ becomes a topological class space.

Really, if $S \in K$, then $S \subset K$ and so $S \cap X = \emptyset$ or $S \cap X \neq \emptyset$. Futher $U \in \tau$ if $U \cap X = \emptyset$ and if $U \cap X$ is a nonempty open set in topological space X. A nonempty set F belongs to σ if $F \cap X = \emptyset$ and if $F \cap X$ is a nonempty closed set in topological space X. Put $\emptyset \in \tau \cap \sigma$.

For two nonempty sets x and y in a class τ, we have $(x \cap y) \cap X = (x \cap X) \cap (y \cap X)$, so axiom 1 holds for topological class spaces.

Let $\{x_j | \; j \in i\} \subset \tau$. From $(\cup x_j) \cap X = \cup_j (x_j \cap X)$ it follows that $\cup x_j$ is in class τ. So axiom 2 holds, for topological class spaces, too.

If $a \in K$ then there exists $x \in \tau$ with $a \in x$. Really, if $a \in X$, the set X is an element of τ and axiom 3 holds.

If $x \notin X$, we can have $x \in V_{\alpha-1}$. Then $\{x\} \in \tau$, and $x \in \{x\} \in \tau$. If $x \notin V_\alpha$, then also $\{x\} \in \tau$ with $x \in \{x\}$. So, axiom 3 of topological class spaces holds.

For $x \in \tau$ and $y \in \sigma$ we have $(x - y) \cap X = (x \cap X) - (y \cap X)$ and it is easy to see that axiom 4 holds.

Similarly, it is easy to see the validity of axiom 1' because for $x \in \sigma$ and $y \in \sigma$, we have $(x \cup y) \cap X = (x \cap X) \cup (y \cap X)$.

We have $(\bigcap x_j) \cap X = \bigcap_j (x_j \cap X)$, and so the intersection of the members of σ will be in σ. This means that axiom 2' holds.

If $x \in K$, then there exists a $y \in \sigma$ with $x \subseteq y$. Really, if $x \in X$ then $x \subseteq V_{\alpha-1} \in \sigma$, since $X \cap V_{\alpha-1} = \emptyset$. If $x \notin X$, that might be $x \in V_{\alpha-1}$, and obviously $x \subseteq V_{\alpha-1} \in \sigma$. Otherwise if $x \notin V_\alpha$, then $\alpha < rank(x)$ and $x \subseteq V_{rank(x)}$. We have $V_{rank(x)} \cap X = X$, hence $V_{rank(x)} \in \sigma$ so axiom 3' holds.

Let us prove that the structure of topological class space $\mathcal{K} = (K, \tau, \sigma)$ is nondiscrete, hence, there exists a set $F \in K$ such that $F \neq \bar{F}$.

If $F \cap X = \emptyset$, then $F \in \sigma$ and so we have $F = \bar{F}$. The topology on set X is nondiscrete topology, and there exists a set $F \subsetneq X$ with $F \neq \bar{F}$, where the adherence is in topological space X. $F \in K$ since $F \in V_{\alpha+1} \subset K$. From $\bar{F} \cap X = \bar{F}$, F will be in $\mathcal{K}$, thus $F \neq \bar{F}$.

Finally, let $S \in K$ be an element of class K. We can have $S \cap X = \emptyset$, and in this case $S = \bar{S}$.

From the regularity axiom we cannot have $S \in S = \bar{S}$, and every such set S will be τ-regular set. Let us suppose that $S \cap X \neq \emptyset$. Then $\bar{S} - S \subset X$, but we cannot have $S \in S$, so, if $S \in \bar{S}$ it must be $S \in \bar{S} - S \subset X$, but this is impossible because $\alpha \leq rank(S)$, so we don't have $S \in X \subset V_\alpha$. Hence $S \notin \bar{S} - S$, e.g., $S \notin \bar{S}$, which means set S is a τ-regular set.

So, every set in a topological class space $\mathcal{K} = (K, \tau, \sigma)$ is a τ-regular set.

Theorem 4. *Let $\mathcal{K} = (K, \tau, \sigma)$ be a topological class space, where K is a model of the Set Theory, such that every element of K. satisfies the condition (*). Then for every nonempty class $C \subset K$ there exists $S \in C$, such that $\bar{S} \cap C = \emptyset$.*

Proof. Let $S \in C$. We can have $\bar{S} \cap C = \emptyset$ or $\bar{S} \cap C \neq \emptyset$. If $\bar{S} \cap C = \emptyset$ the theorem is proved. Let us suppose that $\bar{S} \cap C \neq \emptyset$. We can construct a set T in the following way: Put $S_0 = \bar{S}$, $S_1 = \bigcup_{x \in S_0} \bar{x}$, $S_2 = \bigcup_{x \in S_1} \bar{x}, \ldots$, $S_{n+1} = \bigcup_{x \in S_n} \bar{x}, \ldots, \ldots$. Put $T = \bigcup_{n \in N} S_n$. The set $S_0 = \bar{S} \subset T$, and from $\bar{S} \cap C \neq \emptyset$ it follows $T \cap C = X \neq \emptyset$. As class K is a model of the Set Theory and every element in K satisfies the condition (*), it follows $X \in K$ hence $X \neq \emptyset$ and X satisfies the condition (*), so there exists an element $x \in X$ such that $\bar{x} \cap X = \emptyset$. Let us prove that $\bar{x} \cap C = \emptyset$. $x \in X = T \cap C$ means $x \in T$, so there exists $n \in N$ such that $x \in S_n$, and $\bar{x} \subset S_{n+1} \subset T$. Hence, for $y \in \bar{x}$ we have $y \in \bar{x}$ and $y \in C$, e.g. $y \in T$ and $y \in C$, thus $y \in X \cap \bar{x} = \emptyset$, and this is a contradiction.

Definition 5. *Let $\mathcal{K} = (K, \tau, \sigma)$ be a topological class space. The set T is a τ-transitive set if for every $x \in \bar{T}$, it holds that $\bar{x} \subset \bar{T}$.*

If set T is a τ-transitive set, then $\bar{T}$ is a transitive set. It is easy to see that set T can be transitive, but not τ-transitive.

Definition 6. *Let $\mathcal{K} = (K, \tau, \sigma)$ be a topological class space as in Theorem 4. Then we can define a relation $\prec$ on K in the next way: for $\alpha \in K$ and $\beta \in K$ we put $\alpha \prec \beta$ if $\alpha \in \bar{\beta}$.*

For nondiscrete structure of topological class space, there exists $\alpha \in K$ such that $\alpha \neq \bar{\alpha}$. Neither $\alpha \prec \bar{\alpha}$, nor $\bar{\alpha} \prec \alpha$ hold. The first relation means that $\alpha \in \bar{\alpha}$ but this is impossible because every element $\alpha \in K$ is a τ-regular set, also, $\bar{\alpha} \prec \alpha$ means $\bar{\alpha} \in \bar{\alpha}$ which is impossible because of the axiom of regularity. If $\bar{\alpha} = \bar{\beta}$, then the elements α and β are incomparable elements in respect to the relation $\prec$.

Definition 7. *If $\mathcal{K} = (K, \tau, \sigma)$ is a topological class space with the conditions for K as in Theorem 3, we can define the notion of τ-ordinals.*

For a set α, we say it is a τ-ordinal, if it is τ-transitive, and for $\bar{\alpha}$ we say it is a well-ordered set with the relation $\prec$.

Theorem 8. *Let $\mathcal{K} = (K, \tau, \sigma)$ be a topological class space, where K satisfies the conditions as in Theorem 3, then:*

(a) *$\emptyset = \bar{\emptyset}$ is a τ-ordinal.*

(b) *If α is a τ-ordinal and $\beta \in \bar{\alpha}$, then β is a τ-ordinal.*

(c) *If α and β are τ-ordinals, with $\bar{\alpha} \subset \bar{\beta}$, then there exists a τ-ordinal γ such that $\bar{\gamma} = \bar{\alpha}$ and $\gamma \in \bar{\beta}$.*

(d) *If $\mathcal{K}$ is a T_1-topological class space, then for every τ-ordinal α, $\bar{\alpha} \cup \{\alpha\}$ is a τ-ordinal, too.*

(e) *Let ORD_τ be a class of all τ-ordinals, then the relation $\prec$ defined above is a partial ordering on ORD_τ.*

(f) *If $C \subset ORD_\tau$ is a nonempty class of τ-ordinals, then $\bigcap C$ is also a τ-ordinal.*

Proof.

(a) The proof is obvious.

(b) If α is a τ-ordinal and $\beta \in \bar{\alpha}$ then β is a τ-ordinal. Really, $\bar{\alpha}$ is a τ-transitive set and from $\beta \in \bar{\alpha}$ we have $\bar{\beta} \subset \bar{\alpha}$. The set $\bar{\alpha}$ is a well-ordered set with the relation $\prec$, and so $\bar{\beta}$ will be well-ordered with $\prec |_{\bar{\beta}}$. The set β is a τ-transitive set. Really, if $\gamma \in \bar{\beta}$, and $\delta \in \bar{\gamma}$ then $\delta \in \bar{\beta}$, hence $\bar{\gamma} \subset \bar{\beta}$, e.g., the set β is a τ-transitive set. This means set β is a τ-ordinal.

(c) Let α and β be τ-ordinals, with $\bar{\alpha} \subset \bar{\beta}$. The set $\bar{\beta} - \bar{\alpha}$ is a nonempty subset of well-ordered set $\bar{\beta}$, so this set has a minimal element γ with respect to relation $\prec$. The set $\gamma \in \bar{\beta}$ will be a τ-ordinal corresponding to (b) and $\bar{\alpha}$ is an initial segment for $\gamma \in \bar{\beta}$. Also, the initial segment for γ will also be $\bar{\gamma}$, so $\bar{\gamma} = \bar{\alpha}$.

(d) Let us suppose that α is a τ-ordinal in T_1-topological class space $\mathcal{K}$. Let us prove that $\bar{\alpha} \cup \{\alpha\}$ is a τ-ordinal too. Really, the set $\bar{\alpha} \cup \{\alpha\}$ is a τ-transitive set. For $\beta \in \overline{\bar{\alpha} \cup \{\alpha\}} = \bar{\alpha} \cup \overline{\{\alpha\}} = \bar{\alpha} \cup \{\alpha\}$. We can have $\beta \in \bar{\alpha}$ and in this case we have $\bar{\beta} \subset \bar{\alpha} \subset \overline{\bar{\alpha} \cup \{\alpha\}}$, because α is τ-transitive. But we can also have $\beta = \alpha$ and $\bar{\beta} \subset \bar{\alpha}$. Obviously, the set $\overline{\bar{\alpha} \cup \{\alpha\}}$ is a well-ordered set with the relation $\prec$. Let us note that sets $\bar{\alpha} \cup \{\beta\}$ are τ-ordinals for every β with property $\bar{\beta} = \bar{\alpha}$.

(e) Let ORD_τ be a class of all τ-ordinals in topological class space $\mathcal{K}$. For ordinals α and β in ORD_τ, the relation $\alpha \prec \beta$, defined by $\alpha \in \bar{\beta}$, is transitive, since β is τ-transitive. Also, we cannot have $\alpha \prec \beta$ and $\beta \prec \alpha$ because $\alpha \prec \alpha$ is impossible by τ-regularity for all elements in K. For a nondiscrete structure of topological class space $\mathcal{K}$ there exists incomparable elements for relation $\prec$. Also, we have $\bar{\alpha} = \{\beta \in ORD_\tau \mid \beta \prec \alpha\}$.

(f) Let $C \subset ORD_\tau$ be a nonempty class of τ-ordinals. Then $\gamma = \bigcap C$ is a τ-ordinal too. Really, γ is τ-transitive, for $\beta \in \bar{\gamma}$ if and only if $\beta \in \bar{\alpha}$, for every $\alpha \in C$. If $\delta \in \bar{\beta}$ from the transitivity of $\prec$, in $\bar{\alpha}$, it follows that $\delta \in \bar{\gamma}$ and finally $\bar{\beta} \subset \bar{\gamma}$.

Obviously, the set γ is a well ordered set by $\prec$.

Definition 9. *The set S is a τ-inductive set, if $\emptyset \in S$ and if for every $x \in S$ it holds that $\bar{x} \cup \{x\} \in S$.*

If $N_\tau = \bigcap\{X \mid X \text{ is a } \tau - \text{inductive set}\}$, then N_τ will be a τ-inductive set. In T_1-topological class space $\mathcal{K}$, it holds that $N_\tau = N$, where N is the set of all natural numbers.

Proposition 10. *Let $\mathcal{K} = (\mathcal{K}, \tau, \sigma)$ be a T_1-topological class space. If the set X is a τ-inductive set, then $S = \{x \in X \mid \bar{x} \subset \bar{X}\}$ is τ-inductive, too.*

Proof. First, $\emptyset \in S$, because $\emptyset \in X$ and $\emptyset = \bar{\emptyset} \subset \bar{X}$. Let $x \in S$. Then $\bar{x} \cup \{x\} \in X$ and $\overline{\bar{x} \cup \{x\}} = \bar{x} \cup \{x\}$. From $\bar{x} \subset \bar{X}$ and $\{x\} \subset X \subset \bar{X}$ it follows that $\bar{x} \cup \{x\} \subset \bar{X}$.

Proposition 11. *Let $\mathcal{K}$ be a T_1-topological class space. If X is a τ-inductive set, then $S = \{x \in X | x \text{ is a } \tau - \text{transitive set}\}$ is τ-transitive, too.*

Proof. If $x \in S$, then x is τ-transitive. Let us prove that is $\bar{x} \cup \{x\}$ is τ-transitive, too. Really, let $y \in \overline{\bar{x} \cup \{x\}} = \bar{x} \cup \{x\}$. Then $y \in \bar{x}$, and from the τ-transitivity of x, it follows that $\bar{y} \subset \bar{x} \subset \bar{x} \cup \{x\}$ or $y = x$, and so $\bar{y} \subset \bar{x} \subset \bar{x} \cup \{x\}$. $\emptyset \in S$, since $\emptyset \in X$ and $\emptyset$ is τ-transitive.

REFERENCES

[1] Bourbaki N., *Topologie generale*, Hermann, Paris, 1971.

[2] Ćirić D., Mijajlović Ž, *Topologies on classes*, Math. Balkanica (1990).

[3] Ćirić D., Mijajlović Ž, *Class spaces*, forthcoming in Math. Balkanica 1995.

[4] Ćirić D., Mijajlović Ž, *On dimension of class spaces*, forthcoming in FILOMAT 1995 (Proc. of Int. Conf. on Algebra, Logic and Discrete Mathematics, Niš 95).

[5] R. Engelking, *General Topology*, PWN, Warszawa, 1977.

[6] A. Fraenkel and Y. Bar-Hillel, *Foundations of Set Theory*, North-Holland, Amsterdam, 1958.

[7] K. Kunen, *Set Theory*, North-Holland, Amsterdam, 1983.

[8] A. Levy, *Basic Set Theory*, Springer-Verlag, Berlin, 1979.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 127-135

AROUND PSEUDO-DISTANCIAL SPACES †

Ljubiša R.D. KOČINAC
Faculty of Philosophy, University of Niš
18000 Niš, Yugoslavia

Abstract. We consider some classes of spaces which are close to the class of pseudo-distancial (or linearly uniformizable, or τ-metrizable) spaces. A new class of spaces is defined and studied.

Keywords: pseudo-distancial space, biquotient mapping, *lob*-space, biradial space, almost biradial space, almost bisequential space.

AMS Subject Classification (1991): 54E15, 54C10, 54A20, 54A25

1. INTRODUCTION

Let τ be a regular cardinal number. A (Tychonoff) space X is called *τ-metrizable* or *linearly uniformizable* if the topology of X is generated by a uniformity $\mathcal{U}$ on X having a well-ordered base $\mathcal{B}$ of order type τ (where $B_1 \leq B_2 \iff B_1 \supset B_2$ for $B_1, B_2 \in \mathcal{B}$).

These spaces were introduced by Đ. Kurepa in 1934 [13] (in a different form) under the name *pseudo-distancial spaces* (see also [13]-[18], [19], [22]-[24], [25], [5]) as probably the best generalization of the usual metric spaces. This definition is one of the equivalent formulations. Note that pseudo-distancial spaces play an important role in nonlinear numerical analysis. Nowdays there are many papers about τ-metrizable spaces (see the paper [5] and references there in).

A characteristic property of τ-metrizable spaces is τ-additivity. Recall that a space X is *τ-additive* if the intersections of less than τ many open sets in X are open.

Let us mention two characterizations of τ-metrizable spaces.

†* Supported by the Serbian Scientific Foundation, grant N° 04M01.

1.1. (**[26]**) *A regular space X is τ-metrizable if and only if it is τ-additive and has a base which is the union of τ many locally finite collections of open subsets of X.*

1.2. (**[12]**) *A regular space X is τ-metrizable if and only if it is τ-additive and $kdv(X) = \tau$.*

(The *k-developability degree $kdv(X)$* of a space X is the smallest cardinal τ for which there exists a family $\{\mathcal{U}_\alpha : \alpha \in \tau\}$ of open covers of X with the property that for every compact set $C \subset X$ and its every neighborhood V there is a $\mathcal{U}_\alpha$ such that $St(C, \mathcal{U}_\alpha) \subset V$.)

Another typical property of τ-metrizable spaces is that they are *lob*-spaces. A space X is called a *lob*-space if its every point has a linearly ordered (by the reverse inclusion) local base.

A topological group is linearly uniformizable if and only if it is a *lob*-space [21] and a τ-metrizable topological group is a monotonically normal space (R. Heath; see [5]).

In this paper we give some results on spaces which are images of τ-metrizable spaces under different kinds of continuous mappings. A new class of spaces is also considered.

The notation and terminology we use are standard and follow that of [5] and [6]. All spaces are assumed to be regular T_1 and all mappings are continuous surjections. τ is always a regular infinite cardinal.

2. IMAGES OF LINEARLY UNIFORMIZABLE SPACES

In the last years several results concerning images or preimages of τ-metrizable spaces under some kinds of continuous mappings have been obtained. We will mention some of these results.

Let us recall some definitions. A mapping $f : X \to Y$ is:

(**a**) [20] *biquotient* if whenever $\mathcal{F}$ is a filter base in Y accumulating at a point $y \in Y$, then $f^{-1}(\mathcal{F})$ accumulates at some $x \in f^{-1}(y)$;

(**b**) [10] τ-biquotient if whenever $\{B_\alpha : \alpha \in \tau\}$ is a decreasing τ-sequence of subsets of Y accumulating at a point $y \in Y$, then $\{f^{-1}(B_\alpha) : \alpha \in \tau\}$ accumulates at some $x \in f^{-1}(y)$;

(**c**) [6] *pseudo-open* (or *hereditarily quotient*) if whenever B is a subset of Y and $y \in \overline{B}$, then $x \in \overline{f^{-1}(B)}$ for some $x \in f^{-1}(y)$.

Definition 1. ([7]) *A space X is said to be chain-net if for every non-closed set $A \subset X$ there exists a point $x \in \overline{A} \setminus A$ and a τ-sequence $(x_\alpha : \alpha \in \tau)$ in A converging to x. X is radial if for every $A \subset X$ and every $x \in \overline{A}$ there is a τ-sequence $(x_\alpha : \alpha \in \tau)$ in A converging to x.*

Definition 2. ([9]) (see also [2], [4]) *A space X is called biradial if whenever a filter base $\mathcal{F}$ accumulates at a point $x \in X$, then there exists a family $\mathcal{S}$ of subsets of X such that:*

(i) *$\mathcal{S}$ is linearly ordered by $\supset$;*

(ii) $\cap\{S : S \in \mathcal{S}\} = \{x\}$;

(iii) *$\mathcal{S}$ converges to x;*

(iv) *$\mathcal{S}$ is synchronous with $\mathcal{F}$ (i.e. $S \cap F \neq \emptyset$ for every $S \in \mathcal{S}$ and every $F \in \mathcal{F}$).*

Definition 3. ([10]) *A space X is called strongly τ-radial if for every decreasing τ-sequence $(A_\alpha : \alpha \in \tau)$ of subsets of X accumulating at $x \in X$, there exist $x_\alpha \in A_\alpha$ such that the sequence $(x_\alpha : \alpha \in \tau)$ converges to x. X is strongly radial if it is strongly τ-radial for some cardinal τ.*

Images of linearly uniformizable spaces under mappings are shown in the next table. We also indicate a reference about the corresponding result. (Hušek and Kulpa [8;T.2] have proved that the class of all open images of τ-metrizable orderable spaces coincides with the class of all *lob*-spaces X having the property $\chi(x, X) = \tau$ for every non-isolated $x \in X$.)

mapping	space
open	lob [8]
pseudo-open	radial [7]
quotient	pseudo-radial [7]
biquotient	biradial [9]
τ-biquotient	strongly τ-radial [10]
closed irreducible 2-to-1	weak-butterfly [5]
closed irreducible	proto-metrizable [5]

In what follows we use the notion of the *Arhangel'skii number* of a space introduced by the author in 1983 (see [5]).

Definition 4. *The Arhangel'skii number $A(X)$ of a space X is the smallest cardinal τ such that there exists a perfect mapping f from X onto some τ-metrizable space Y.*

Now we are going to give some results involving the Arhangel'skii number.

Definition 5. ([10]) *Let τ be a cardinal. A space X is called a τ-bi-k-space if for every filter base $\mathcal{F}$ accumulating at a point $x \in X$ there exists a* *decreasing τ-sequence $(S_\alpha : \alpha \in \tau)$ of subsets of X converging to a compact set $C \subset X$ and synchronous with $\mathcal{F}$. If X is a τ-bi-k-space for some cardinal τ we say that X is a linearly bi-k-space.*

Definition 6. ([11]) **(a)** *A space X is called monotonically bi-k if there is a cardinal τ such that for each decreasing τ-sequence $(A_\alpha : \alpha \in \tau)$ of subsets of X and each $x \in \cap\{\overline{A_\alpha} : \alpha \in \tau\}$ there exists a decreasing τ-sequence $(S_\alpha : \alpha \in \tau)$ of subsets of X converging to the compact set $C = \cap\{S_\alpha : \alpha \in \tau\}$ and having the property $x \in \cap\{S_\alpha : \alpha \in \tau\}$ and $x \in \overline{S_\alpha \cap A_\alpha}$ for every $\alpha \in \tau$.*

(b) *If in the definition above all the sets A_α are equal to a set A we call X a linearly singly-bi-k-space.*

Definition 7. ([10], [11]) *A space X is linearly singly bi-k (linearly bi-k, monotonically bi-k) if and only if it is a pseudo-open (biquotient, τ-biquotient) image of a space Y with $A(Y) = \tau$ for some cardinal τ.*

Therefore, we have the following table regarding the images of spaces which admit perfect mappings onto linearly uniformizable spaces.

mapping	space
pseudo-open	linearly singly bi-k
biquotient	linearly bi-k
τ-biquotient	monotonically bi-k

3. ALMOST BIRADIAL SPACES

In this section we introduce a new class of spaces and prove some assertions concerning this class.

A filter base $\mathcal{F}$ is called *ω-directed filter base* if for every countable family $\mathcal{A} \subset \mathcal{F}$ there exists a member $F \in \mathcal{F}$ such that $F \subset \cap\mathcal{A}$.

We call a space X *almost biradial* if for every ω-directed filter base $\mathcal{F}$ on X accumulating at a point $x \in X$ there is a (decreasing) chain $\mathcal{S}$ converging to x and synchronous with $\mathcal{F}$.

For an almost biradial space X let $lc(X) \leq \tau$ mean that the length of every chain $\mathcal{S}$ in the definition of almost biradial spaces is $\leq \tau$. When $lc(X) = \omega$, X is said to be an almost bisequential space (see [3], where almost bisequential spaces are defined and studied under the name weakly bisequential spaces as a generalization of bisequential spaces [1], [20]).

Clearly, every biradial space is almost biradial.

Proposition 1. *Every almost biradial space is radial.*

Proof. Let A be a subset of an almost biradial space X and $x \in \overline{A}$. The family $\{A\}$ is an ω-directed filter base that accumulates at x. Since X is almost biradial there is a decreasing chain $\mathcal{S}$ of subsets of X converging to x and synchronous with $\{A\}$. For every $S \in \mathcal{S}$ take a point $x_S \in A \cap S$. We have a chain $(x_S : S \in \mathcal{S})$ of points of A which converges to x and witnesses that X is a radial space.

Recall that a space X is said to be a P-space if the intersection of any countable family of open subsets of X is also open. Then we have:

Proposition 2. *Let X be a P-space. If every subspace of X having cardinality $\leq lc(X)d(X)$ is almost biradial, then X is also almost biradial.*

Proof. Let A be a dense subset of X such that $|A| \leq \tau \equiv lc(X)d(X)$ and let $\mathcal{F}$ be an ω-directed filter base on X accumulating at some $x \in X$. As X is a P-space, the collection

$$\mathcal{F}_A = \{A \cap U \ : \ U \text{ is open in } X \text{ and contains some } F \in \mathcal{F}\}$$

is an ω-directed filter base on A and accumulates at x (in fact, converges to x). By assumption, the set $A \cup \{x\}$ is almost biradial, hence there exists a chain $\mathcal{S}_A$ of subsets of $A \cup \{x\}$ converging to x and synchronous with $\mathcal{F}_A$. The family $\mathcal{K} = \{Cl_X(M) : M \in \mathcal{F}_A\}$ is an ω-directed filter base on X that converges to x. The proof will be completed if we prove that $\mathcal{K}$ is synchronous with $\mathcal{F}$. Suppose it is not and take $F \in \mathcal{F}$ and $Cl_X(M) \in \mathcal{K}$ such that $F \cap Cl_X(M) = \emptyset$. Then $X \setminus Cl_X(M)$ is an open subset of X which contains $F \in \mathcal{F}$ so that $X \setminus Cl_X(M) \in \mathcal{F}_A$. Therefore, $(X \setminus Cl_X(M)) \cap M \neq \emptyset$ which is impossible. This contradiction proves that $\mathcal{K}$ and $\mathcal{F}$ are synchronous.

Remark. If every subspace A of a space X such that $|A| \leq lc(X)d(X)$ is biradial, then X is itself biradial.

Theorem 1. *Every almost radial space X of countable pseudocharacter is almost bisequential.*

Proof. Let $\mathcal{F}$ be an ω-directed filter base on X accumulating at a point $x \in X$ and let $\{U_i : i \in \mathrm{N}\}$ be a family of open subsets of X such that $\cap\{U_i : i \in \mathrm{N}\} = \{x\}$. (We suppose that $x \notin F_0$ for some $F_0 \in \mathcal{F}$, because otherwise the proof is trivial.) Since X is almost biradial, there is a chain $\mathcal{S}$ of subsets of X which converges to x and synchronous with $\mathcal{F}$. For every $i \in \mathrm{N}$ choose an element S_i in $\mathcal{S}$ such that $x \in S_i \subset U_i$ and put $\mathcal{S}^* = \{S_i : i \in \mathrm{N}\}$. It is clear that $\mathcal{S}^*$ is a chain synchronous with $\mathcal{F}$. Let us show that $\mathcal{S}^*$ converges to x. Suppose, on the contrary, that there exists a neighbourhood W of x such that $S_i \setminus W \neq \emptyset$ for each $i \in \mathrm{N}$. One can find a member $T \in \mathcal{S}$ for which $x \in T \subset W$ holds. From the fact that $\mathcal{S}$ is a chain it follows that $T \subset S_i$ for every $i \in \mathrm{N}$ so that we have

$$T \subset \cap\{S_i : i \in \mathrm{N}\} \subset \{U_i : i \in \mathrm{N}\} = \{x\}.$$

However, $T \cap F_0 \neq \emptyset$ implies $T \cap (X \setminus \{x\}) \neq \emptyset$ and we obtain a contradiction which proves that $\mathcal{S}^*$ converges to x. So, X is an almost bisequential space.

Theorem 2. *Every regular ccc almost biradial P-space X is almost bisequential.*

Proof. Let $\mathcal{F}$ be an ω-directed filter base on X accumulating at a point $x \in X$. Let

$$\mathcal{A} = \{A : A \text{ is open in } X \text{ and } A \supset F \text{ for some } F \in \mathcal{F} \text{ or } x \in int(\overline{A})\}.$$

Since X is a P-space the family $\mathcal{A}$ has the countable intersection property and so there is an ω-directed filter base $\mathcal{B}$ containing $\mathcal{A}$. Clearly, $\mathcal{B}$ accumulates at x. As X is an almost biradial space there exists a chain $\mathcal{S}$ of subsets of X converging to x and synchronous with $\mathcal{B}$. Consider the family $\mathcal{S}^* = \{int(\overline{S}) : S \in \mathcal{S}\}$. The regularity of X implies that the chain $\mathcal{S}^*$ converges to x. We are going to show that for every $S^* \in \mathcal{S}^*$ and every $F \in \mathcal{F}$ we have $\overline{S^*} \cap F \neq \emptyset$. Suppose that there are $S^* \in \mathcal{S}^*$ and $F \in \mathcal{F}$ such that $\overline{S^*} \cap F = \emptyset$. Then $X \setminus \overline{S^*} \in \mathcal{A}$ and, since the set $M = \overline{S^*} \setminus int(\overline{S^*})$ is nowhere dense, $X \setminus M \in \mathcal{A}$. We have

$$S^* \cap (X \setminus M) \cap (X \setminus \overline{S^*}) \subset int(\overline{S^*}) \cap (X \setminus int(\overline{S^*})) = \emptyset.$$

However, this contradicts the fact that $(X \setminus M) \cap (X \setminus \overline{S^*}) \neq \emptyset$.

The chain $\{\overline{S^*} : S^* \in \mathcal{S}^*\}$ shows that in X one can take chains having cardinality $\leq |\mathcal{S}^*|$. As elements of $\mathcal{S}^*$ are regular open sets and every chain consisting of such sets has cardinality $\leq c(X)$ we conclude that $lc(X) \leq \omega$, i.e. X is an almost bisequential space.

Theorem 3. *Let X be a regular almost biradial space with the Baire property. Then every point x in X is the limit of a chain consisting of (regular) open subsets of X.*

Proof. Let x be an arbitrary point in X. Let $\mathcal{N}$ denote the collection of all subsets of X which are unions of countably many nowhere dense subsets of X. Of course, if $\mathcal{K}$ is a countable subcollection of $\mathcal{N}$, then $\cup\mathcal{K} \in \mathcal{N}$. Therefore, the family $\mathcal{F} = \{X \setminus M : M \in \mathcal{N}\}$ is an ω-directed filter base on X, and, since every $F \in \mathcal{F}$ is a dense subset of X (because X has the Baire property), $\mathcal{F}$ accumulates at x. Since X is almost biradial there exists a chain $\mathcal{S}$ of subsets of X converging to x and synchronous with $\mathcal{F}$. It follows from this that $\mathcal{N} \cap \mathcal{S} = \emptyset$, hence $\mathcal{S}$ does not contain a member which is nowhere dense in X, i.e. for every $S \in \mathcal{S}$ we have $int(\overline{S}) \neq \emptyset$. Let $\mathcal{A} = \{int(\overline{S}) : S \in \mathcal{S}\}$. Then all elements of the chain $\mathcal{A}$ are regular open sets, $\mathcal{A}$ is synchronouus with $\mathcal{F}$ and since X is a regular space, this chain converges to x.

Corollary 1. *If X is a regular ccc almost biradial space with the Baire property, then $|X| \leq 2^{\omega}$.*

Proof. Every chain in X consisting of regular open subsets of X cannot have more than $c(X)$ members. By the previous theorem X is an almost bisequential space. The cardinality of every regular almost bisequential space is $\leq 2^{c(X)}$ [3;Prop. 6].

REFERENCES

[1] Arhangel'skii A.V., *On invariants of character and weight type*, Trudy Moskov. Mat. Obshch. 38 (1979), 3–23.

[2] Arhangel'skii A.V., *On biradial topological spaces and groups*, Topology Appl. 36 (1990), 173–180.

[3] Arhangel'skii A.V., *Bisequential spaces, tightness of products, and metrizability conditions in topological groups*, Trudy Moskov. Mat. Obshch. 55 (1994), 268–284.

[4] Arhangel'skii A.V., Bella A., *The product of biradial compact spaces*, Topology Appl. 45 (1992), 157–162.

[5] Cammaroto F., Kočinac Lj., *Some results on ω_μ-metrizable and related spaces*, Bolletino Un. Mat. Italiana (7) 7-B (1993), 607–629

[6] Engelking R., *General Topology*, PWN, Warszawa, 1977.

[7] Herrlich H., *Quotienten geordneter Räume und Folgenkonvergenz*, Fund. Math. 61 (1967), 79–81.

[8] Hušek M., Kulpa W., *Open images of orderable spaces*, Proc. Amer. Math. Soc. 88 (1983), 711–715.

[9] Kočinac Lj., *Bi-quotient images of ordered spaces*, Publ. Inst. Math. (Beograd) 39(53) (1986), 173–177.

[10] Kočinac Lj., *Two theorems concerning bi-quotient maps*, Matem. vesnik 38 (1986), 58–64.

[11] Kočinac Lj., *Linearly singly bi-k-spaces*, Publ. Inst. Math. (Beograd) 45(59) (1988), 169–177.

[12] Kočinac Lj., *A characterization of ω_μ-metrizable spaces*, Zbornik rad. Fil. fak. (Niš), Ser. Mat. 4 (1990), 79–82.

[13] Kurepa Đ., *Tableaux ramifiés d'ensembles. Espaces pseudodistanciés*, C.R. Acad. Sci. Paris 198 (1934), 1563–1565.

[14] Kurepa Đ., *Ensembles ordonnés et ramifiés*, Publ. Math. Univ. Belgrade 4 (1935), 1–138.

[15] Kurepa Đ., *Le problème de Souslin et les espaces abstraits*, C.R. Acad. Sci. Paris 203 (1936), 1049-1052.

[16] Kurepa Đ., *Sur les classes $(\mathcal{E})$ et $(\mathcal{D})$*, Publ. Math. Univ. Belgrade 5 (1936), 124–132.

[17] Kurepa Đ., *Sur l'écart abstrait*, Glasnik Mat.-Fiz. Astr. 11 (1956), 105–134.

[18] Kurepa Đ., *On the existence of pseudo-metric non-totally orderable spaces*, Glasnik Mat.-Fiz. Astr. 18 (1963), 183–194.

[19] Mamuzić Z.P., *Introduction to General Topology*, Beograd, 1960; Noordhoff, Groningen, 1963.

[20] Michael E., *A quintuple quotient quest*, Gen. Top. Appl. 2 (1972), 91–138.

[21] Nyikos P., Reichel H.-C., *Topologicaly orderable groups*, Gen. Top. Appl. 5 (1975), 195–204.

[22] Papić P., *Sur les espaces pseudo-distanciés*, Glasnik Mat.-Fiz. Astr. 9 (1954), 217–228.

[23] Papić P., *Sur les espaces pseudo-distanciés complètes*, Glasnik Mat.-Fiz. Astr. 11 (1956), 135–142.

[24] Papić P., *Sur une classe d'ensembles ordonnés et les espaces pseudo-distanciés*, Glasnik Mat.-Fiz. Astr. 11 (1956), 161–168.

[25] Todorčević S., *On a theorem of Đ. Kurepa*, in: *Topology and Order Structures* I, Math. Centre Tracts 142, Amsterdam (1981), 173-176.

[26] Wang Shu-Tang, *Remarks on ω_μ-additive spaces*, Fund. Math. 55 (1964), 101–112.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 137-146

A SEQUENCE OF KUREPA'S FUNCTIONS †

Gradimir V. MILOVANOVIĆ
Faculty of Electronic Engineering, Department of Mathematics
University of Niš, P.O. Box 73, 18000 Niš, Yugoslavia

Dedicated to the memory of Professor DJ. Kurepa

Abstract. In this paper we define and study a sequence of functions $\{K_m(z)\}_{m=-1}^{+\infty}$, where $K_{-1}(z) = \Gamma(z)$ is the gamma function and $K_0(z) = K(z)$ is the Kurepa function [5]–[6]. We give several properties of $K_m(z)$ including a discussion on their zeros and poles.

Keywords: Gamma function, Kurepa function, left factorial, meromorphic function, zeros, poles.

AMS Subject Classification (1991): Primary 33B15, 65D20.

1. INTRODUCTION

The left factorial function $z \mapsto K(z)$ was defined by Professor DJ. Kurepa (see [5]–[6]) in the following way

$$K(z) = \int_0^\infty \frac{t^z - 1}{t - 1} e^{-t}\, dt \qquad (\mathrm{Re} z > 0). \tag{1}$$

Firstly, he introduced the so-called left factorial as

$$!0 = 0, \quad !n = 0! + 1! + \cdots + (n-1)! \quad (n \in \mathbf{N})$$

and then extended it to the right side of the complex plane by (1). The function $K(z)$ can be extended analytically to the whole complex plane by

$$K(z) = K(z+1) - \Gamma(z+1), \tag{2}$$

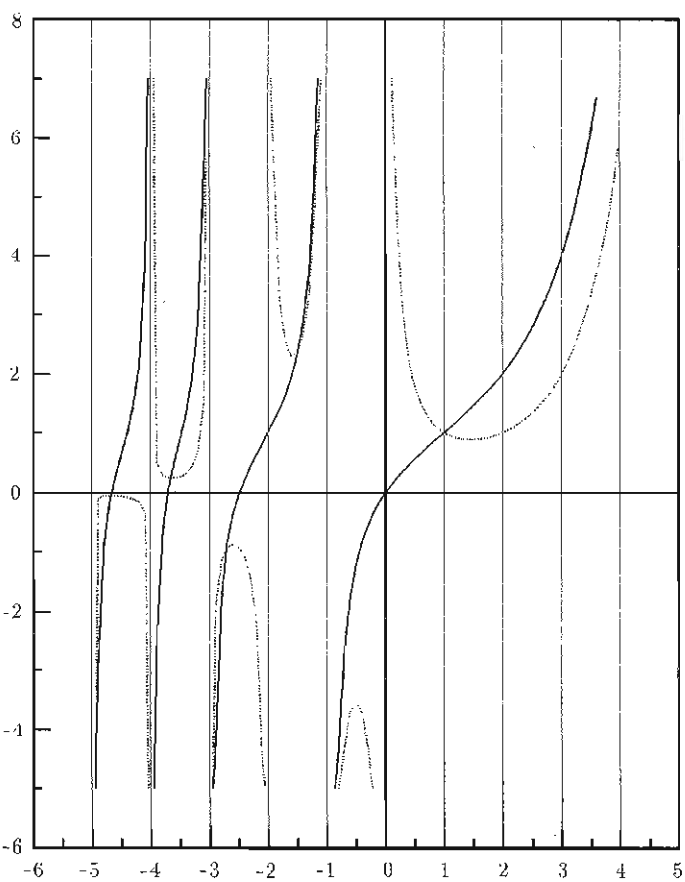
†This work was supported in part by the Serbian Scientific Foundation, grant number 04M03.

where $\Gamma(z)$ is the gamma function defined by

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1}\, dt \quad (\mathrm{Re}\, z > 0) \quad \text{and} \quad z\Gamma(z) = \Gamma(z+1).$$

Kurepa [6] proved that $K(z)$ is a meromorphic function with simple poles at the points $z_k = -k$ ($k \in \mathrm{N} \setminus \{2\}$). Graphs of the gamma and Kurepa functions for real values of z are displayed in Fig. 1.

Fig. 1: The gamma function $\Gamma(x) = K_{-1}(x)$ (dotted line) and the Kurepa function $K(x) = K_0(x)$ (solid line)



Slavić [10] found the representation

$$K(z)=-\frac{\pi}{e}\cot \pi z+\frac{1}{e}\left(\sum_{n=1}^{\infty}\frac{1}{n!n}+\gamma\right)+\sum_{n=0}^{\infty}\Gamma(z-n),$$

where γ is Euler's constant. These formulas were also mentioned in the book [8]. A number of problems and hypotheses, especially in number theory, were posed by Kurepa and then considered by several mathematicians. For details and a complete list of references see a recent survey written by Ivić and Mijajlović [4].

In this paper we define and study a sequence of complex functions $\{K_m(z)\}_{m=-1}^{+\infty}$, such that the first two terms are the gamma function and the Kurepa function, i.e., $K_{-1}(z)=\Gamma(z)$ and $K_0(z)=K(z)$. In Section 2 we give the basic definition of the sequence $\{K_m(z)\}_{m=-1}^{+\infty}$ and main properties of such functions including their graphs for the real values of z. Zeros and poles of $K_m(z)$ are discussed in Section 3. Numerical calculations, series expansions, as well as some applications of such functions will be given elsewhere.

2. BASIC DEFINITIONS AND PROPERTIES

Definition 1. *The polynomials $t\mapsto Q_m(t;z)$, $m=-1,0,1,2,\ldots$, are defined by*

$$Q_{-1}(t;z)=0,\quad Q_m(t;z)=\sum_{\nu=0}^{m}\binom{m+z}{\nu}(t-1)^{\nu}. \tag{3}$$

For example, a few first polynomials are given by

$$\begin{aligned}
Q_0(t;z)&=1,\\
Q_1(t;z)&=1+(z+1)(t-1),\\
Q_2(t;z)&=1+(z+2)(t-1)+\frac{1}{2}(z^2+3z+2)(t-1)^2,\\
Q_3(t;z)&=1+(z+3)(t-1)+\frac{1}{2}(z^2+5z+6)(t-1)^2\\
&\qquad+\frac{1}{6}(z^3+6z^2+11z+6)(t-1)^3.
\end{aligned}$$

It is easy to see that the following result holds:

Lemma 1. *For every $m \in \mathbf{N}_0$ we have*

$$Q_m(t;z) = Q_{m-1}(t;z+1) + \frac{1}{m!}(z+1)(z+2)\cdots(z+m)(t-1)^m.$$

This Lemma can be useful for constructing the polynomials $Q_m(t;z)$.

If we define Δ_z as the standard forward difference operator

$$\Delta_z f(z) = f(z+1) - f(z),$$

then equality (2) can be expressed in the form

$$\Delta_z K_0(z) = K_{-1}(z+1),$$

where we put $K(z) = K_0(z)$ and $\Gamma(z) = K_{-1}(z)$. Our goal here is to define the functions $K_m(z)$, $m = 1, 2, \ldots$, such that

$$\Delta_z K_m(z) = K_{m-1}(z+1), \qquad m = 0, 1, \ldots .$$

In our considerations we also use the k-th order difference operator Δ_z^k, defined inductively as

$$\Delta_z^0 f(z) \equiv f(z), \qquad \Delta_z^k f(z) = \Delta_z \left(\Delta_z^{k-1} f(z)\right) \quad (k \in \mathbf{N}).$$

Firstly, we prove the following auxiliary result:

Lemma 2. *For every $m \in \mathbf{N}_0$ we have*

$$\Delta_z Q_m(t;z) = (t-1) Q_{m-1}(t;z+1).$$

Proof. According to previous definition we have

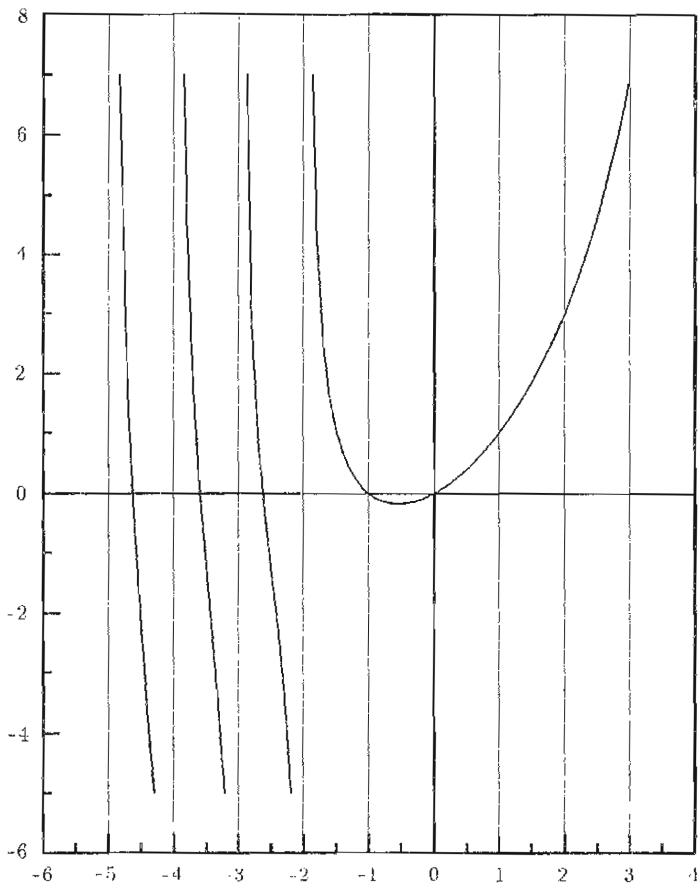
$$\begin{aligned}
\Delta_z Q_m(t;z) &= Q_m(t;z+1) - Q_m(t;z) \\
&= \sum_{\nu=0}^{m} \binom{m+z+1}{\nu}(t-1)^\nu - \sum_{\nu=0}^{m} \binom{m+z}{\nu}(t-1)^\nu \\
&= \sum_{\nu=1}^{m} \binom{m+z}{\nu-1}(t-1)^\nu \\
&= (t-1)\sum_{\nu=0}^{m-1} \binom{m-1+z+1}{\nu}(t-1)^\nu \\
&= (t-1) Q_{m-1}(t;z+1).
\end{aligned}$$

Definition 2. *The sequence $\left\{K_m(z)\right\}_{m=-1}^{+\infty}$ is defined by*

$$K_m(z) = \int_0^{+\infty} \frac{t^{z+m} - Q_m(t;z)}{(t-1)^{m+1}}\, e^{-t}\, dt \qquad (Re z > 0), \tag{4}$$

where $Q_m(t;z)$ given by (3).

Fig. 2: The function $K_1(x)$



Theorem 1. *For* $Re z > 0$ *we have*

$$\Delta_z K_m(z) \equiv K_m(z+1) - K_m(z) = K_{m-1}(z+1)$$

and

$$\Delta_z^i K_m(z) = K_{m-i}(z+i), \qquad i = 1, 2, \ldots, m+1.$$

Proof. Using Lemma 2 we obtain

$$\begin{aligned} \Delta_z\big(t^{z+m} - Q_m(t;z)\big) &= t^{z+1+m} - t^{z+m} - \Delta_z Q_m(t;z) \\ &= (t-1)\big[t^{z+m} - Q_{m-1}(t;z+1)\big]. \end{aligned}$$

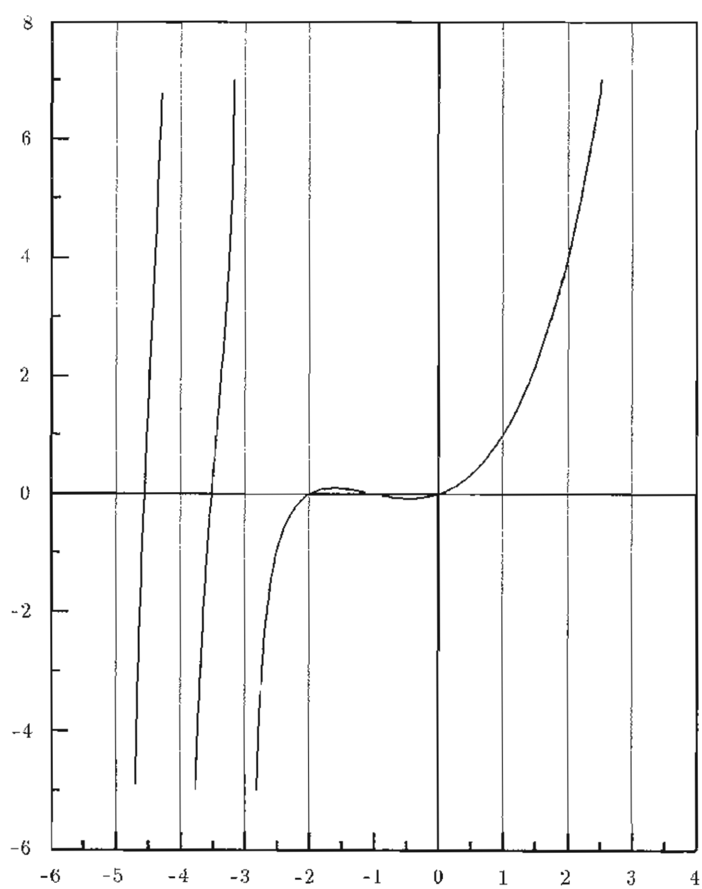
Then

$$\begin{aligned}
\Delta_z K_m(z) &= \int_0^{+\infty} \Delta_z \left[\frac{t^{z+m} - Q_m(t;z)}{(t-1)^{m+1}} \right] e^{-t}\, dt \\
&= \int_0^{+\infty} \frac{t^{z+m} - Q_{m-1}(t;z+1)}{(t-1)^m} e^{-t}\, dt \\
&= K_{m-1}(z+1).
\end{aligned}$$

Iterating we obtain

$$\Delta_z^i K_m(z) = \Delta_z^{i-1} K_{m-1}(z+1) = \Delta_z^{i-2} K_{m-2}(z+2) = \cdots = K_{m-i}(z+i).$$

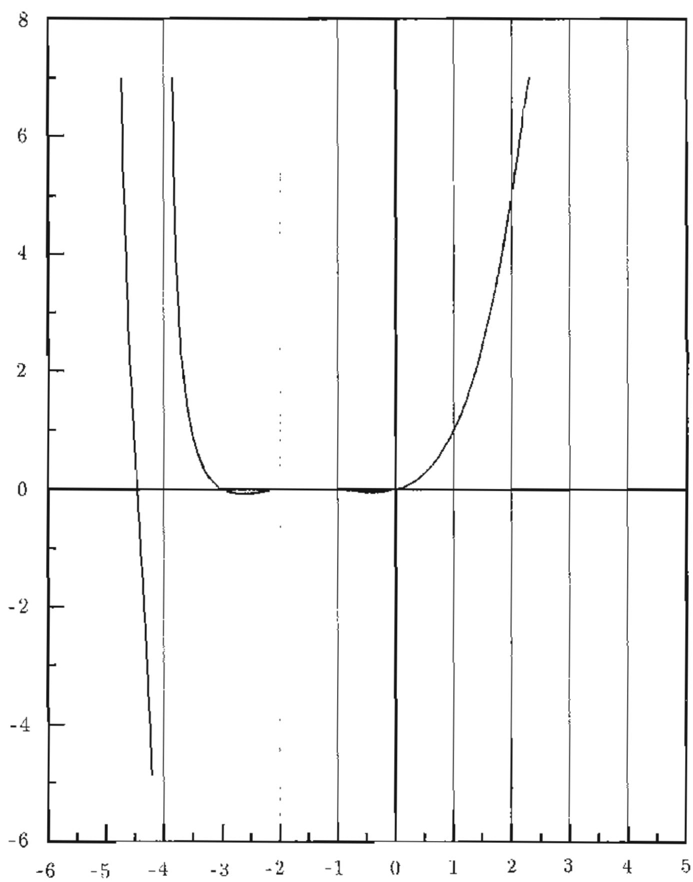
Fig. 3: The function $K_2(x)$



For $i = m + 1$ we find

$$\Delta_z^{m+1} K_m(z) = K_{-1}(z + m + 1) = \Gamma(z + m + 1).$$

Fig. 4: The function $K_3(x)$



It is easy to see that for nonnegative integers the following result holds:

Theorem 2. *For $n, m \in \mathbf{N_0}$ we have*

$$K_m(n) = \sum_{i=0}^{n-1} \frac{(-1)^i}{i!} \sum_{\nu=i}^{n-1} \nu! \binom{m+n}{\nu + m + 1}, \qquad K_m(0) = 0.$$

If we put

$$S_\nu = \nu! \sum_{i=0}^{\nu} \frac{(-1)^i}{i!} \qquad (\nu \geq 0),$$

i.e., $S_\nu = \nu S_{\nu-1} + (-1)^\nu$ with $S_0 = 1$, then $K_m(n)$ can be expressed in the following form

$$K_m(n) = \sum_{\nu=0}^{n-1} \binom{m+n}{\nu+m+1} S_\nu.$$

Since

$$S_0 = 1, \quad S_1 = 0, \quad S_2 = 1, \quad S_3 = 2, \quad S_4 = 9, \quad S_5 = 44, \quad \text{etc.},$$

we have

$$\begin{aligned} &K_m(0) = 0, \quad K_m(1) = 1, \quad K_m(2) = m + 2, \\ &K_m(3) = \frac{1}{2}(m^2 + 5m + 8), \\ &K_m(4) = \frac{1}{6}(m^3 + 9m^2 + 32m + 60), \end{aligned}$$

etc.

The function $K_m(z)$, $m \in \mathrm{N}$, can be extended analytically to the whole complex plane by

$$K_m(z) = K_m(z+1) - K_{m-1}(z+1). \tag{5}$$

Suppose that we have analytic extensions for all functions $K_\nu(z)$, $\nu < m$. Using (4) and (5) we define $K_m(z)$ at first for z satisfying $\mathrm{Re}z > -1$, then for $\mathrm{Re}z$ such that $\mathrm{Re}z > -2$, etc. In this way we obtain the function $K_m(z)$ in the whole complex plane.

An evaluation of the Kurepa function $K_0(z)$ for some specific z in $(0,1)$, using quadrature formulas with relatively small accuracy, was made by Slavić and the author of this paper (see [6]). Recently, we [9] gave power series expansions of the Kurepa function $K_0(a+z)$, $a \geq 0$, and determined numerical values of their coefficients $b_\nu(a)$ for $a = 0$ and $a = 1$, in high precision (Q-arithmetic with machine precision $\approx 1.93 \times 10^{-34}$). Using an asymptotic behaviour of $b_\nu(a)$, when $\nu \to \infty$, we gave a transformation of series with much faster convergence. Also, we obtained the Chebyshev expansions for $K_0(1+z)$ and $1/K_0(1+z)$. For similar expansions of the gamma function see, e.g. Davis [2], Luke [7], Fransén and Wrigge [3], and Bohman and Fröberg [1].

Graphs of functions $K_m(x)$, $m = 1, 2, 3$, for real values of x are displayed in figures 2, 3, and 4, respectively.

3. ZEROS AND POLES

Firstly, we prove:

Theorem 3. *For each $m \in \mathrm{N}_0$ we have*

$$K_m(-n) = 0, \qquad n = 0, 1, \ldots, m. \tag{6}$$

Proof. For $m = 0$ the statement is true ($K_0(0) = 0$).

Suppose that (6) holds. Then, for $n = 0, 1, \ldots, m$, we have

$$K_{m+1}(-n-1) = K_{m+1}(-n) - K_m(-n) = K_{m+1}(-n).$$

Since $K_{m+1}(0) = 0$, we conclude that

$$K_{m+1}(-1) = K_{m+1}(-2) = \cdots = K_{m+1}(-m) = K_{m+1}(-m-1).$$

Similarly as in [6] we can conclude that the function $K_m(z)$ ($m \geq 1$) has an infinite strictly decreasing sequence $\{\xi_k^{(m)}\}_{k=0}^{+\infty}$ of zeros

$$\begin{aligned} &\xi_0^{(m)} = 0, \quad \xi_1^{(m)} = -1, \quad \ldots, \quad \xi_m^{(m)} = -m, \\ &\xi_k^{(m)} \in (-k-1, -k) \qquad (m+1 \leq k \in \mathrm{N}). \end{aligned}$$

Poles of $K_m(z)$ are in the points $z_n^{(m)} = -n$ ($n = m+1, m+2, \ldots$), except the point $z_2^{(0)}$ when $K_0(z_2^{(0)}) = K_0(-2) = 1$.

The poles of gamma function $\Gamma(z) = K_{-1}(z)$ are $z_n^{(-1)} = -n$, $n = 0, 1, \ldots$, with the corresponding residues

$$\mathrm{Res}_{z=-n}\Gamma(z) = \frac{(-1)^n}{n!} \qquad (n = 0, 1, \ldots).$$

Putting

$$R_n^{(m)} = \mathrm{Res}_{z=-n} K_m(z) \qquad (n \geq m+1),$$

we can prove by induction the following result:

Theorem 4. *For every $n \geq m+3$ we have that*

$$R_n^{(m)} = R_{m+2}^{(m)} - \sum_{\nu=m+2}^{n-1} R_\nu^{(m-1)},$$

where

$$R_{m+1}^{(m)} = (-1)^{m+1}, \qquad R_{m+2}^{(m)} = m(-1)^{m+1}.$$

For $m = 0$ Theorem 4 reduces to Kurepa's result [6], §6:

$$R_1^{(0)} = \mathrm{Res}_{z=-1} K_0(z) = -1,$$
$$R_n^{(0)} = \mathrm{Res}_{z=-n} K_0(z) = -\sum_{\nu=2}^{n-1} \frac{(-1)^\nu}{\nu!} = -\frac{1}{e}. \quad (n \geq 3).$$

We note that $z = -2$ is not a pole of $K_0(z)$ ($R_2^{(0)} = 0$).

REFERENCES

[1] J. Bohman and C.-E. Fröberg, *The Γ-function revisited: power series expansions and real-imaginary zero lines*, Math. Comp. 58 (1992), 315–322.

[2] H.T. Davis, *Tables of the Higher Mathematical Functions*, vol. I, Principia Press, Bloomington, IN, 1933.

[3] A. Fransén and S. Wrigge, *High-precision values of the gamma function and some related coefficients*, Math. Comp. 34 (1980), 553–566.

[4] A. Ivić and Ž. Mijajlović, *On Kurepa problems in number theory*, Publ. Inst. Math. (N.S.) 57 (71) (1995), 19–28.

[5] DJ. Kurepa, *On the left factorial function* $!n$, Math. Balkanica 1 (1971), 147–153.

[6] DJ. Kurepa, *Left factorial function in complex domain*, Math. Balkanica 3 (1973), 297–307.

[7] Y.L. Luke, *Mathematical Functions and Their Approximations*, Mir, Moscow, 1980 (Russian).

[8] O.I. Marichev, *Handbook of Integral Transformation of Higher Transcendental Functions: Theory and Algorithmic Tables*, Ellis Horwood Ltd., Chichester, 1983

[9] G.V. Milovanović, *Expansions of the Kurepa function*, Publ. Inst. Math. (N.S.) 57 (71) (1995), 81–90.

[10] D.V. Slavić, *On the left factorial function of the complex argument*, Math. Balkanica 3 (1973), 472–477.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 147-153

A CRITERION FOR UNIFORM CONTINUITY USING SEQUENCES

N. SHEKUTOVSKI
Faculty of Natural Sciences and Mathematics,
University "Kiril i Metodij"
Institute of Mathematics, P.O. 162, FYR Macedonia

Abstract. In this paper a characterization of uniform continuity is presented, analogous to the well-known characterization of continuity using sequences.

Keywords: metric space, uniform continuity, end, sequence tending to an end

Let X and Y be metric spaces. The following theorem is well-known.

Theorem 1. *The map $f : X \to Y$ is continuous if for every sequence (x_n) in X from*

$$\lim_{n\to\infty} x_n = x_0 \tag{1}$$

it follows

$$\lim_{n\to\infty} f(x_n) = f(x_0). \tag{2}$$

The purpose of this paper is to state a theorem of the above type for uniformly continuous maps.

Of course directly from the definition the following theorem can be stated: The map $f : X \to Y$ is uniformly continuous if for any two sequences (x_n), (y_n) in X from

$$\lim_{n\to\infty} d(x_n, y_n) = 0 \tag{3}$$

it follows

$$\lim_{n\to\infty} d(f(x_n), f(y_n)) = 0. \tag{4}$$

The idea is to eliminate some of the sequences. It is clear that if a map is continuous then for the uniform continuity only the sequences which are not contained in a compact subset are of interest. We can make a further selection, i.e. only sequences are of interest which tend to an end of X. The notion of end can be found in the papers of Freudenthal [3], [4] (also [1], [2]). The crucial point in the proof is Theorem 4, a modified version of Theorem 3 from [5].

Lemma 2. *Let X be connected, and let Q be an open nonempty subset of X, $Q \neq X$. Then $\overline{Q}\backslash Q \neq \emptyset$.*

Proof. If the statement of the lemma is not true, i.e. $\overline{Q} = Q$, then Q is a closed subset of X. Then Q is open and closed at the same time and $Q \neq X$, and it follows that X is not connected.

This is a contradiction, so the statement of the lemma must be true.

Theorem 3. *Let X be connected, locally connected, T_2 and locally compact. Let $K \subset X$, K compact. If $K \subset V$ where V is open, then V contains all components of $X\backslash K$, but finitely many.*

Proof. Let $\{Q_b \mid b \in B\}$ be the set of components of $X\backslash K$. Let $\{Q_a \mid a \in A\}$ be the set of components of $X\backslash K$ which are not completely contained in V, i.e.

$$Q_a \cap (X\backslash V) \neq \emptyset. \tag{5}$$

Since X is locally connected each Q_a is open. From Lemma 2, $\overline{Q}_a\backslash Q_a \neq \emptyset$, and it follows that

$$\overline{Q}_a \cap K \neq \emptyset. \tag{6}$$

Since X is locally compact, it follows that for every $x \in K$, there exists a neighborhood $U_x \subset V$ such that $\overline{U}_x$ is compact.

The covering of K, $\{U_x \mid x \in K\}$ contains a finite subcovering $\{U_{x_1}, U_{x_2}, \ldots, U_{x_n}\}$. If we put

$$U = U_{x_1} \cup U_{x_2} \cup \cdots \cup U_{x_n}, \tag{7}$$

then $\overline{U}$ is compact.

From $Q_a \cap K \neq \emptyset$ it follows that $Q_a \cap U \neq \emptyset$. Also $Q_a \cap U$ is open in Q_a and we can apply Lemma 2 in the space Q_a. Then we have

$$(Q_a \cap \overline{U}) \backslash (Q_a \cap U) \neq \emptyset \tag{8}$$

i.e.

$$Q_a \cap (\overline{U} \backslash U) \neq \emptyset. \tag{9}$$

Also each point of $\overline{U} \backslash U$ is in some Q_b, $b \in B$.

Since $\overline{U} \backslash U$ is compact it follows that there exists a finite covering $\{Q_{b_1}, Q_{b_2}, \ldots, Q_{b_p}\}$ of $\overline{U} \backslash U$. But for every $a \in A$,

$$Q_a \in \{Q_{b_1}, Q_{b_2}, \ldots, Q_{b_p}\} \tag{10}$$

i.e. A is finite.

Definition. *We call a subspace of X essential if it is not contained in any compact subset of X.*

Theorem 4. *Let X be T_2, connected, locally connected and locally compact. Then for every compact C there exists a compact D, $C \subset D$, such that $X \backslash D$ has only a finite number of components.*

Moreover, D can be chosen such that each of the finite number of components of $X \backslash D$ is essential.

Proof. Let C be compact in X. Since X is locally compact there exists an open set U and a compact K' such that $C \subset U \subset K'$. By the previous theorem, U contains all the components of $X \backslash C$, except a finite number.

For a given nonessential component of $X \backslash C$ not contained in U, there exists a compact set which contains this component. Because the number of these components is finite, it follows that there exists a compact set K, such that $K \subset K'$, and K contains all nonessential components of $X \backslash C$.

Let $\{Q_a \mid a \in A\}$ be all nonessential components of $X \backslash C$. We put

$$D = \bigcup_{a \in A} Q_a \bigcup C. \tag{11}$$

Then $X \backslash D$ is a union of essential components of $X \backslash C$. Since X is locally connected, every component of $X \backslash C$ is open, and it follows that $X \backslash D$ is open, i.e. D is closed. From $D \subset K$ we have that D is compact.

Moreover, $X \backslash D$ contains only a finite number of components and all components are essential.

To state the main theorem we need the definition of the notion of *end*. We denote by $S(X\backslash C)$ the set of components of connectedness of $X\backslash C$ for any compact set C in X. Then the set of ends $E(X)$ of X is:

$$E(X) = \varprojlim S(X\backslash C) \tag{12}$$

where the inverse limit is taken over all compact subsets C of X.

Definition. *A sequence of compacts (C_n) in X,*

$$C_1 \subset C_2 \subset C_3 \subset \cdots$$

is cofinal if:

$$\bigcup_{n=1}^{\infty} C_n = X \tag{13}$$

and any compact subset C in X is contained in some C_n.

Lemma 5. *In any locally compact, T_2 space with a countable basis there exists a cofinal sequence of compacts.*

Proof. Let $\mathcal{B}$ be a countable basis for X. Then for each point $x \in X$ there exists a $B_x \in \mathcal{B}$ and a compact set K_x such that $x \in B_x \subset K_x$. The set $\mathcal{A} = \{B_x \mid x \in X\}$ is also countable, i.e.

$$\mathcal{A} = \{B_x \mid x \in X\} = \{B_1, B_2, B_3, \ldots\}, \tag{14}$$

and

$$\bigcup_{n=1}^{\infty} \overline{B}_n = X. \tag{15}$$

Let $C_1 = \overline{B}_1$, $C_2 = \overline{B}_1 \cup \overline{B}_2$, $C_3 = \overline{B}_1 \cup \overline{B}_2 \cup \overline{B}_3, \ldots$. Then (C_n) is the required cofinal sequence of compacts. To prove this let C be an arbitrary compact subset of X. Then C can be covered by a finite number of sets from $\mathcal{A}$. Let n be the largest index of these sets. Then $C \subset C_n$.

For spaces X in which there exists a cofinal sequence of compacts (C_n),

$$C_1 \subset C_2 \subset C_3 \subset \cdots$$

the set of ends $E(X)$ consists of all sequences

$$Q_1 \supset Q_2 \supset Q_3 \supset \cdots$$

where $Q_i \in S(X \backslash C_i)$.

The definition does not depend on the choice of the cofinal sequence (C_n). It follows that for fixed cofinal sequence (C_n) in X each end of X is uniquely determined by a sequence $Q_1 \supset Q_2 \supset Q_3 \supset \cdots$.

Definition. *The sequence (x_k) in X has the limit at the end q determined by the sequence $Q_1 \supset Q_2 \supset Q_3 \supset \cdots$ ((x_k) tends to the end q) if for any n, there exists an index k_0 such that*

$$x_k \in Q_n, \tag{17}$$

for all $k \geq k_0$.

Theorem 6. *Let X be a separable metric space, locally compact, connected and locally connected. Let Y be a metric space and $f : X \to Y$ be a continuous map. Then f is uniformly continuous if for any sequence (x_n) in X such that (x_n) tends to some end of X, and another sequence (y_n) from*

$$\lim_{n\to\infty} d(x_n, y_n) = 0 \tag{18}$$

it follows that

$$\lim_{n\to\infty} d(f(x_n), f(y_n)) = 0. \tag{19}$$

Proof. Suppose that the theorem is not true, i.e. function f is not uniformly continuous. This means: there exists an $\varepsilon > 0$, such that for any $\delta > 0$ there exist $x, y \in X$, such that $d(x,y) < \varepsilon$, and $d(f(x), f(y)) \geq \varepsilon$.

We choose a cofinal sequence of compact sets

$$C_1 \subset C_2 \subset C_3 \subset \cdots$$

such that each $X \backslash C_n$ has only a finite number of components.

We define subset $P(X \backslash C_n)$ of the set of components of $X \backslash C_n$ in the following way: $Q_n \in P(X \backslash C_n)$ if and only if there exists a point $x_n \in Q_n$, and a point $y_n \in X$ such that:

$$d(x_n, y_n) < \frac{1}{n}, \tag{20}$$

and

$$d(f(x_n), f(y_n)) \geq \varepsilon. \tag{21}$$

Let

$$\prod_n = \{(Q_1, Q_2, \ldots, Q_n) \mid Q_i \in P(X \backslash C_i),\ i = 1, 2, \ldots, n, \tag{22}$$
$$Q_1 \supset Q_2 \supset \cdots \supset Q_n\}$$

and

$$\prod_\infty = \{(Q_1, Q_2, \ldots, Q_n, \ldots) \mid Q_i \in P(X \backslash C_i),\ i = 1, 2, \ldots, \tag{23}$$
$$Q_1 \supset Q_2 \supset \cdots \supset Q_n \supset \cdots\}$$

Since f is uniformly continuous on each compact C_n, it follows that $P(X \backslash C_n) \neq \emptyset$, $n = 1, 2, \ldots$. For any $Q_{i+1} \in P(X \backslash C_{i+1})$ there exists a component $Q_i \in P(X \backslash C_i)$ such that $Q_{i+1} \subset Q_i$. It follows that $\prod_n \neq \emptyset$ for all indices n.

We have to prove that $\prod_\infty \neq \emptyset$.

There exists $Q_1^0 \in P(X \backslash C_1)$ such that for each n there exists an n-tuple (which depends on n), $(Q_1, Q_2, \ldots, Q_n) \in \prod_n$ (if this is not true, then starting from some n_0, $\prod_\infty = \emptyset$ for all $n \geq n_0$ because $P(X \backslash C_1)$ is finite).

Suppose we have

$$(Q_1^0, Q_2^0, \ldots, Q_j^0) \in \prod_j \tag{24}$$

such that for each $n \geq j$ there exists an n-tuple (which depends on n),

$$(Q_1^0, Q_2^0, \ldots, Q_j^0, Q_{j+1}, \ldots, Q_n) \in \prod_n \tag{25}$$

Then there exists $Q_{j+1}^0 \in P(X \backslash C_{j+1})$, such that for each n there exists an n-tuple (which depends on n),

$$(Q_1^0, Q_2^0, \ldots, Q_j^0, Q_{j+1}^0, \ldots, Q_n) \in \prod_n \tag{26}$$

(if this is not true then we will have that there exists n_0, such that $\prod_n = \emptyset$, for all $n \geq n_0$).

In this way a sequence

$$(Q_1^0, Q_2^0, \ldots, Q_j^0, Q_{j+1}^0, \ldots) \in \prod_\infty, \tag{27}$$

is constructed, such that

$$Q_1^0 \supset Q_2^0 \supset \cdots \supset Q_j^0 \supset Q_{j+1}^0 \supset \cdots$$

i.e. this sequence determines an end of X.

From the construction there exists a point $x_j \in Q_j^0$, and a point $y_j \in X$ such that:

$$d(x_j, y_j) < \frac{1}{j}, \tag{28}$$

and

$$d(f(x_j), f(y_j)) \geq \varepsilon. \tag{29}$$

Further, since the sequence (x_n) tend to an end of X, we have a contradiction.

The conclusion is that the map f must be uniformly continuous.

Corollary. *Let X be a connected manifold (with or without boundary), Y a metric space and $f : X \to Y$ a continuous map. Then f is uniformly continuous if for any sequence (x_n) in X such that (x_n) tends to some end of X, and another sequence (y_n) from*

$$\lim_{n \to \infty} d(x_n, y_n) = 0 \tag{30}$$

it follows that

$$\lim_{n \to \infty} d(f(x_n), f(y_n)) = 0. \tag{31}$$

REFERENCES

[1] Ball B., Sher R., *A theory of proper shape for locally compact spaces*, Fund. Math. LXXXVI.2 (1974), 163–192.

[2] Ball B., *Proper shape retracts*, Fund. Math. LXXXIX.2 (1975), 177–189.

[3] Freudenthal H., *Über die Enden topologisher Räume und Gruppen*, Math. Z. 33 (1931), 692–713.

[4] Freudenthal H., *Kompaktisierungen und Bikompaktisierungen*, Indag. Math. 13 (1951), 184–192.

[5] Hocking J., Young G., *Topology*, Addison-Wesley Publ. Comp. Inc., 1961.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 155-166

THE CONSTRUCTION OF POTENTIAL DIFFERENTIAL - FUNCTIONAL EQUATION

Milan MARTINOVIĆ
Faculty of Natural Sciencies and Mathematics,
University of Monte Negro,
Cetinjski put bb, 81000 Podgorica, Yugoslavia

Milenko PIKULA
Faculty of Management, Trebinje, Republic of Srpska

Abstract. We consider an operator which arises when a linear functional is added to the Sturm-Liouville operator. The corresponding inverse spectral problem is studied. In particular, we prove the uniqueness theorem.

Keywords: differential operator, eigenvalues, inverse problem.

Let us give a definition of linear operator L, or more precisely $L(q,a)$, which is the object of our investigations. Its domain consists of functions $y(x)$, from $L_2(0,\pi)$, satisfying $y'' \in L_2(0,\pi)$, and $y(0) = y(\pi) = 0$. For such functions $y(x)$, we put $Ly = -y'' + q(x)y + y(a)$. Here, $q(x)$ is a complex-valued function, called a potential. We assume that the potential is a summable function. i.e. $q(x) \in L_1(0,\pi)$, if not specified otherwise. Here, $0 < a < \pi$. Let us also formulate a corresponding boundary problem with a spectral paramether $\lambda = z^2$:

$$-y''(x) + q(x)y(x) + y(a) = \lambda y(x), \tag{1}$$

$$y(0) = 0, \tag{2}$$

$$y(\pi) = 0. \tag{3}$$

In the first part of the present paper, we will investigate the asymptotic behaviour of the eigenvalues of the problem (1)-(3). In the second part, an integral equation, on the unknown $q(x)$, is derived. Using this equation, the main result of the present paper is obtained. We prove that one sequence of eigenvalues determines both potential q and parameter a, under certain restrictions, in the case of symmetric potential, $q(\pi - x) = q(x)$. Various questions of spectral theory, for the operator L and for similar operators, were studied in earlier papers. In paper [1], an expansion problem is solved. In our paper [2], a trace formula is obtained.

Theorem 1. *If the function $q(x)$ has a bounded variation derivation, then the following formula, for eigenvalues λ_n $(n = 1, 2, \ldots)$ of the problem (1)-(3), holds:*

$$\lambda_n = n^2 + C_1 + \frac{1}{n} \cdot C_2(n) + O(n^{-2}), \tag{4}$$

where

$$C_1 = \frac{1}{\pi} \int_0^\pi q(t)dt, \quad C_2(n) = \frac{2}{\pi}[1 - (-1)^n] \sin an.$$

Proof. Equation (1) with condition (2) is equivalent to the integral equation

$$y(x, z) = c_2 \sin zx + y(a) \cdot \frac{1 - \cos zx}{z^2} + \frac{1}{z} \int_0^x q(t) \sin z(x - t) y(t, z) dt, \tag{5}$$

where c_2 denotes an arbitrary constant. Equation (5) can be solved by the method of successive approximations. We will introduce the necessary notations:

$$T_l = (t_1, t_2, \ldots, t_l), \quad l \geq 1, \quad Q(T_l) = \prod_{k=1}^{l} q(t_k), \quad dT_l = dt_1 dt_2 \ldots dt_l,$$

$$D_l = \{T_l \mid 0 \leq t_1 \leq \pi, 0 \leq t_2 \leq t_1, \ldots, 0 \leq t_l \leq t_{l-1}\},$$

$$D_l^1 = \{T_l | 0 \leq t_1 \leq a, 0 \leq t_2 \leq t_1, \ldots, 0 \leq t_l \leq t_{l-1}\},$$

$$S(z, T_l) = \sin z(\pi - t_1) \sin zt_l \prod_{k=1}^{l-1} \sin z(t_k - t_{k+1}),$$

$$S^1(z, T_l) = \sin z(a - t_1) \sin zt_l \prod_{k=1}^{l-1} \sin z(t_k - t_{k+1}),$$

$$C(z,T_l) = \sin z(\pi - t_1)[1 - \cos zt_l] \prod_{k=1}^{l-1} \sin z(t_k - t_{k+1}),$$

$$C^1(z,T_l) = \sin z(a - t_1)[1 - \cos zt_l] \prod_{k=1}^{l-1} \sin z(t_k - t_{k+1}),$$

$$\begin{aligned}\varphi(\pi,z) = \sin \pi z &+ \frac{1}{z}\int_0^{\pi} q(t_1)\sin z(\pi - t_1)\sin zt_1 dt_1 \\ &+ \sum_{l=2}^{\infty} \frac{1}{z^l}\int_{D_l} Q(T_l)S(z,T_l)dT_l,\end{aligned}$$

$$\psi(\pi,z) = \frac{1-\cos\pi z}{z^2} + \sum_{l=1}^{\infty}\frac{1}{z^{l+2}}\int_{D_l} Q(T_l)C(z,T_l)dT_l.$$

Now, let $x = \pi$ in equation (5), and let the boundary condition (3) be used. So, the characteristic equation for L, i.e. an equation having eigenvalues (their square roots) as solutions, is derived:

$$\Phi(z,a) = \varphi(\pi,z)[1 - \psi(a,z)] + \varphi(a,z)\psi(\pi,z) = 0. \tag{6}$$

To determine the asymptotic behaviour of the zeros of $\Phi(z,a)$, Horn's iterative method can be applied. So, if we write

$$z_n = n + \frac{1}{n}c_1(n) + \frac{1}{n^2}c_2(n) + O(n^{-3}),$$

then using (6), we find,

$$\begin{aligned} c_1(n) &= \frac{1}{2\pi}\int_0^{\pi} q(t)dt - \frac{1}{2\pi}\int_0^{\pi} q(t)\cos 2ntdt, \\ c_2(n) &= \frac{1-(-1)^n}{\pi}\sin an + \frac{C_1}{2\pi}\int_0^{\pi}(\pi - 2t)q(t)\sin 2ntdt. \end{aligned}$$

The function $q'(x)$ has a bounded variation, therefore one integration by parts is allowed, and we find

$$\begin{aligned} c_1(n) &= \frac{1}{2\pi}\int_0^{\pi} q(t)dt + O(n^{-2}), \\ c_2(n) &= \frac{1-(-1)^n}{\pi}\sin an + O(n^{-1}). \end{aligned}$$

Finally,

$$z_n = n + \frac{1}{2\pi n}\int_0^\pi q(t)dt + \frac{1}{n^2}\cdot\frac{1-(-1)^n}{\pi}\sin an + O(n^{-3}).$$

All that remains is to raise to the square, i.e. to compute $\lambda_n = z_n^2$ $(n \in N)$. Formula (4) is derived. The proof is completed.

With the help of the previous theorem, part of the inverse spectral problem may be solved right away. Let the sequence $\{\lambda_n\}_{n=1}^\infty$, representing the spectrum of some operator L from the earlier-described class, be given. Can the number a be determined (according to this sequence)? Let us form a sequence:

$$\beta_n(a) = \frac{\lambda_{2n+3} - (2n+3)^2 - \lambda_{2n-3} + (2n-3)^2}{\lambda_{2n+1} - (2n+1)^2 - \lambda_{2n-1} + (2n-1)^2}.$$

From (4), it is easy to conclude:

$$\lim_{n\to\infty}\beta_n(a) = \frac{\sin 3a}{\sin a}. \tag{7}$$

Thus, when the spectrum is known, the expression on the right side of (7) is known, i.e. a is found.

We want to form an equation on unknown $q(x)$. Let us consider some operator $L(q,a)$ defined by (1)-(3). Let $\{\lambda_n\}_{n=1}^\infty$ be the sequence of its eigenvalues. The characteristic function $\Phi(z)$, from (6), of this operator, can be presented in two manners:

$$\Phi(z) = A\cdot z\cdot\prod_{n=1}^\infty\left(1-\frac{z^2}{\lambda_n}\right),\quad A = \pi\prod_{n=1}^\infty\frac{n^2}{\lambda_n}, \tag{8}$$

$$\Phi(z) = \varphi(\pi,z,q)[1-\psi(a,z,q)] + \varphi(a,z,q)\psi(\pi,z,q). \tag{9}$$

The function $\varphi(x,z,q)$ is defined by:

$$\varphi(x,z,q) = \sin xz + \frac{1}{z}\int_0^x q(t_1)\sin z(x-t_1)\sin zt_1 dt_1 +$$

$$\frac{1}{z^2}\int_0^x\int_0^{t_1} q(t_1)q(t_2)\sin z(x-t_1)\sin zt_2 \sin z(t_1-t_2)dt_2dt_1 + \cdots,$$

and analogously for $\psi(x,z,q)$.

Since the two representations (8) and (9) hold for every $z \in C$, we set $z = m$ (m - natural), and thus we get the sequence of equations:

$$\begin{aligned} &\pi \prod_{n=1}^{\infty} \frac{n^2}{\lambda_n} \cdot m \cdot \prod_{n=1}^{\infty} (1 - \frac{m^2}{\lambda_n}) \\ (10) \qquad &= \varphi(\pi, m, q)[1 - \psi(a, m, q)] + \varphi(a, m, q)\psi(\pi, m, q). \end{aligned}$$

We wish to establish a connection between the λ_n (given objects) and Fourier's coefficients (unknown) of the potential q. Let

$$a_{2m} = \int_0^{\pi} q(t) \cos 2mt dt.$$

Let us introduce the new notations:

$$\begin{aligned} \Delta_2\varphi(\pi, m, q) &= \sum_{l=2}^{\infty} \frac{1}{m^l} \int_{D_l} Q(T_l) S(m, T_l) dT_l, \\ \Delta_1\varphi(a, m, q) &= \sum_{l=1}^{\infty} \frac{1}{m^l} \int_{D_l^1} Q(T_l) S^1(m, T_l) dT_l = \\ &= \frac{1}{m} \int_0^a q(t) \sin m(a - t) \sin mt dt + \Delta_2\varphi(a, m, q), \\ \Delta_1\psi(\pi, m, q) &= \frac{(-1)^{m+1}}{m^3} \int_0^{\pi} q(t) \sin mt[1 - \cos mt] dt + \\ &+ \sum_{l=2}^{\infty} \frac{1}{m^{l+2}} \int_{D_l} Q(T_l) \cdot C(m, T_l) dT_l = \\ &= \frac{(-1)^{m+1}}{m^3} \int_0^{\pi} q(t) \sin mt[1 - \cos mt] dt + \Delta_2\psi(\pi, m, q), \\ \Delta_1\psi(a, m, q) &= \frac{1}{m^3} \int_0^a q(t) \sin m(a - t)[1 - \cos mt] dt + \Delta_2\psi(a, m, q). \end{aligned}$$

Hence,

$$\begin{aligned} \varphi(\pi, m, q) &= \frac{(-1)^m}{2m} a_{2m} + \Delta_2\varphi(\pi, m, q), \\ \psi(a, m, q) &= \frac{1 - \cos am}{m^2} + \Delta_1\psi(a, m, q), \\ \varphi(a, m, q) &= \sin am + \Delta_1\varphi(a, m, q), \\ \psi(\pi, m, q) &= \frac{1 - (-1)^m}{m^2} + \Delta_1\psi(\pi, m, q). \end{aligned}$$

Consequently,

$$\varphi(\pi,m,q)[1-\psi(a,m,q)]=\frac{(-1)^m}{2m}[1-\frac{1-\cos am}{m^2}]a_{2m}+$$

$$(11)\qquad [1-\frac{1-\cos am}{m^2}]\Delta_2\varphi(\pi,m,q)-\varphi(\pi,m,q)\Delta_1\psi(a,m,q),$$

and

$$\varphi(a,m,q)\psi(\pi,m,q)=\frac{1-(-1)^m}{m^2}\sin am+\frac{(-1)^{m+1}}{m^3}\sin am\cdot$$

$$(12)\qquad [\int_0^\pi q(t)\sin mtdt-\frac{1}{2}\int_0^\pi q(t)\sin 2mtdt]+\frac{1-(-1)^m}{m^3}\cdot$$

$$\int_0^a q(t)\sin m(a-t)\sin mtdt+\varphi(a,m,q)\Delta_2\psi(\pi,m,q)+\Delta_2\varphi(a,m,q)\psi(\pi,m,q).$$

The equations (10) take a new form, after the relations (11) and (12) are used (for $m\in N$):

$$\Phi(m)-\frac{1-(-1)^m}{m^2}\sin am=\frac{(-1)^m}{2m}[1-\frac{1-\cos am}{m^2}]a_{2m}+$$

$$+\frac{(-1)^{m+1}}{m^3}\sin am\cdot[\int_0^\pi q(t)\sin mtdt-\frac{1}{2}\int_0^\pi q(t)\sin 2mtdt]+$$

$$(13)\qquad +\frac{1-(-1)^m}{m^3}\int_0^a q(t)\sin m(a-t)\sin mtdt+$$

$$+[1-\frac{1-\cos am}{m^2}]\Delta_2\varphi(\pi,m,q)-\varphi(\pi,m,q)\Delta_1\psi(a,m,q)+$$

$$+\varphi(a,m,q)\Delta_2\psi(\pi,m,q)+\Delta_2\varphi(a,m,q)\psi(\pi,m,q).$$

At this stage, we have to put some limitations on the potential. We assume that

$$(14)\qquad \int_0^\pi q(t)dt=0.$$

Also, let

$$(15)\qquad q(\pi-x)=q(x).$$

These two conditions do not restrict the generality of our analysis. The last condition implies $b_{2m}=\int_0^\pi q(t)\sin 2mtdt=0$, i.e. the potential q is determined by its Fourier coefficients a_{2m} (because $0 \le x \le \pi$). Let us look at the relation (13) when the number m is even. Then, on the right side, the second and the third member vanish, if the additional supposition is made - that q is already determined by its coefficients of the form a_{4m}. Therefore, let also

$$q(\frac{\pi}{2} - x) = q(x). \tag{16}$$

In view of the previous suppositions, (13) can be rewritten as

$$\begin{aligned}\Phi(2m) = \frac{1}{4m}\Big[1 - \frac{1 - \cos 2am}{(2m)^2}\Big]a_{4m} + \Big[1 - \frac{1 - \cos 2am}{(2m)^2}\Big]\Delta_2\varphi(\pi, 2m, q)\\ - \varphi(\pi, 2m, q)\Delta_1\psi(a, 2m, q) + \varphi(a, 2m, q)\Delta_2\psi(\pi, 2m, q) +\\ + \Delta_2\varphi(a, 2m, q)\psi(\pi, 2m, q).\end{aligned} \tag{17}$$

Set

$$\xi_m = \frac{1}{4m}\Big[1 - \frac{1 - \cos 2am}{(2m)^2}\Big],$$

and besides $\xi_m \neq 0$ for every $m \in N$. Let

$$A_{4m} = \frac{1}{\xi_m}\Phi(2m).$$

We also put

$$\begin{aligned}\alpha_{4m}(q, a) = -4m\Delta_2\varphi(\pi, 2m, q) +\\ + \frac{1}{\xi_m}\big[\varphi(\pi, 2m, q)\Delta_1\psi(a, 2m, q) - \varphi(a, 2m, q)\Delta_2\psi(\pi, 2m, q) -\\ - \Delta_2\varphi(a, 2m, q)\psi(\pi, 2m, q)\big].\end{aligned} \tag{18}$$

Now, (17) receives the form

$$a_{4m} = A_{4m}(\lambda_n, a) + \alpha_{4m}(q, a), \quad m \in N. \tag{19}$$

It is easy to show that $A_{4m} \to 0$ when $m \to \infty$, and that the series

$$\sum_{m=1}^{\infty} A_{4m} \cos 4mx$$

converges. Thereby, the numbers A_{4m} are Fourier's coefficients of some function from $L_1(0, \pi)$. Let $L'_1(0, \pi)$ denote the subspace of $L_1(0, \pi)$ consisting of all elements satisfying the conditions (14)-(16). By means of previous examinations, the following result is already proved.

Theorem 2. *Let $q \in L'_1(0,\pi)$. Let λ_n $(n = 1, 2, \ldots)$ be the eigenvalues of the operator $L(q,a)$; for $L(q,a)$, see formulas (1)-(3). Then, the system of equations (19) holds, expressing a connection between the Fourier's coefficients a_{4m} of this potential q and these eigenvalues λ_n.*

Set

$$f(x) = \frac{2}{\pi} \sum_{m=1}^{\infty} A_{4m} \cos 4mx. \tag{20}$$

We shall multiply every equation of system (19) by $\frac{2}{\pi} \cos 4mx$, and, after that, we shall sum up. As a result,

$$q(x) = f(x) + \frac{2}{\pi} \sum_{m=1}^{\infty} \alpha_{4m}(q,a) \cos 4mx. \tag{21}$$

After the substitution $h(x) = q(x) - f(x)$ is carried out, (21) transforms into

$$h(x) = \frac{2}{\pi} \sum_{m=1}^{\infty} \alpha_{4m}(f+h,a) \cos 4mx. \tag{22}$$

For equation (22), the uniqueness of the solution will be studied, whereby we rewrite (22) as:

$$h = T(h), \tag{23}$$

where, clearly, $T(h)$ is a nonlinear operator acting by the rule:

$$T(h) = \frac{2}{\pi} \sum_{m=1}^{\infty} \alpha_{4m}(f+h,a) \cos 4mx.$$

We shall demonstrate that Banach's theorem about a fixed-point can be applied to the operator T and to some ball with centre f. We assume that T acts in the space $L_1(0,\pi)$, i.e. we use the norm $\|h\| = \int_0^\pi |h(t)|dt$. With a certain liberty, we shall write $F(T_l) + H(T_l)$ for $\prod_{k=1}^{l} (f(t_k) + h(t_k))$.

In accordance with the right side of formula (18), the expression for $T(h)$ can be divided into four parts $T_0(h) + T_1(h) + T_2(h) + T_3(h)$, where:

$$T_0(h)(x) = -\sum_{m=1}^{\infty} 4m \cdot \cos 4mx \cdot \sum_{l=2}^{\infty} \frac{1}{(2m)^l} \int_{D_l} [F(T_l) + H(T_l)] S(2m, T_l) dT_l, \tag{24}$$

$$\begin{aligned} T_1(h)(x) = \sum_{m=1}^{\infty} \frac{1}{\xi_m} \cdot \cos 4mx \cdot \sum_{l_1=1}^{\infty} \frac{1}{(2m)^{l_1}} \int_{D_{l_1}} [F(T_{l_1}) + H(T_{l_1})] S(2m, T_{l_1}) dT_{l_1} \cdot \\ \sum_{l_2=1}^{\infty} \frac{1}{(2m)^{l_2+2}} \int_{D^1_{l_2}} [F(T_{l_2}) + H(T_{l_2})] C^1(2m, T_{l_2}) dT_{l_2}, \end{aligned} \tag{25}$$

and for $T_2(h)$, $T_3(h)$ by analogy.

At this stage, we wish to make an appropriate estimate for $T_0(h_1) - T_0(h_2)$. If we introduce the notation

$$\sigma_0 = \sup_{l, T_l} \sum_{m=1}^{\infty} \left| \frac{1}{m(2m)^{l-1}} S(2m, T_l) \right| \int_0^{\pi} |\cos 2mx| dx, \tag{26}$$

then we have

$$|T_0(h_1)(x) - T_0(h_2)(x)| \leq \sum_{l=2}^{\infty} \int_{T_l} |[F(T_l) + H_1(T_l)] - [F(T_l) + H_2(T_l)]| dT_l \cdot$$

$$\sup_{l, T_l} \left| \sum_{m=1}^{\infty} \frac{1}{m(2m)^{l-1}} S(2m, T_l) \right| \cdot |\cos 2mx|,$$

and

$$\begin{aligned} \|T_0(h_1) - T_0(h_2)\| \leq \sigma_0 \sum_{l=2}^{\infty} \int_0^{\pi} |[f(t_1) + h_1(t_1)] \int_0^{t_1} [f(t_2) + h_1(t_2)] \cdots \\ \int_0^{t_{l-1}} [f(t_l) + h_1(t_l)] dt_l \cdots dt_2 - [f(t_1) + h_2(t_1)] \int_0^{t_1} [f(t_2) + h_2(t_2)] \cdots \\ \cdots \int_0^{t_{l-1}} [f(t_l) + h_2(t_l)]\, dt_l \ldots dt_2 \Big| dt_1, \\ \|T_0(h_1) - T_0(h_2)\| \leq \sigma_0 \sum_{l=1}^{\infty} \frac{1}{l!} (\|f\| + \rho)^l \cdot \|h_1 - h_2\| = \\ = \sigma_0 (e^{\|f\| + \rho} - 1) \|h_1 - h_2\|, \end{aligned} \tag{27}$$

where $\rho = \max\{\|h_1\|, \|h_2\|\}$.

In order to estimate T_1, we write

$$\sigma_1 = \sup \sum_{m=1}^{\infty} \left| \frac{1}{\xi_m} \cdot \frac{1}{(2m)^{l_1}} S(2m, T_{l_1}) \cdot \frac{1}{(2m)^{l_2+2}} C^1(2m, T_{l_2}) \right| \cdot \int_0^{\pi} |\cos 2mx| dx,$$

where "sup" is taken over $T_{l_1} \in D_{l_1}$, $T_{l_2} \in D^1_{l_2}$, $l_1 \geq 1$, $l_2 \geq 1$. The following inequality is valid:

$$\|T_1(h_1) - T_1(h_2)\| \leq \sigma_1 \sum_{l_1=1}^{\infty} \sum_{l_2=1}^{\infty} \int_0^{\pi} |[f(t_1) + h_1(t_1)] \cdot$$

$$\cdot \int_0^{t_1} \cdots \int_0^{t_{l_1-1}} [f(t_{l_1}) + h_1(t_{l_1})] dt_{l_1} \cdots dt_2 \cdot$$

$$\cdot \int_0^{a} [f(t_1^1) + h_1(t_1^1)] \int_0^{t_1^1} \cdots \int_0^{t^1_{l_2-1}} [f(t^1_{l_2} + h_1(t^1_{l_2})] dt^1_{l_2} \cdots dt^1_1 -$$

$$-[f(t_1) + h_2(t_1)] \int_0^{t_1} \cdots \int_0^{t_{l_1-1}} [f(t_{l_1}) + h_2(t_{l_1})] dt_{l_1} \cdots dt_2 \cdot$$

$$\cdot \int_0^{a} [f(t_1^1) + h_2(t_1^1)] \int_0^{t_1^1} \cdots \int_0^{t^1_{l_2-1}} [f(t^1_{l_2}) + h_2(t^1_{l_2})] dt^1_{l_2} \ldots dt^1_1 | dt_1.$$

In view of the fact that the previous double series absolutely converges, an arbitrary amalgamation in this series is legal. Let the k-th member consist of all integrals having a dimension k. So,

$$\|T_1(h_1) - T_1(h_2)\| \leq \sigma_1 \cdot \sum_{k=2}^{\infty} \gamma_k,$$

$$\gamma_2 = \int_0^{\pi} |[f(t_1) + h_1(t_1)] \int_0^{a} [f(t_1^1) + h_1(t_1^1)] dt_1^1 -$$

$$- [f(t_1) + h_2(t_1)] \int_0^{a} [f(t_1^1) + h_2(t_1^1)] dt_1^1 | dt_1,$$

and for γ_3, etc. by analogy.

Can the numbers surpassing γ_k be found? We have

$$\gamma_2 \leq \|h_1 - h_2\|(\|f\| + \rho), \quad \rho \geq \max\{\|h_1\|, \|h_2\|\}.$$

Analogously,

$$\gamma_3 \leq \int_0^{\pi} |[f(t_1) + h_1(t_1)] \int_0^{t_1} [f(t_2) + h_1(t_2)] dt_2 \int_0^{a} [f(t_1^1) + h_1(t_1^1)] dt_1^1 -$$

$$-[f(t_1)+h_2(t_1)]\int_0^{t_1}[f(t_2)+h_2(t_2)]dt_2\int_0^a[f(t_1^1)+h_2(t_1^1)]dt_1^1\Big|dt_1+$$

$$+\int_0^\pi\Big|[f(t_1)+h_1(t_1)]\int_0^a[f(t_1^1)+h_1(t_1^1)]\int_0^{t_1^1}[f(t_2^1)+h_1(t_2^1)]dt_2^1dt_1^1-$$

$$-[f(t_1)+h_2(t_1)]\int_0^a[f(t_1^1)+h_2(t_1^1)]\int_0^{t_1^1}[f(t_2^1)+h_2(t_2^1)]dt_2^1dt_1^1\Big|dt_1,$$

then,

$$\gamma_3\le\big(\frac{3}{2\cdot 2!}(\|f\|+\rho)^2+\frac{1}{8}(\|f\|+\rho)^2\big)\|h_1-h_2\|=\frac{7}{8}(\|f\|+\rho)^2\|h_1-h_2\|.$$

In the same manner,

$$\begin{aligned}\gamma_4\le\big(&\frac{4}{2\cdot 3!}(\|f\|+\rho)^3+\frac{3}{2\cdot 2^3}(\|f\|+\rho)^3+\\&+\frac{1}{2^4}(\|f\|+\rho)^3\big)\|h_1-h_2\|=\frac{7}{12}(\|f\|+\rho)^3\|h_1-h_2\|,\end{aligned}$$

and, in general,

$$\gamma_k\le(\|f\|+\rho)^k\|h_1-h_2\|.$$

Hence,

$$\|T_1(h_1)-T_1(h_2)\|<$$

(28)

$$\sigma_1\sum_{k=1}^{\infty}(\|f\|+\rho)^k\|h_1-h_2\|=\sigma_1\frac{\|f\|+\rho}{1-\|f\|-\rho}\|h_1-h_2\|,\ \ \|f\|+\rho<1.$$

The next two estimates are obtained in a similar way:

$$\|T_2(h_1)-T_2(h_2)\|\le\sigma_2\big[\frac{\|f\|+\rho}{1-\|f\|-\rho}+e^{\|f\|+\rho}-1\big]\|h_1-h_2\|, \tag{29}$$

$$\|T_3(h_1)-T_3(h_2)\|\le\sigma_3\frac{\|f\|+\rho}{1-\|f\|-\rho}\|h_1-h_2\|. \tag{30}$$

Putting together the inequalities (27)-(30), we have:

$$\|T(h_1)-T(h_2)\|\le \tag{31}$$

$$\big\{(\sigma_0+\sigma_3)(e^{\|f\|+\rho}-1)+(\sigma_1+\sigma_2+\sigma_3)\frac{\|f\|+\rho}{1-\|f\|-\rho}\big\}\|h_1-h_2\|=\alpha\|h_1-h_2\|.$$

At this point, it is easy to determine radius r of the ball having its centre in point f, such that the transformation T is a contraction in this ball $\{\|h\| \leq r\}$, i.e. $\{\|q-f\| \leq r\}$. Preliminarily, it must be

$$\|f\| + r < 1 \quad (\|f\| = \int_0^\pi |f(t)|dt). \tag{32}$$

The quantity r must also satisfy the inequality

$$(\sigma_0 + \sigma_3)(e^{\|f\|+r} - 1) + (\sigma_1 + \sigma_2 + \sigma_3)\frac{\|f\| + r}{1 - \|f\| - r} < 1. \tag{33}$$

We would remark that $\sigma_j \geq 0$ depends only on the number a, while f depends on $\{\lambda_n\}$ and a. It only remains to demonstrate that T transforms the ball $B(f,r) = \{p, \|p-f\| \leq r\}$ into itself. That can be easily shown, performing estimates of the type (27)-(30). The next theorem represents the main result of the present paper. Its proof is contained in previous investigations.

Theorem 3. *Let the sequence of numbers $\{\lambda_n\}_{n=1}^\infty$ be given; and let these numbers be the eigenvalues of some problem having the form (1)-(3) with the potential from the set $L'_1(0,\pi)$; for $L'_1(0,\pi)$, see formulas (14)-(16). Let the function $f(x)$ be defined by formulas (8), (20). Let us suppose that there exists a number $r > 0$, such that the inequalities (32), (33) are satisfied. Then, in the ball $B(f,r) = \{p, \|p-f\| \leq r\}$ there exists the unique function $q \in L'_1(0,\pi)$, such that $L(q,a)$ has exactly the given numbers λ_n as eigenvalues.*

Remarks: The unknown potential $q(x)$ can be reconstructed by solving equation (21). In further work, we shall attempt to "magnify" the radius r by means of more precise estimates for the multiple integrals. The presented method can be also applied (for proving uniqueness) when $y(a)$, from (1), is replaced by a more general linear functional. The method can be used for various non-self-adjoint boundary problems.

REFERENCES

[1] Krall A.M., *Differential-boundary operators*, Transactions of AMS, 154(1971) 429-458.

[2] Sadovnichii V.A., Lyubishkin V.A., Martinovich M., *Finite-dimensional perturbations of discrete operators and trace formulas*, Soviet Math. Dokl., 35(1987) 417-419.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 167-170

DOUBLE OPERATOR INTEGRALS AND GENERALIZED NORMAL DERIVATIONS

Danko R. JOCIĆ
Faculty of Mathematics, University of Belgrade
Studentski trg 16, P.O. Box 550, 11000 Belgrade, Yugoslavia

Let $B(H)$ and C_2 denote respectively the space of all bounded and Hilbert-Schmidt operators acting on a separable, infinite-dimensional, complex Hilbert space H. If E and F are the spectral measures on $\mathbf{C}$, then for all X and Y in C_2, the function $\gamma \times \delta \longrightarrow trace(E(\gamma)XF(\delta)Y^*)$ defined on Borel rectangles in $\mathbf{C}^2$ can be extended to an σ-additive complex Borel measure $\mu_{X,Y}$ on $\mathbf{C}^2$ satisfying $|\mu_{X,Y}| \leq \|X\|_2\|Y\|_2$ as shown in [4]. This is a basic point for the existence of essentially bounded functional calculus in two (generally non-commuting) normal operator variables which was first introduced by Birman and Solomyak in [1] and [2] as double operator integrals.

Thus to every essentially bounded function in two complex variables there corresponds a linear transformation (**transformator**) acting on C_2 as its domain. As shown earlier by several authors, a global extension of this domain to the whole $B(H)$ is not possible (see [3] for example). However, we show in this paper how double operator integrals can be efficiently used in some problems for generalized normal derivations, despite of the fact that they represent $B(H)$ transformations. So, we start with the following theorem from [7].

Theorem 1. *For normal A and B and an arbitrary X in $B(H)$ the function*

$$f(s) = \| \, |A|^{s-1}AX|B|^{p-s} + |A|^{p-s}XB|B|^{s-1}\|_2$$

is convex and symmetric on $[0, p]$, non-increasing on $[0, p/2]$ and non-decreasing on $[p/2, p]$.

Here we present a short proof of this theorem based on the following integral representation formula:

$$\| \, |A|^{s-1}AX|B|^{p-s} + |A|^{p-1}XB|B|^{s-1}\|_2^2 = \tag{1}$$

$$\int\limits_{\sigma(A)}\int\limits_{\sigma(B)} \left| \frac{z|z|^{s-1}|w|^{p-s} + |z|^{p-s}w|w|^{s-1}}{z|z|^{t-1}|w|^{p-t} + |z|^{p-t}w|w|^{t-1}} \right|^2 d\mu_{|A|^{t-1}AX|B|^{p-t}+|A|^{p-t}XB|B|^{t-1}}(z,w)$$

for all $\frac{p}{2} \leq s \leq t \leq p$.

It was shown in [4] that the classical Fuglede-Putnam theorem and the main result in [9] immediately follow from the following integral representation formula

$$\|f(A)X - Xf(B)\|_2^2 = \iint\limits_{\sigma(A)\times\sigma(B)} \left| \frac{f(z) - f(w)}{z - w} \right|^2 d\mu_{AX-XB}(z,w) \tag{2}$$

whenever $AX - XB$ is in C_2 and f is a Lipshitz function on $\sigma(A) \cup \sigma(B)$.

Formula (1) itself can easily be shown for Hilbert-Schmidt class operators X, according to the essentially bounded calculus (see also [4] for details). To show its validity for all bounded X, we follow the lines of the proof of formula (2) in [4] and we first prove it for off-diagonal pieces E_iXF_j of X for all $|i-j| > 1$. The formula is completed by the limit argument based on the uniform boundedness principle. As $\left| \frac{z|z|^{s-1}|w|^{p-s} + |z|^{p-s}w|w|^{s-1}}{z|z|^{t-1}|w|^{p-t} + |z|^{p-t}w|w|^{t-1}} \right| \leq 1$ for all complex z, w and $\frac{p}{2} \leq s \leq t \leq p$, then according to the fact that

$$\mu_{|A|^{t-1}AX|B|^{p-t}+|A|^{p-t}XB|B|^{t-1}} = \| \, |A|^{t-1}AX|B|^{p-t} + |A|^{p-t}XB|B|^{t-1}\|_2^2,$$

the conclusion follows from the above integral representation formula.

We conclude with an application of Jensen's convexity theorem to double operator integrals.

Theorem 2. *Let A, B and X be in $B(H)$ with A and B being self-adjoint and X being in C_2. Then for every even natural number n we have*

$$\| \, |AX + XB|^n\|_2 \leq 2^{n-1}\|X\|_2^{n-1}\|A^nX + XB^n\|_2. \tag{3}$$

Proof. As shown in representation formula (2)

$$\|A^nX + XB^n\|_2^2 = \iint |x^n + y^n|^2 d\mu_X(x,y) \tag{4}$$

for some measure μ satisfying $\mu_X(\mathbf{C}^2) = \|X\|_2^2$. Having the function $x \longrightarrow x^n$ convex on $\mathbf{R}$ for n even, we get

$$|x^n + y^n| \geq 2^{n-1}|x+y|^n \tag{5}$$

and, in view of (4)

$$\|A^n X + XB^n\|_2^2 \geq 2^{2-2n}\|X\|_2^2 \int\int |x+y|^{2n} d\mu_X^0(x,y) \tag{6}$$

The measure $\mu_X^0 = \mu_X / \|X\|_2^2$ is a probability measure; now the application of Jensen's inequality gives

$$\begin{aligned} \|A^n X + XB^n\|_2^2 &\geq 2^{2-2n}\|X\|_2^{2-2n} \left(\int\int |x+y|^2 d\mu_X^0(x,y) \right)^n \\ &= 2^{2-2n}\|X\|_2^{2-2n}\|AX + XB\|_2^{2n}. \end{aligned} \tag{7}$$

For compact operator $AX - XB$ we have that

$$\|AX + XB\|_2^{2n} \geq \|AX - XB\|_{2n}^{2n} = \|\, |AX + XB|^n \|_2^2 \tag{8}$$

and now, the assertion of the theorem follows directly from (7) and (8).

Remark 1. This theorem is also true for all natural n because (5) is true for all add natural numbers n; the proof requires mathematical induction only.

For other applications of double operator integrals see [6] and [8].

REFERENCES

[1] Birman M.Š., Solomyak M.Z., *Double operator Stiltjes integrals*, Dokladi Akademii nauk SSSR, (6)T165 (1965), 1223–1226.

[2] Birman M.Š., Solomyak M.Z., *Double operator Stiltjes integrals*, Problemy matematicheskoi fiziki, vyp 1, Izd. LGU, 33–76, 1966.

[3] Farfarovskaya Yu.B., *Example of Lipschitz function of self-adjoints operators that gives a nonnuclear increment under a nuclear perturbation*, J. Soviet Math. 4 (1975), 426–433.

[4] Jocić D.R., *Kvazifunkcionalni račun i teoreme Fuglede-Putnam-Rosenblum-ovog tipa*, Doktorska disertacija, Matematički fakultet, Beograd 1991.

[5] Jocić D.R., *A Hilbert-Schmidt norm equality associated with the Fuglede-Putnam-Rosenblum type theorem for generalized multipliers*, J. Operator Theory 30 (1993), 31–40.

[6] Jocić D.R., *Norm inequalities for self-adjoint derivations*, J. Functional Anal. (to appear).

[7] Jocić D.R., *Norm inequalities for self-adjoint derivations II*, Topics in Operator Theory, operator Algebras and Applications, IMAR, Bucharest 1995.

[8] Jocić D.R., Kittaneh F., *Some pertubation inequalities for self-adjoint operators*, J. Operator Theory 31 (1994), 3–10.

[9] Kittaneh F., *On Lipshitz functions of normal operators*, Proc. Amer. Math. Soc 94 (1985), 416–418.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 171-178

ON A CLASS OF p-VALENT MEROMORPHIC FUNCTIONS WITH POSITIVE COEFFICIENTS

Santosh B. JOSHI

Department of Mathematics, Walchand College of Engineering
Sangli 416415, India

Abstract. In the present paper we introduce a class $Q_p^*\ (A, B, \alpha)$ of p-valent meromorphic functions with positive coefficients and obtain coefficient inequality, distortion theorem, radius of convexity and convex linear combination. Various results obtained in this paper are shown to be sharp.

Keywords: distortion, meromorphic functions

AMS Subject Classification 32A20, 30C45

1. INTRODUCTION

Let Q_p denote a class of functions of the form

$$f(z) = \frac{1}{z^p} + \sum_{n=0}^{\infty} a_{p+n} z^{p+n} \quad (a_{p+n} \geq 0) \tag{1}$$

which are analytic and p-valent in the punctured disc $U^* = \{z \mid 0 < |z| < 1\}$. For A, B fixed we define a class $Q_p^*\ (A, B, \alpha)$ as follows.

Definition. *A function $f \in Q_p$ is in $Q_p^*\ (A, B, \alpha)$ if $zf'(z)/f(z)$ has the form*

$$\frac{zf'(z)}{f(z)} = -\frac{p + [pB + (A - B)(p - \alpha)]w(z)}{1 + Bw(z)} \tag{2}$$

Here $w(z)$ is regular in $U = \{z \mid |z| < 1\}$ and satisfies $w(0) = 0$ and $|w(z)| < 1$ for $z \in U$; $0 \leq \alpha < p, -1 \leq A < B \leq 1,\ 0 < B \leq 1$.

The condition (2) is equivalent to

$$\left|\frac{\dfrac{zf'(z)}{f(z)}+1}{\dfrac{Bzf'(z)}{f(z)}+[pB+(A-B)(p-\alpha)]}\right|<1. \tag{3}$$

The study of the aforementioned class $Q_p^*(A,B,\alpha)$ was especially motivated by the recent work of Uralegaddi and Ganigi [6] and similar work has been done by many researchers, Aouf [1], Cho et al [2], Mogra-Reddy and Juneja [5], Clunie [3], Miller [4].

In the present paper we have obtained coefficient inequality, distortion theorem, radius of convexity and closure theorem. Various results are shown to be sharp.

2. COEFFICIENT INEQUALITIES

The following theorem gives a necessary and sufficient condition for a function to be in $Q_p^*(A,B,\alpha)$.

Theorem 1. *A function* $f(z)=\dfrac{1}{z^p}+\displaystyle\sum_{n=0}^{\infty}a_{p+n}z^{p+n}$ $(a_{p+n}\geq 0)$ *is in* $Q_p^*(A,B,\alpha)$ *if and only if*

$$\sum_{n=0}^{\infty}\{(2+A+B)p+n(1+B)+(B-A)\alpha\}a_{p+n}\leq(B-A)(p-\alpha).$$

The result is sharp.

Proof. Let $|z|=1$. Then from (3) we have

$$\left|\frac{zf'(z)}{f(z)}+p\right|-\left|\frac{Bzf'(z)}{f(z)}+[pB+(A-B)(p-\alpha)]\right|=\left|\sum_{n=0}^{\infty}(2p+n)a_{p+n}z^{2p+n}\right|-$$

$$\left|(B-A)(p-\alpha)-\sum_{n=0}^{\infty}[B(p+n)+(B-A)\alpha+Ap]a_{p+n}z^{2p+n}\right|$$

$$\leq\sum_{n=0}^{\infty}[(2+A+B)p+n(1+B)+(B-A)\alpha]a_{p+n}-(B-A)(p-\alpha)\leq 0;\ \text{by hypothesis.}$$

Hence by the maximum modulus theorem $f \in Q_p^*(A,B,\alpha)$. For the converse assume that

$$\left|\frac{\dfrac{zf'(z)}{f(z)}+1}{\dfrac{Bzf'(z)}{f(z)}+[pB+(A-B)(p-\alpha)]}\right| =$$

$$= \left|\frac{\displaystyle\sum_{n=0}^{\infty}(2p+n)a_{p+n}z^{2p+n}}{(B-A)(p-\alpha)-\displaystyle\sum_{n=0}^{\infty}[B(p+n)+(B-A)\alpha+Ap]a_{p+n}z^{2p+n}}\right| < 1.$$

Since $|\mathrm{Re}\,(z)| \leq |z|$ for all z we have

$$(4)\ \mathrm{Re}\left[\frac{\displaystyle\sum_{n=0}^{\infty}(2p+n)a_{p+n}z^{2p+n}}{(B-A)(p-\alpha)-\displaystyle\sum_{n=0}^{\infty}[B(p+n)+(B-A)\alpha+Ap]a_{p+n}z^{2p+n}}\right] < 1.$$

Choose values of z on the real axis so that $zf'(z)/f(z)$ is real. Upon clearing the denominator in (4) and letting $z \to 1$ through real values we obtain

$$\sum_{n=0}^{\infty}(2p+n)a_{p+n} \leq (B-A)(p-\alpha) - \sum_{n=0}^{\infty}[B(p+n)+(B-A)\alpha+Ap]a_{p+n}.$$

This gives the required relation. The result is sharp, the extremal function being

$$f(z) = \frac{1}{z^p} + \frac{(B-A)(p-\alpha)z^{p+n}}{\{(2+A+B)p+(1+B)n+(B-A)\alpha\}},\ n = 0,1,2,\ldots$$

3. DISTORTION THEOEM AND RADIUS OF CONVEXITY

Theorem 2. *If* $f(z) = \dfrac{1}{z^p} + \displaystyle\sum_{n=0}^{\infty} a_{p+n}z^{p+n}$ $(a_{p+n} \geq 0)$ *is in* $Q_p^*(A,B,\alpha)$ *then*

$$\frac{1}{r^p} - \frac{(B-A)(p-\alpha)r^p}{[(2+A+B)p+(B-A)\alpha]} \leq |f(z)|$$
$$\leq \frac{1}{r^p} + \frac{(B-A)(p-\alpha)r^p}{[(2+A+B)p+(B-A)\alpha]}$$

and

$$\frac{p}{r^{p+1}} - \frac{(B-A)(p-\alpha)r^{p-1}}{[(1+B)+(B-A)\alpha]} \leq |f'(z)|$$
$$\leq \frac{p}{r^{p+1}} + \frac{(B-A)(p-\alpha)r^{p-1}}{[(1+B)+(B-A)\alpha]}$$

Proof. From Theorem 1, we have

$$\begin{aligned}[(2+A+B)p+(B-A)\alpha]\sum_{n=0}^{\infty} a_{p+n} \\ &\leq \sum_{n=0}^{\infty}\{(2+A+B)p+(1+B)n+(B-A)\alpha\}a_{p+n} \\ &\leq (B-A)(p-\alpha)\end{aligned}$$

and we have,

$$\sum_{n=0}^{\infty} a_{p+n} \leq \frac{(B-A)(p-\alpha)}{[(2+A+B)p+(B-A)\alpha]}.$$

Hence

$$\begin{aligned}|f(z)| &\leq \frac{1}{r^p} + \sum_{n=0}^{\infty} a_{p+n}r^{p+n} \\ &\leq \frac{1}{r^p} + r^p\sum_{n=0}^{\infty} a_{p+n} \\ &\leq \frac{1}{r^p} + \frac{(B-A)(p-\alpha)r^p}{[(2+A+B)p+(B-A)\alpha]}\end{aligned}$$

and

$$\begin{aligned}|f(z)| &\geq \frac{1}{r^p} - \sum_{n=0}^{\infty} a_{p+n}r^{p+n} \\ &\geq \frac{1}{r^p} - r^p\sum_{n=0}^{\infty} a_{p+n} \\ &\geq \frac{1}{r^p} - \frac{(B-A)(p-\alpha)r^p}{[(2+A+B)p+(B-A)\alpha]}\end{aligned}$$

The bounds for $|f(z)|$ are sharp and are attained for the function:

$$f(z) = \frac{1}{z^p} + \frac{(B-A)(p-\alpha)z^p}{[(2+A+B)p+(B-A)\alpha]} \text{ at } z = r,\ re^{i\frac{\pi}{2p}}$$

Also we have

$$
\begin{aligned}
|f'(z)| &\le \frac{p}{r^{p+1}} + \sum_{n=0}^{\infty}(p+n)a_{p+n}r^{p+n-1} \\
&\le \frac{p}{r^{p+1}} + r^{p-1}\sum_{n=0}^{\infty}(p+n)a_{p+n}
\end{aligned}
$$

$$
\begin{aligned}
|f'(z)| &\ge \frac{p}{r^{p+1}} - \sum_{n=0}^{\infty}(p+n)a_{p+n}r^{p+n-1} \\
&\ge \frac{p}{r^{p+1}} - r^{p-1}\sum_{n=0}^{\infty}(p+n)a_{p+n}.
\end{aligned}
$$

Since

$$(B-A)\alpha+(1+B)(p+n) \le (1+B)n+(2+A+B)n+(B-A)\alpha,\ n = 0,1,2,\ldots$$

we have

$$
\begin{aligned}
&(1+B)+(B-A)\alpha\sum_{n=0}^{\infty}(p+n)a_{p+n} \\
&\quad \le \sum_{n=0}^{\infty}[(1+B)n+(B-A)\alpha+(2+A+B)p]a_{p+n} \\
&\quad \le (B-A)(p-\alpha)
\end{aligned}
$$

The bounds for $|f'(z)|$ follow.

Theorem 3. *If $f(z)$ is in $Q_p^*(A,B,\alpha)$ then $f(z)$ is meromerphically p-valently convex in*

$$0 < |z| < C_p = \inf_n \left[\frac{[(1+B)n+(2+A+B)p+(B-A)\alpha]p^2}{(B-A)(p-\alpha)(p+n)^2}\right]^{\frac{1}{(2p+n)}},\ n = 0,1,2,\ldots$$

The result is sharp.

Proof. It suffices to show that

$$\left|\frac{1+\frac{zf''(z)}{f'(z)}+p}{1+\frac{zf''(z)}{f'(z)}-p}\right| < 1,\ \text{for } 0 < |z| < C_p.$$

We have

$$\left|\frac{1+\frac{zf''(z)}{f'(z)}+p}{1+\frac{zf''(z)}{f'(z)}-p}\right| = \left|\frac{\sum_{n=0}^{\infty}(2p+n)(p+n)a_{p+n}z^{2p+n}}{2p^2+\sum_{n=0}^{\infty}n(p+n)a_{p+n}z^{2p+n}}\right|$$

$$\leq \frac{\sum_{n=0}^{\infty}(2p+n)(p+n)a_{p+n}|z|^{2p+n}}{2p^2-\sum_{n=0}^{\infty}n(p+n)a_{p+n}|z|^{2p+n}}.$$

The last expression is bounded above by 1 if

$$\sum_{n=0}^{\infty}\left[\frac{(p+n)}{p}\right]^2 a_{p+n}|z|^{2p+n} \leq 1. \tag{5}$$

By Theorem 1, we have

$$\sum_{n=0}^{\infty}\frac{[(2+A+B)p+(1+B)n+(B-A)\alpha]}{(B-A)(p-\alpha)} \leq 1. \tag{6}$$

In view of (6), (5) is true if

$$\left[\frac{(p+n)}{p}\right]^2 |z|^{2p+n} \leq \frac{[(2+A+B)p+(1+B)n+(B-A)\alpha]}{(B-A)(p-\alpha)},\quad n=0,1,\ldots \tag{7}$$

Solving (7) for $|z|$ we obtain

$$|z| \leq \left[\frac{[(1+B)n+(2+A+B)p+(B-A)\alpha]p^2}{(B-A)(p-\alpha)(p+n)^2}\right]^{\frac{1}{(2p+n)}}. \tag{8}$$

Putting $|z| = C_p$ in (8) the result follows.

The estimate is sharp for the function

$$f(z) = \frac{1}{z^p} + \frac{(B-A)(p-\alpha)z^{p+n}}{[(1+B)+(2+A+B)p+(B-A)\alpha]} \text{ for some } n.$$

4. CONVEX LINEAR COMBINATION

In this section we shall prove that the class $Q_p^*(A, B, \alpha)$ is closed under convex linear combinations.

Theorem 4. *Let* $f_{p-1}(z) = \dfrac{1}{z^p}$ *and*

$$f_{p+n}(z) = \frac{1}{z^p} + \frac{(B-A)(p-\alpha)z^{p+n}}{[(1+B)n + (2+A+B) + (B-A)\alpha]},\ n = 0, 1, 2, \ldots .$$

Then $f \in Q_p^*(A, B, \alpha)$ *if and only if it can be expressed in the form* $f(z) = \sum_{n=-1}^{\infty} \lambda_{p+n} f_{p+n}(z)$ *where* $\lambda_{p+n} \geq 0$ *and* $\sum_{n=-1}^{\infty} \lambda_{p+n} = 1$.

Proof. Suppose

$$\begin{aligned} f(z) &= \sum_{n=-1}^{\infty} \lambda_{p+n} f_{p+n}(z) \\ &= \frac{1}{z^p} + \sum_{n=0}^{\infty} \lambda_{p+n} f_{p+n}(z) \end{aligned}$$

Then

$$\sum_{n=0}^{\infty} \lambda_{p+n} \frac{(B-A)(p-\alpha)}{[(1+B)n + (2+A+B)p + (B-A)\alpha]} \cdot$$

$$\cdot \frac{[(1+B)n + (2+A+B)p + (B-A)\alpha]}{(B-A)(p-\alpha)} = \sum_{n=0}^{\infty} \lambda_{p+n} = 1 - \lambda_{p-1} \leq 1.$$

Hence, $f \in Q_p^*(A, B, \alpha)$. Then

$$a_{p+n} \leq \frac{(B-A)(p-\alpha)}{[(1+B)n + (2+A+B)p + (B-A)\alpha]},\ n = 0, 1, 2, \ldots$$

Setting

$$\lambda_{p+n} = \frac{[(1+B)n + (2+A+B)p + (B-A)\alpha]}{(B-A)(p-\alpha)} a_{p+n},\ n = 0, 1, 2, \ldots$$

and

$$\lambda_{p-1} = 1 - \sum_{n=0}^{\infty} \lambda_{p+n}$$

we obtain

$$f(z) = \sum_{n=-1}^{\infty} \lambda_{p+n} f_{p+n}(z).$$

REFERENCES

[1] Aouf M.K., *A certain subclass of meromorphically starlike functions with positive coefficients*, Rend. Mat. 9 (1989), 225–235.

[2] Cho N.E., Lee S.H., Owa S., *A class of meromorphic univalent functions with positive coefficients*, Kobe J., Math. 4 (1987), 43–50.

[3] Clunie J., *On meromorphic Schlicht functions*, J. London Math. Soc. 34 (1959), 215–216.

[4] Miller J., *Convex meromorphic mappings and related functions*, Proc. Amer Math. Soc. 25 (1970), 220–228.

[5] Mogra M.L., Reddy T.D., Juneja O.P., *Meromorphic univalent functions with positive coefficients*, Bull. Austral. Math. Soc. 32 (1985), 161-176.

[6] Uralegaddi B.A., Ganigi M.D., *Meromorphic multivalent functions with positive coefficients*, Nep. Math. Sci. Rep. 11 (1986), 95–102.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 179-183

ON CALCULATION PROCEDURES FOR SUMS OF INDEPENDENT RANDOM VARIABLES

Pavle MLADENOVIĆ
Faculty of Mathematics, University of Belgrade
Studentski trg 16, 11000 Belgrade, Yugoslavia

Abstract. In this paper we consider a double array of random variables which are independent in each row and have geometric distributions with different parameters. The condition of uniformly asymptotically negligibility is not satisfied for this double array, and consequently the limiting distribution of the sum S_n of the rundom variables in the nth row cannot be derived from classical limit theorems. Nevertheless, the exact distribution of the sum of random variables in the nth row can be determined.

Keywords: sum of independent random variables, geometric distribution, condition of uniformly asymptotically negligibility.

1. INTRODUCTION

Let

$$
\begin{array}{cccc}
X_{11}, & X_{12}, & \ldots & X_{1k_1}, \\
X_{21}, & X_{22}, & \ldots & X_{2k_2}, \\
\ldots & \ldots & \ldots & \ldots \\
X_{n1}, & X_{n2}, & \ldots & X_{nk_n}, \\
\ldots & \ldots & \ldots & \ldots
\end{array}
\tag{1}
$$

be a double array of random variables. For each $n \in \{1, 2, \ldots\}$ let us denote

$$
S_n = \sum_{i=1}^{k_n} X_{ni}, \quad \sigma_n^2 = \sum_{i=1}^{k_n} \mathrm{Var}(X_{ni}).
$$

We suppose that $k_n \to \infty$ as $n \to \infty$ and the random variables X_{ni} are independent in each row.

In dealing with the double array (X_{ni}) it is essential

(a) to determine the exact distribution of the sum S_n (if it is possible) or

(b) to determine the asymptotic distribution of the sum S_n and to investigate the question of the "speed" of convergence, i.e. the difference between the approximating expression and its limit.

The exact distribution of the sum S_n can be determined in some cases, for example if the terms X_{nj} have a normal distribution, Poisson distribution or uniform distribution on the segment $[0, 1]$, [2]. Of course, there are many cases for which it is not easy to find the exact distribution. In such cases the limit theorems for the sum S_n are of essential interest. If a sequence of distribution functions F_n (of random variables S_n) converges weakly to normal (0,1) distribution, as in the central limit theorem, then an estimate of the remainder term $F_n(x) - \Phi(x)$ is necessary for numerical calculation in many mathematical applications.

The classical assumption is that individual terms in the sum S_n are negligible. More precisely, the double array (X_{nj}) is uniformly asymptotically negligible if the following (UAN) condition is satisfied:

$$\lim_{n\to\infty} \max_{1\leq i\leq k_n} P\left\{ \frac{|X_{ni} - EX_{ni}|}{\sigma_n} \geq \varepsilon \right\} = 0 \quad \text{for every } \varepsilon > 0.$$

Classical limit theorems were established on the basis of the UAN condition and an exelent presentation of this theory is given in [1]. If the UAN condition is not satisfied one can use the classical method of truncation or the more sophisticated method of centres and scatters of probability distribution introduced in [3]. In this paper we shall consider an example of a double array of random variables for which the UAN condition is not satisfied.

2. A DOUBLE ARRAY OF RANDOM VARIABLES WITH GEOMETRIC DISTRIBUTIONS

In this paper we consider a double array of random variables

$$(2)\qquad \begin{array}{llll} X_{11}, & & & \\ X_{21}, & X_{22}, & & \\ \dots & \dots & \dots & \\ X_{n1}, & X_{n2}, & \dots & X_{nn}, \\ \dots & \dots & \dots & \dots \end{array}$$

where for every $n \geq 1$ the following conditions are satisfied:

(a) The random variables X_{ni}, $1 \leq i \leq n$ are defined at the same probability space $(\Omega, \mathcal{A}, \mathcal{P})$ and mutually independent;

(b) For every $i \in \{1, 2, \ldots, n\}$, the random variable X_{ni} has a geometric distribution with the parameter $p_{ni} = \dfrac{n-i+1}{n}$, i.e.

$$P\{X_{n1} = 1\} = p_{11} = 1,$$
$$P\{X_{ni} = k\} = p_{ni}(1-p_{ni})^{k-1}, \quad k = 1, 2, 3, \ldots; \quad 2 \leq i \leq n.$$

Theorem 1. *For the double array (2) with specified geometric distributions of the random variables X_{nj} the UAN condition is not satisfied.*

Proof. It is easy to prove that the following equalities hold true:

$$E(X_{ni}) = \frac{1}{p_{ni}}, \quad \operatorname{Var}(X_{ni}) = \frac{1}{p_{ni}^2} - \frac{1}{p_{ni}};$$

$$E(S_n) = n \sum_{i=1}^{n} \frac{1}{i} \sim n \ln n \quad (n \to \infty);$$

$$Var\,(S_n) = n^2 \sum_{i=1}^{n} \frac{1}{i^2} - n \sum_{i=1}^{n} \frac{1}{i} \sim \frac{n^2 \pi^2}{6} \quad (n \to \infty).$$

If $0 < \varepsilon < \sqrt{6}/(2\pi)$ and n is an even positive integer, then the following relations hold true:

$$\max_{1 \leq i \leq n} P\left\{ \frac{|X_{ni} - E(X_{ni})|}{\sigma_n} \geq \varepsilon \right\} \geq P\left\{ \frac{|X_{nn} - E(X_{nn})|}{\sigma_n} \geq \varepsilon \right\}$$
$$\geq P\{|X_{nn} - n| \geq \varepsilon\sigma_n\} > P\{X_{nn} \leq n - \varepsilon\sigma_n\} \geq P\left\{ X_{nn} \leq \frac{n}{2} \right\}$$
$$= 1 - P\left\{ X_{nn} > \frac{n}{2} \right\} = 1 - \left(1 - \frac{1}{n}\right)^{n/2} \to 1 - e^{-1/2} \quad (n \to \infty),$$

i.e. the double array (X_{ni}) is not uniformly asymptotically negligible (and consequently the Lindeberg condition is not satisfied).

Let $S_n = X_{n1} + X_{n2} + \ldots + X_{nn}$ be the sum of the random variables in the nth row of double array (2). It is interesting that the exact distribution of the sum S_n can be determined using some combinatorial reasoning. Actually, the following theorem holds true:

Theorem 2. *For every $n \geq 1$ the random variable S_n takes values from the set $\{n, n+1, \ldots\}$ with probability* 1, *and for every $k \in \{n, n+1, \ldots\}$ the following equality holds true:*

$$(3) \qquad P\{S_n = k\} = \sum_{i=0}^{n-2} (-1)^i \binom{n-1}{i} \left(\frac{n-i-1}{n}\right)^{k-1}.$$

Proof. We shall give the following interpretation of the random variables X_{ni}. For fixed $n \geq 1$, let us consider a sequence of independent trials, where each trial consists of choosing a number from the set $N_n = \{1, 2, \ldots, n\}$. We suppose that each of the numbers 1, 2, ..., n has the probability $1/n$ to be chosen in every trial. The trials should be repeated until each of the numbers 1, 2, ..., n occurs. Let us now suppose that after some number of trials exactly $i-1$ of numbers from the set N_n have been chosen, and let X'_{ni} be the number of new trials that should be made until the ith number from the set N_n occurs. Then, the random variable X'_{ni} has the same distribution as the random variable X_{ni}. We shall identify these two random variables. The sum $S_n = X_{n1} + X_{n2} + \ldots + X_{nn}$ can be interpreted as the number of trials until each of the numbers from set N_n occurs. In the future we shall use this interpretation of the random variables X_{ni} and S_n.

Let $\omega = c_1 c_2 \ldots c_m$ be one of the possible sequences of numbers from set N_n that can be chosen in the first m trials. We shall call the outcome ω: m-variation of the elements 1, 2, ..., n. Then, the probability of outcome ω after the first m trials is equal to $\dfrac{1}{n^m}$. Let $k \in \{n, n+1, \ldots\}$ be fixed and let A_k be the event that after the first $k-1$ trials each of the numbers 1, 2, ..., $n-1$ has been chosen, but the number n has not been chosen. Let us denote: M – the set of all $(k-1)$-variations of elements 1, 2, ..., $n-1$; M_i – the set of all $(k-1)$-variations of elements from the set $\{1, 2, \ldots, n-1\} \setminus \{i\}$, where $1 \leq i \leq n-1$. Then, we have $A_k = M \setminus (M_1 \cup M_2 \cup \ldots \cup M_{n-1})$. Using the principle of inclusion and exclusion (see [1]) we get that the number of elements of the set A_k is given by

$$\begin{aligned} |A_k| &= |M \setminus (M_1 \cup M_2 \cup \ldots \cup M_{n-1})| \\ &= \sum_{i=0}^{n-2} (-1)^i \binom{n-1}{i} (n-i-1)^{k-1}. \end{aligned}$$

Since every $(k-1)$-variation from A_k has the probability n^{-k+1}, it follows that the following equality holds true:

$$P(A_k) = \sum_{i=0}^{n-2} (-1)^i \binom{n-1}{i} \left(\frac{n-i-1}{n}\right)^{k-1}.$$

Let B_k be the event that after the first $k-1$ trials exactly $n-1$ of the numbers from the set N_n have been chosen. Then, we have $P(B_k) = nP(A_k)$. Since the numbers from set N_n will occur in the kth trial with the probability $\frac{1}{n}$, it follows that

$$\begin{aligned} P\{S_n = k\} &= \frac{1}{n} P(B_k) = P(A_k) \\ &= \sum_{i=0}^{n-2} (-1)^i \binom{n-1}{i} \left(\frac{n-i-1}{n} \right)^{k-1}. \end{aligned}$$

This completes the proof of Theorem 2.

3. A NUMERICAL EXAMPLE

Modern computers give us new possibilities in calculating the probabilities concerning the sums of independent random variables. The equalities (3) allow us to calculate the probabilities concerning the random variable S_n. For example, in the following table we give the values of probability p_n of the event $P\{S_n \leq ES_n\}$ for some values of n (for every $n \in \{6585, 6586, \ldots, 8614\}$ we have $p_n = 0.5708\ldots$):

n	p_n	n	p_n	n	p_n
2	0.750000	101	0.571868	6585	0.570880
3	0.617284	102	0.574419	6586	0.570864
4	0.622924	103	0.573733	6587	0.570894
5	0.606364	104	0.573028	...	0.5708...
6	0.582845	105	0.572307	8612	0.570877
7	0.611154	106	0.574636	8613	0.570854
8	0.579272	107	0.573858	8614	0.570857

REFERENCES

[1] Gnedenko B.V., Kolmogorov A. N. *Predelnie raspredelenije dlja nezavisymih slučajinyh veličin*, Gostehizdat, Leningrad-Moskva, 1949.

[2] Feller W., *An Introduction to Probability Theory and Its Applications*, Vol. I, 3rd ed., Wiley, New York, 1968.

[3] Zolotarev, V.M. *Centres and scatters of probability distributions*, Soviet. Math. Dokl. Vol. 11, No. 50, 1329–1331.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 185-195

VARIATIONAL INEQUALITIES WITH POINT-TO-SET OPERATORS

Đurica S. JOVANOV
Faculty of Organizational Sciences, University of Belgrade
Jove Ilića 154, 11040 Belgrade, Yugoslavia

Abstract. The paper considers variational inequalities with cone constraints. The characterization of solutions and sufficient conditions for the existence of solutions are given. An algorithm for the numerical solution of a problem is proposed and the convergence of the generated sequence of points is proved.

Keywords: variational inequality, iterative method, cone constraint.

1. INTRODUCTION

Let X be a real Banach space, $U \subset X$ a convex closed subset of the space X and $F : X \to 2^{X^*}$ a point-to-set mapping from X to its dual space X^*.

The variational inequality $VI(F,U)$ is the problem:

Find $u \in U$ such that there exists $y \in F(u)$ satisfying

$$(\forall\ v \in U) < y, v - u > \geq 0. \tag{1}$$

There are many problems which can be formulated as variational inequalities, for example convex programming problems, some problems from game theory, equilibrium problems in economics and traffic, linear complementarity problems, optimal control problems, etc.

This paper considers variational inequalities in a real Banach space, when the feasible set of points is described by the cone constraint. In Section 3 some existence theorems are proved. In Section 4 the characterization of solutions is given. In Section 5 an approximation of solutions of variational inequalities $VI(F,U)$ is presented and the convergence of the generated sequence of points to a solution is proved.

2. DEFINITION, NOTATION, PRELIMINARIES

A multi-valued mapping $F : X \to 2^{X^*}$ is said to be a monotone operator if

$$(\forall x_1, x_2 \in dom(F))(\forall y_1 \in F(x_1))(\forall y_2 \in F(x_2)) < y_1 - y_2, x_1 - x_2 > \geq 0,$$

where $dom(F) = \{u \in X | F(u) \neq \emptyset\}$. The monotone operator F is said to be strictly monotone if this inequality is strict for $x_1 \neq x_2$, and maximal monotone if its graph

$$\Gamma(F) = \{(x,y) \in X \times X^* | y \in F(x)\}$$

is not properly contained in the graph of any other monotone operator. The dual cone of a set $E \subset X$ is the closed convex cone

$$E^+ = \{v \in X^* | (\forall u \in E) < v, u > \geq 0\}$$

Let X, Y be Banach spaces; let $S \subset Y$ be a closed convex cone with a nonempty interior. We shall say that the mapping $g : X \to Y$ is S-convex if, for every $\alpha \in (0,1)$ and $u_1, u_2 \in X$, the following holds:

$$\alpha g(u_1) + (1-\alpha)g(u_2) - g(\alpha u_1 + (1-\alpha)u_2) \in S$$

Let $U \subset X$ be a closed convex subset of X. A point z is called a radial point of U if

$$(\forall w \in U)(\exists \alpha > 0) \ [z, z - \alpha(w - z)] \subset U.$$

If U has a nonempty relative interior riU with respect to some norm for X then U has a radial point.

3. EXISTENCE OF SOLUTIONS

In this Section we shall suppose that X is a real Hilbert space and $<,>$: $X \times X \to R$ is a bounded bilinear form. Then there is a linear continuous operator $A : X \to X$, such that

$$(\forall u, v \in X) < u, v >= (Au, v),$$

where $(,)$ denotes the inner product.

Lemma 3.1. *Let $G : X \to 2^X$ be a point-to-set mapping, which is upper semicontinuous and $G(u)$ is closed for every $u \in X$. Furtheremore, let us suppose that sequences of points $\{u_n\}$, $\{v_n\}$ and $\{\varepsilon_n\}$ satisfy conditions*

$$u_n \in dom(G), v_n \in X, \varepsilon_n \in \mathbf{R}_+$$

and

$$\lim_{n\to\infty} u_n = u \in dom(G), \lim_{n\to\infty} v_n = v, \lim_{n\to\infty} \varepsilon_n = 0$$

Then from

$$v_n \in G(u_n) + \varepsilon_n B$$

it follows that $v \in G(u)$.

Proof. Since G is upper semicontinuous, it follows that

$$(\forall\varepsilon > 0)(\exists\delta > 0)G(u + \delta B) \subset G(u) + \varepsilon B.$$

From

$$\lim_{n\to\infty} u_n = u, \lim_{n\to\infty} v_n = v \text{ and } \lim_{n\to\infty} \varepsilon_n = 0,$$

it follows that, for every $\varepsilon > 0$ there exist n_0, n_1, $n_2 \in N$ such that

$$(\forall n > n_0)\|v_n - v\| < \frac{\varepsilon}{3} \tag{2}$$

$$(\forall n > n_1)G(u_n) \subset G(u) + \frac{\varepsilon}{3}B \tag{3}$$

$$(\forall n > n_2)0 < \varepsilon_n < \frac{\varepsilon}{3} \tag{4}$$

Since $v_n \in G(u_n) + \varepsilon_n B$, for $n > \max\{n_0, n_1, n_2\}$, it follows that

$$v_n \in G(u_n) + \varepsilon_n B \subset G(u) + \frac{\varepsilon}{3}B + \frac{\varepsilon}{3}B = G(u) + \frac{2\varepsilon}{3}B,$$

and from (2), (4) we have

$$(\forall\varepsilon > 0)v \in G(u) + \varepsilon B.$$

Hence

$$v \in \bigcap_{\varepsilon>0}(G(u) + \varepsilon B) = G(u)$$

Theorem 3.1. *Let X be a real Hilbert space, $U \subset X$ a nonempty, convex and compact subset, $F : X \to 2^X$ a point-to-set mapping which is upper semicontinuous, $U \subset dom(F)$, $F(u)$ is a convex, compact subset of the space X, for every $u \in U$. Then there exist solutions of variational inequality $VI(F, U)$.*

Proof. Let $G : U \to 2^X$ be defined by

$$G(u) := u - A(F(u)),$$

where $A : X \to X$, is a linear continuous operator from X in X, which is associated to bilinear form $<,>$. Then G is upper semicontinuous, and for every $u \in U$, $G(u)$ is convex and compact. From Theorem 1.2.2. ([1]) it follows that for every $\varepsilon > 0$ there exists a locally Lipschitzean map $g_\varepsilon : X \to X$ such that

$$\Gamma(g_\varepsilon) \subset \Gamma(G) + \varepsilon B.$$

We define $f_\varepsilon : U \to U$, for $\varepsilon > 0$, with

$$f_\varepsilon(u) := P_U(g_\varepsilon(u)).$$

Since g_ε and P_U are continuous, it follows that f_ε is continuous too. Let $\{\varepsilon_n\}$ be a sequence of points with

$$\varepsilon_n \in \mathbf{R}_+ \quad \lim_{n\to\infty} \varepsilon_n = 0.$$

From Schauders Theorem, it follows that there exists $u_n \in U$ such that

$$u_n = f_{\varepsilon_n}(u_n) = P_U(g_{\varepsilon_n}(u_n) \in P_U(G(u_n) + \varepsilon_n B) \subset P_U(G(u_n)) + \varepsilon_n B$$

Let $H(u) := P_U(G(u))$. Then H is upper semicontinuous, and has the property

$$u_n \in H(u_n) + \varepsilon_n B \tag{5}$$

Since U is compact we can suppose that there exists $u \in U$ such that $\lim\limits_{n\to\infty} u_n = u$. From Lemma 3.1. and (5) it follows that

$$u \in H(u) = P_U(G(u)) = P_U(u - A(F(u)) \tag{6}$$

From (6) it follows that there exists $y \in F(u)$, such that

$$u = P_U(u - Ay).$$

From that we have

$$(\forall v \in U)(u, v-u) \geq (u - Ay, v-u)$$

$$(\forall v \in U) < y, v-u > \geq 0.$$

So u is a solution of the variational inequality $VI(F,U)$.

Theorem 3.2. *Let X be a real Hilbert space, $U \subset X$, a nonempty, convex and compact subset, $F : X \to 2^X$ a point-to-set mapping, which is lower semicontinuous, $U \subset dom(F)$, $F(u)$ is a convex, compact subset of the space X, for every $u \in U$. Then there exist solutions of variational inequality $VI(F,U)$.*

Proof. Let G be defined as in the proof of Theorem 3.1. Since $F(u)$, for $u \in U$, is compact and convex, from the lower semi-continuity of F, it follows that F is lower semi-continuous in the ε-sense. Then G is lower semi-continuous, and $G(u)$ is convex and compact for $u \in U$. From Theorem 1.11.1.([1]) it follows that there exists $g : U \to X$, which is a continuous selection from G. We define $f : U \to U$ with

$$h(u) := P_U(g(u)).$$

Then h is continuous. Since U is a convex, compact subset of X, from Schauders Theorem it follows that there exists $u \in U$ such that

$$u = h(u) = P_U(g(u)) \in P_U(G(u)) = P_U(u - A(F(u))),$$

and there exists $y \in F(u)$, such that

$$u = P_U(u - Ay).$$

As in the proof of Theorem 3.1 it follows, that u is a solution of the variational inequality $VI(F,U)$.

4. CHARACTERIZATION OF SOLUTIONS

In this Section the characterization of the solutions of variational inequality $VI(F,U)$ is given.

Let

$$U := \{u \in U_0 \mid -g(u) \in S\},$$

where $U_0 \subset X$ is a convex set and $g : X \to Y$ is S-convex, S is a closed convex cone with a nonempty interior.

Let us denote by H a set of generators for S^+; thus $H \subset S^+$ and S^+ is the closure of $\{\alpha h \mid h \in coH\}$.

We define

$$H_= := \{q \in H \mid \ < q, g(u) >= 0 \quad \text{for all} \quad u \in U\}$$

$$H_< := H \setminus H_=.$$

If Q is a weak* compact subset of $H_<$ such that $0 \notin coQ$ then we define

$$\Delta := \{u \in U_0 \mid \ < q, g(u) > \leq 0, \quad \text{for all } q \in H \setminus Q\}.$$

Let $K = Q^+$. Then $K^+ = cone(clcoQ)$.

The associated problem for $VI(F, U)$ is

$$(u, \lambda) \in \Delta \times K^+ \wedge (\forall (v, \mu) \in \Delta \times K^+)(< \lambda_0 F(u) + \partial_u < \lambda, g(u) >, v - u >$$

$$\geq 0 \wedge < -g(u), \mu - \lambda > \geq 0) \tag{7}$$

Theorem 4.1. *Let $F : X \to 2^{X^*}$ be a point-to-set mapping from the Banach space X to its dual X^*, $S \subset Y$ a closed convex cone in Banach space Y and S has a nonempty interior. Furtheremore let $U_0 \subset X$ be a convex subset of X and $U = \{u \in U_0 \mid -g(u) \in S\}$, where $g : X \to Y$ is S-convex, and there exists $(qg)'(u; d)$ for $q \in Q$, that is bounded on the set $\{d \in X \| \|d\| \leq 1\}$.*

(a) If (u, λ) is a solution of (7) for some $\lambda_0 > 0$ then u is a solution of $VI(F, U)$. If u is a solution of $VI(F, U)$ then there exists $\lambda_0 \geq 0, \lambda \in K^+$, such that $\lambda_0 + \|\lambda\| \neq 0$, so that pair (u, λ) is a solution of (7).

(b) If U has a radial point $\neq u$ and if u is a solution of $VI(F, U)$, then for some $\lambda \in K^+$ the pair (u, λ) is a solution of the system (7) for $\lambda_0 = 1$.

Proof. (a) From the proof of Theorem 1 ([12]), it follows that

$$u \in U \wedge -g(u) \in S \text{ if and only if } u \in \Delta \wedge -g(u) \in K. \tag{8}$$

Let (u, λ) be the solution of the system (7). From (7) it follows that

$$(\forall \mu \in K^+) < \mu, g(u) > \leq < \lambda, g(u) >,$$

and from that $< \lambda, g(u) >= 0$. From the S-convexity of g and $\lambda \in K^+$ it follows that $\lambda g : U_0 \to R$ is convex, and

$$(9) \qquad (\forall r \in \partial_u < \lambda, g(u) >) < r, v-u > \leq < \lambda, g(v) > - < \lambda, g(u) > \leq 0$$

whenever $v \in U$.

From (7) it follows that $(\forall \mu \in K^+) < \mu, g(u) > \leq 0$, and $-g(u) \in K$. From (7), (8) and (9) it follows that u is a solution of variational inequality $VI(F, U)$.

Let $u \in U$ be a solution of variational inequality (1). Let us define a continuous linear map $f : Y \to C(Q)$ by

$$f(y)(q) := q(y).$$

Continuity follows since Q is weak* compact. We denote by J the cone of nonnegative functions in $C(Q)$. Then

$$u \in \Delta \wedge -g(u) \in K \text{ if and only if } u \in \Delta \wedge -(fg)(u) \in J.$$

Hence there does not exist a solution $v \in \Delta$ such that

$$< F(u), v-u >< 0 \wedge -(fg)(v) \in J.$$

Alternative theorems ([11], Theorem 2) and ([10], Theorem 2.5.1) show that there exists $\lambda_0 \geq 0$ and $\rho \in J^+$, not both zero, for which

$$(10) \qquad (\forall v \in \Delta) < \lambda_0 F(u), v-u > + \rho(fg)(v) \geq 0$$

Then $\lambda = \rho f \in K^+$. From (10) it follows that $< \lambda, g(u) >= 0$ and the saddlepoint condition for the Lagrangian $L(v, \lambda, \lambda_0) = \lambda_0 < F(u), v-u > + < \lambda, g(v) >$ is

$$(\forall v \in \Delta)(\forall \mu \in K^+) L(u, \mu, \lambda_0) \leq L(u, \lambda, \lambda_0) \leq L(v, \lambda, \lambda_0).$$

From this it follows that (7) holds.

(b) From (a) it follows that there exist $\lambda_0 \geq 0$ and $\lambda \in K^+$ such that (7) holds. Suppose that $\lambda_0 = 0$; then $\rho \neq 0$. Since Q is compact, $\rho \in C(Q^*)$ is represented by a regular Radon measure ω on Q, where

$$(\forall \psi \in C(Q)) \rho\psi = \int_Q \psi d\omega.$$

Since $\rho(J) \subset R_+$, ω is a nonnegative measure. As $< \lambda, g(u) >= 0$ we have $\int_Q < q, g(u) > d\omega = 0$. Since $< q, g(u) >\leq 0$ and $\omega \geq 0$, ω vanishes except on $Q_1 = \{q \in Q | < q, g(u) >= 0\}$. But $\rho \neq 0$ and as ω is regular, $\omega(A) > 0$ for some compact $A \subset Q_1$. Then the following holds

$$(\forall q \in Q_1)\psi(q) = (qg)'(u; v-u) < 0.$$

$\psi(q)$ is negative and bounded away from zero on A, hence $(\forall r \in \partial_u < \lambda, g(u) >) < r, v-u >< 0$. But this is contradicted by (7) with $\lambda_0 = 0$. Hence $\lambda_0 \neq 0$, and $\lambda_0 = 1$ may be assumed. So (u, λ) satisfies the system (7) for $\lambda_0 = 1$.

5. APROXIMATION OF SOLUTIONS

Let X, Y be Hilbert spaces, and let $U_0 \subset X$ be a convex and closed set. Furthermore let $U = \{u \in U_0 \mid -g(u) \in S\}$, where $g : X \to Y$ is $S-$convex. The associated system for the variational inequality is

$$\begin{gathered}(u, \lambda) \in U_0 \times S^+ \wedge (\forall (v, \mu) \in U_0 \times S^+) \\ < F(u) + \partial_u < \lambda, g(u) >, v-u >\geq 0 \wedge < -g(u), \mu - \lambda >\geq 0.\end{gathered} \tag{12}$$

Then the following propositions hold.

Lemma 5.1. *If a pair (u, λ) is a solution of system (12) then u is a solution of $VI(F, U)$.*

Proof. Suppose that a pair (u, λ) is a solution of system (12). Then from (12) it follows $< \lambda, g(u) >= 0$ and for every $\mu \in S^+$ $< \mu, g(u) >\leq 0$. It follows from this that

$$-g(u) \in (S^+)^+ = S.$$

So $u \in U$. Let $v \in U$. Then for $r \in \partial_u < \lambda, g(u) >$

$$< r, v-u >\leq 0 \leq< \lambda, g(v) > - < \lambda, g(u) >=< \lambda, g(v) >$$

holds. From (12) it follows that

$$u \in U \wedge (\exists y \in F(u))(\forall v \in U) < y, v-u >\geq 0.$$

So u is a solution of variational inequality $VI(F, U)$.

Lemma 5.2. *If mapping F in variational inequality $VI(F, U)$ is maximal monotone, and $U \subset dom(F)$, then mapping $\Phi(u, \lambda) = (F(u) + \partial_u < \lambda, g(u) >, -g(u))$ in system (12) is maximal monotone too.*

Proof. $\Phi(u,\lambda) = (F(u) + \partial_u < \lambda, g(u) >, -g(u))$.

Since F is monotone and $\lambda g : U_0 \to R$ is convex, it follows that

$$< \Phi(u_1,\lambda_1) - \Phi(u_2,\lambda_2), (u_1,\lambda_1) - (u_2,\lambda_2) >=< F(u_1) - F(u_2), u_1 - u_2 > +$$
$$< \lambda_2, g(u_1) > - < \lambda_2, g(u_2) > - < \partial_u < \lambda_2, g(u_2) >, u_1 - u_2 > +$$
$$< \lambda_1, g(u_2) > - < \lambda_1, g(u_1) > - < \partial_u < \lambda_1, g(u_1) >, u_2 - u_1 >\geq 0.$$

So Φ is monotone and since F and $\partial_u(\lambda g)$ are maximal monotone it follows that Φ is maximal monotone ([18]).

Let the sequence of points $\{(u_n,\lambda_n)\}$ be defined by induction:

$$u_0 \in U_0, \lambda_0 \in S^+$$

and if $(u_n,\lambda_n) \in U_0 \times S^+$, (u_{n+1},λ_{n+1}) is defined by

$$u_{n+1} \in P_{U_0}\left(u_n - \alpha_n(F(u_n) + \varepsilon_n u_n)\right) \tag{13a}$$

$$\lambda_{n+1} = P_{S^+}(\lambda_n - \alpha_n(-g(u_n) + \varepsilon_n\lambda_n)) \tag{13b}$$

Using the results from [5] we can show that the sequence of points generated by algorithm (13) converges to the solution of system (12). Namely the following theorem holds:

Theorem 5.1. *Let X, Y be Hilbert spaces, $U \neq \emptyset$, $U_0 \subset dom(F)$ $F : X \to 2^X$, maximal monotone, $g : X \to Y$, $S-$convex, and there exists $L > 0$ such that*

$$(\forall u \in U_0)(\exists y \in F(u))\|y\| \leq L(1 + \|u\|)$$

$$(\forall u \in U_0, \lambda \in S^+)(\exists y \in \partial_u < \lambda, g(u) >)\|y\| \leq L(1 + \|u\|)$$

$$(\forall u \in U_0)\|g(u)\| \leq L(1 + \|u\|))$$

Furthermore suppose that for sequences (α_n) and (ε_n) the following holds:

$$\alpha_n > 0, \varepsilon_n > 0, \alpha_n \to 0, \varepsilon_n \to 0, \alpha_n = o(\varepsilon_n),$$

$$\sum_{n=1}^{\infty} \alpha_n\varepsilon_n = \infty, \lim_{n\to\infty} \frac{|\varepsilon_n - \varepsilon_{n+1}|}{\alpha_n\varepsilon_n^2} = 0.$$

Then a sequence of points $\{(u_n, \lambda_n)\}$ *generated by algorithm* (13) *converges to the solution of system* (12), *and a sequence of points* (u_n) *converges to some solution of the variational inequality* $VI(F, U)$.

Proof. The proposition follows from Theorem 3.1 ([5]). We will prove that operator $\Phi : U_0 \times S^+ \to X \times S$ satisfies the hypotheses of Theorem 3.1 ([5]). Indeed, if F is maximal monotone then Φ is maximal monotone too. Furthermore we have

$$\|\Phi(u, \lambda)\| = \|(F(u) + \partial_u < \lambda, g(u) >, -g(u))\| \leq$$

$$\|F(u)\| + \|\partial_u < \lambda, g(u) > \| + \|g(u)\| \leq$$

$$\|F(u)\| + \|\lambda\|\|\partial_u < \lambda, g(u) > \| + \|g(u)\| \leq 2L(1 + \|u\| + \|\lambda\|).$$

So Φ satisfies the condition in Theorem 3.1 ([5]). From this it follows that the sequence of points (u_n, λ_n) converges to some solution of (7), and then the sequence of points $\{u_n\}$ converges to some solution of $VI(F, U)$.

REFERENCES

[1] Aubin, J., P., Celina, A., *Differential Inclusions (Set-Valued Map and Viability Theory)*, Springer-Verlag, Berlin – Heidelberg – New York – Tokyo, 1984.

[2] Bakusinskij, A., B. *Reguljarizujusci algorithm na osnove metoda Njutona-Kantorovica dlja resenija variacionnih neravenstv* 16 (1976), 1397–1404.

[3] Bakusinskij, A., B. *Metodi resenija monotonih variacionnih neravenstv, osnovanie na principe iterativnoj reguljarizacii*, 6 (1976), 1350–1362.

[4] Bakusinskij, A., B. *Ekvivalentnie preobrazovanija variacionnih neravenstv i ih ispolzovanie*, DAN SSSR 247 (1979), 1297–1300.

[5] Bakusinskij, A.B., Goncarskij, A.V., *Iterativnie metodi resenija nekorrektnih zadac*, Nauka, Moscow 1989.

[6] Bakusinskij, A.B., Poljak, B.T., *O resenii variacionnih neravenstv*, DAN SSSR 219 (1974), 1038–1041.

[7] Browder F.E., *Existence and approximation of solutions of variational inequalities*, 56 (1966), 1080–1086.

[8] Chifot, M., *Variational Inequalities and Flow in Porous Media*, Springer- Verlag, New York – Berlin – Heidelberg – Tokyo, 1982.

[9] Craven, B.D., *Nonlinear programming in locally convex spaces*, JOTA 10 (1972), 197–210.

[10] Craven, B.D., *Mathematical Programming and Control Theory*, Chapman and Hall, London, 1978.

[11] Craven, B.D., Mond B., *Transposition theorems for cone-convex functions*, SIAM J.Appl.Math 24 (1973), 603–612.

[12] Craven, B.D., Zlobec, S., *Complete characterization of optimality for convex programming in Banach spaces*, Applicable Analysis 11 (1980), 67–78.

[13] Dafermos, S., *Traffic equilibrium and variational inequalities*, Transportation Science 14 (1980), 42–54.

[14] Dafermos, S.C., *An iterative scheme for variational inequalities*, Mathematical Programming 26 (1983), 40–47.

[15] Karamardian, S., *The complementarity problem*, Mathematical Programming 2 (1972), 107–129.

[16] Mancino, O., Stampachia, G., *Convex Programming and Variational Inequalities*, JOTA 9 (1972), 3–23.

[17] Pang, J., Chan, D., *Iterative methods for variational and complementarity problems*, Mathematical Programming 24 (1982), 284–313.

[18] Rockafellar R.L., *On the maximality of sums of nonlinear monotone operators*, Transactions of the American Mathematical Society 149 (1970), 75–88.

[19] Smith, M.J., *A descent algorithm for solving monotone variational inequalities and monotone complementarity problems*, JOTA 44 (1984), 485–495.

[20] Tobin R.L., *Sensivity analysis for variational inequalities*, JOTA 48 (1986), 191–204.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 197-199

SEPARATION OF CONVEX SETS VIA LINEAR INEQUALITIES

Đorđe DUGOŠIJA
Faculty of Mathematics, University of Belgrade,
Studentski trg 16, 11000 Belgrade, Yugoslavia

Abstract. The theory of linear inequalities is usually treated as a consequence of separation theorems for convex sets. We show that the conversion is also valid.

Keywords: Separation by hyperplanes, linear inequalities.

One of the most useful results in optimization theory is the following separation theorem for disjoint convex sets:

Theorem 1. *Let X and Y be two nonempty disjoint convex sets in R^n. Then there exists a hyperplane $\{x \mid a^T x = b\}$, $a \neq 0$ which separates them, that is,*

$$a^T x \leq b, \text{ for } x \in X$$
$$a^T y \geq b, \text{ for } y \in Y.$$

This fundamental result is used in deriving theorems of the Kuhn-Tucker type, necessary optimality conditions and some other results in convex analysis. Thus for instance, the following unsolvability theorem for linear inequalities is a consequence of Theorem 1:

Theorem 2. *The system of linear inequalities and equalities*

$$A_1 x \geq b_1$$
$$A_2 x = b_2$$

has no solution, iff there exist vectors $\lambda \geq 0$ and μ such that

$$\lambda^T A_1 + \mu^T A_2 = 0$$

$$\lambda^T b_1 + \mu^T b_2 > 0.$$

For the sake of completeness, we shall sketch the proof of the nontrivial part.
If the system has no solution, then

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \notin Y,$$

where

$$Y = \{\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} x - \begin{pmatrix} I \\ 0 \end{pmatrix} x_1 \mid x \in R^n,\ x_1 \geq 0\}.$$

Since Y is closed, there exists a convex ball X with center

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

disjoint with Y. Separating X and Y we find vectors λ and μ such that

$$(\lambda^T A_1 + \mu^T A_2)x - \lambda^T I x_1 \leq 0$$

for all $x \in R^n$ and all $x_1 \geq 0$. This gives

$$\lambda^T A_1 + \mu^T A_2 = 0$$

and $\lambda \geq 0$. On the other hand,

$$\lambda x_1 + \mu x_2 \geq 0, \text{ for } \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in X$$

gives $\lambda b_1 + \mu b_2 > 0$.
Bearing in mind that Theorem 2 has an elementary proof by induction in number of variables (see for instance Vajda [1]), it is interesting to note that we can directly prove Theorem 1 from Theorem 2 using as topological fact the well-known finite intersection theorem for compact sets (Berge-Ghouila Houri [2]).

Proof of Theorem 1. Choosing any $x^1, x^2, \ldots, x^k$ from X and $y^1, y^2, \ldots, y^l$ from Y we claim that the system in $(p_1, p_2, \ldots, p_k, q_1, q_2, \ldots, q_l)$

$$\sum_{i=1}^{k} p_i x^i - \sum_{j=1}^{l} q_j y^j = 0$$

$$\sum_{i=1}^{k} p_i = 1$$
$$\sum_{j=1}^{l} q_j = 1$$
$$p_i \geq 0,\ i = 1, 2, \ldots, k$$
$$q_j \geq 0,\ j = 1, 2, \ldots, l$$

has no solution. Using Theorem 2, we find a vector a, nonnegative numbers λ_i, $i = 1, 2, \ldots, k$ and μ_j, $j = 1, 2, \ldots, l$ and numbers λ and μ, such that

$$a^T x^i + \lambda + \lambda_i = 0, \text{ i=1, 2,..., k}$$
$$-a^T y^j + \mu + \mu_j = 0, \text{ i=1, 2,..., l}$$
$$\lambda + \mu > 0.$$

From these relations, we conclude that

$$a^T x^i < a^T y^j, \text{ for all } i = 1, 2, \ldots, k; j = 1, 2, \ldots, l,$$

thus $a \neq 0$. For $x \in X$, $y \in Y$ define

$$A_{x,y} = \{a \mid a^T(y - x) \geq 0 \text{ and } 1 \leq ||a|| \leq 2\}.$$

Then $A_{x,y}$ is a nonempty compact set and

$$\bigcap_{x \in \{x^1, x^2, \ldots, x^k\},\ y \in \{y^1, y^2, \ldots, y^l\}} A_{x,y} \neq \emptyset,$$

for any $x^1, x^2, \ldots, x^k; y^1, y^2, \ldots, y^l$. From the finite intersection theorem, it follows that

$$\bigcap_{x \in X,\ y \in Y} A_{x,y} \neq \emptyset.$$

Let a be any of its elements. Then, $a^T(y - x) \geq 0$ for all $x \in X$, $y \in Y$. The separating hyperplane is $a^T x = b$, where $b = \inf_{y \in Y} a^T y$.

REFERENCES

[1] Vajda S., *Mathematical Programming*, Addison-Wesley Publishing Company, Reading, Mass., 1961.

[2] Berge C., Ghouila Houri A., *Programming, Games and Transportation Networks*, John Wiley & Sons, Inc., New York,1965.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 201-212

BASIC DATA ENVELOPMENT ANALYSIS MODELS[†]

Milan M. MARTIĆ, Jovan J. PETRIĆ
Laboratory for Operations Research,
Faculty of Organizational Sciences, University of Belgrade,
Jove Ilića 154, 11000 Belgrade, Yugoslavia

Abstract. Data Envelopment Analysis (DEA) is a decision making tool based on linear programming for measuring the relative efficiencies of a set of comparable units. Besides the fact that it can identify which units are relatively efficient and which are not, for each inefficient unit, DEA identifies the sources and level of inefficiency for each of the inputs and outputs. This paper is a survey of basic DEA models and their comparison is given. The effect of model orientation (input or output) on the efficiency frontier and the effect of convexity requirements on returns to scale are examined.

Keywords: efficiency, DEA models, efficiency frontier.

1. INTRODUCTION

One of the most important principles in any business is the principle of efficiency which means that the best possible economic effects (outputs) are attained with as little economic sacrifices as possible (inputs). Efficiency can also be defined as a requirement that the desired targets be achieved with minimum use of the available resources. In order to assess the relative efficiency of a business unit, it is necessary to consider the conditions and operation results of other units of the same kind and determine the real standing of the results of such a comparison.

In the simplest case, where units have a single output and a single input, efficiency is defined as their ratio. However, more typical organizational units have multiple incommensurate inputs and outputs and the relative

[†]This paper has been partially supported by the Ministry of Science and Technology of Serbia (project code 10 T 02)

efficiency of any unit is defined as the ratio of a weighted sum of outputs and a weighted sum of inputs.

Recently, the Data Envelopment Analysis method has become popular in assessing the relative efficiency of business entities. DEA is a technique of mathematical programming that determines the efficiency of a unit based on its inputs and outputs, and compares it to other units involved in the analysis. It can best be described as data-oriented in that it effects performance evaluations and other inferences directly from observed data with minimal assumptions [5]. The efficiency of a Decision Making Unit (DMU) is measured relative to all other DMUs with the simple restriction that all DMUs lie on or below the extremal frontier. DEA is non-parametric method as it does not require any assumption about the functional form (e.g., a regression equation, a production function, etc.). It is a methodology directed to frontier rather than central tendencies. While statistical procedures are based on central tendencies, DEA is an extremal process. DEA analyzes each DMU separately and calculates a maximal performance measure for each.

At presents this is one of the most popular fields in operations research, a fact that is confirmed by a large number of papers published in the course of a year. The bibliographies released periodically by Seiford [9] already list more than 1000 references. These papers inform about the ample possibilities of using DEA to evaluate the performances of bank branches, schools, university departments, farming estates, hospitals and social institutions, military services, whole economic systems (regions) and others. DEA is a methodology comprised of several different interactive approaches and models used for the assessment of the relative efficiency of DMU and the assessment of the efficiency frontier. It supplies information important for managing the operations of both efficient and non-efficient units. This paper is a survey of basic DEA models and some of ways in which these models can be used are also given.

2. DEA MODELS

DEA methodology, originally proposed in [4], is used to assess the relative efficiency of a number of entities using a common set of incommensurate inputs to generate a common set of incommensurate outputs. The original motivation for DEA was to compare the productive efficiency of similar organizations, referred to as DMUs.

Let x_{ij} - denote the observed magnitude of an i - type input for entity j ($xij > 0$, $i = 1,2,...,m, j = 1,2,...,n$), and y_{rj} - the observed magnitude of an r-type output for entity j $(y_{rj} > 0, r = 1,2,...,s, j = 1,2,...,n)$. Then, the Charnes-Cooper- Rhodes (CCR) model is formulated in the following form for a selected entity k :

MODEL (M1)

$$\text{(1)} \qquad \text{Maximize} \quad h_k = \frac{\sum_{r=1}^{s} u_r y_{rk}}{\sum_{i=1}^{m} v_i x_{ik}}$$

subject to

$$\text{(2)} \qquad \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1, \ (j = 1, 2, \ldots, n)$$

$$\text{(3)} \qquad u_r \geq \varepsilon, \ \ (r = 1, 2, \ldots, s)$$

$$\text{(4)} \qquad v_i \geq \varepsilon, \ \ (i = 1, 2, \ldots, m)$$

where:

v_i - *the weights to be determined for input* i; m - *the number of inputs*;
u_r - *the weights to be determined for output* r; s - *the number of outputs*;
h_k - *the relative efficiency of* DMU_k; n - *the number of entities*;
ε - *a small positive parameter.*

Relative efficiency h_k, of one decision-making unit k, is defined as the ratio of the weighted sums of their outputs (virtual output) and the weighted sums of their inputs (virtual input). As for decision-making unit k (DMU_k), for which a maximum in objective function (1) is being searched, condition (2) is true, so we obviously have $0 < h_k \leq 1$, for each DMU_k. The weights v_i and u_r show the importance of each input and output and are determined in the model so that each DMU is as efficient as possible. Given that condition (2) is true for every DMU, this means that each of them lies on the efficiency frontier or beyond it. In case Max $h_k = h_k^* = 1$, this means that efficiency is being achieved, so we can say that DMU_k is efficient. Efficiency is not achieved for $h_k^* < 1$, and DMU_k is not efficient then.

DMUk is considered relatively inefficient if it is possible to expand any of its outputs without reducing any of its inputs, and without reducing any other output (output orientation), or if it is possible to reduce any of its inputs without reducing any output, and without expanding some other of its inputs (input orientation).

Problem (1) - (4) is nonlinear, nonconvex, with a linear and fractional objective function and linear and fractional constraints. The above ratio form yields an infinite number of solutions and if (u^*, v^*) is optimal, then $(\alpha u^*, \alpha v^*)$ is also optimal for $\alpha > 0$. Using a simple transformation developed by Charnes and Cooper (1962), the above CCR ratio model can be reduced to the linear programming form (Primal CCR model) so that LP methods can be applied. In this model the denominator has been set equal to 1 and the numerator is maximized. The input oriented CCR primal model is:

MODEL (M2)

$$(5) \qquad \text{Max} \quad h_k = \sum_{r=1}^{s} u_r y_{rk}$$

subject to

$$(6) \qquad \sum_{i=1}^{m} v_i x_{ik} = 1$$

$$(7) \qquad \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, \quad (j = 1, \ldots, n)$$

$$(8) \qquad u_r \geq \varepsilon, \quad r = 1, 2, \ldots, s$$

$$(9) \qquad v_i \geq \varepsilon, \quad i = 1, 2, \ldots, m$$

The mathematical model presented above is linear and can be solved using any of the well-known program packages for linear programming. However, in practice the dual task for problem (5) - (9) is often solved, which is:

MODEL (M3)

$$(10) \qquad \text{Min} \quad Z_k - \varepsilon\left(\sum_{r=1}^{s} s_r^+ + \sum_{i=1}^{m} s_i^-\right)$$

subject to

$$(11) \qquad \sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{rk}, \quad (r = 1, 2, \ldots, s)$$

$$(12) \qquad Z_k x_{ik} - \sum_{j=1}^{n} \lambda_j x_{ij} - s_i^- = 0, \ \ (i = 1, 2, \ldots, m)$$

$$(13) \qquad \lambda_j, s_r^+, s_i^- \geq 0; \ \ Z_k - \quad \text{sign unbound.}$$

The basic idea behind DEA is conveyed the best in the dual CCR model (M3) that is much easier to solve because of calculating size. The dual model for a given unit using input and output values of other units tries to construct a hypothetical composite unit out of existing units. If it is possible, the given unit is inefficient, otherwise it is efficient and lies at the efficiency frontier. The efficiency frontier is a set of segments interconnecting all the efficient DMUs and acts as an envelope for inefficient units. An inefficient unit can be enveloped below (input-oriented model) or above (output-oriented model).

Because the problems described by models (M2) and (M3) are associated, and also because of the duality theorem in linear programming, DMU_k is efficient if and only if the conditions for optimal solution $(\lambda^*, s^{+*}, s^{-*}, Z_k^*)$ are satisfied for problem (10) - (13):

$$(14) \qquad Z_k^* = 1$$

$$(15) \qquad s^{+*} = s^{-*} = 0 \text{ in all alternate optima}$$

Then, using optimal solution $(\lambda^*, s^{+*}, s^{-*}, Z_k^*)$ of problem (10) - (13), we can determine:

$$(16) \qquad X_k'' = Z_k^* X_k - s^{-*}$$

$$(17) \qquad Y_k'' = Y_k + s^{+*}$$

It can be shown that after CCR projection (16) and (17), DMU_k with altered inputs X_k'' and outputs Y_k'' becomes efficient. Differences $\Delta X_k = X_k - X_k''$, and the $\Delta Y_k = Y_k'' - Y_k$, show the estimated amount of input and output inefficiency, respectively. Thus it can be seen for inefficient DMU_k, how to change its inputs and outputs, so it can become efficient. We should emphasize that for each DMU_j $(j = 1, 2, ..., n)$ taken as DMU_k, the appropriate linear programming problem (10) - (13) is solved. Hence, we should solve n linear programming tasks with form (10) - (13), with $(s + m + 1)$ variables and $(s + m)$ constraints per task.

CCR models (dual and primal) with an input orientation are still the most widely known and used DEA models despite the numerous modified models that have appeared. CCR models assume constant returns to scale. DMU operates under constant returns to scale if an increase in inputs results in a proportionate increase in the output levels. These models calculate overall efficiency in which both pure technical efficiency and scale efficiency are aggregated into a single value. The envelopment surface obtained from the CCR model has the shape of a convex cone. The efficient DMUs would lie on top of the facets, while the inefficient ones would be covered under the cone. In a single input and output case the efficiency frontier reduces to a straight line. The CCR model yields the same efficiencies regardless of whether it is input or output oriented.

Certainly the most important extension of the original CCR models is given in [1], where Banker, Charnes and Cooper introduced one more additional constraint in model (M3):

$$\sum_{j=1}^{n} \lambda_j = 1 \tag{18}$$

This constraint makes possible variable returns to scale and provides that the reference set is formed as a convex combination of the DMUs which are in it (the ones that have positive values for λ in the optimal solution). A DMU operates under variable returns to scale if it is suspected that an increase in inputs would not result in a proportional change in the outputs. The convexity constraint ensures that the composite unit is of similar scale size as the unit being measured. This model (BCC model) yields a measure of the pure technical efficiency that ignores the impact of sale-size by comparing a DMU only to other units of similar scale. Often, small units are qualitatively different from large units and a comparison between the two may distort measures of comparative efficiency. This measure of efficiency is always at least equal to the one given by the CCR model. The envelopment surface obtained from the BCC model results in a convex hull.

DEA model can be input- or output-oriented. The input-oriented model contracts inputs as far as possible while controlling outputs. In input-oriented models, an inefficient unit is made efficient through the proportional reduction of its inputs, while its output proportions are held constant. The output-oriented model expands outputs as far as possible while controlling inputs. In output-oriented models, an inefficient unit is made efficient through the proportional increase of its outputs, while the input' proportions remain unchanged. The input and output measures are always the same in the CCR model, but frequently differ in the BCC model. Thus, if we are using the CCR model, we can solve one model and give either interpretation. If we solve the BCC input model, we can give only an input interpretation, and we must solve the BCC output model for an output interpretation. Another difference between the BCC and CCR models deals with scalar transformations of all data for a given DMU. The efficiency measure in the CCR model is unchanged by scalar transformations, since the efficiency ratio of the scaled DMU is unchanged. On the other hand, scalar transformations of a given DMU change the scalar size and could easily affect the efficiency measures from the BCC model.

A family of related basic DEA models (input- and output-oriented) is presented in Table 1, see [10]. Model PI_0 is an input-oriented primal CCR model. Models PI_1, PI_2 and PI_3 are obtained from PI_0 by adding a variable u_*. PI_3 is an input-oriented primal BCC model. Model DI_0 is an input-oriented dual CCR model. Models DI_1, DI_2 and DI_3 are obtained from DI_0 by adding a constraint on the sum of the multipliers λ. DI_3 is an input-oriented dual BCC model. Similarly, we get models PO_ρ, $(\rho = 1, 2, 3)$ from PO_0 by adding a variable ν_*, and DO_ρ, $(\rho = 1, 2, 3)$ from DO_0 by adding the same envelopment constraint.

Models I_1 (PI_1 and DI_1) and O_1 (PO_1 and DO_1) allow increasing returns to scale. A DMU operates at increasing returns to scale if a proportionate increase in all of its inputs results in a greater than proportionate increase in its outputs. These models are hybrid models in which the efficient frontier consists of two parts from the previous models, the lower portion of the ray segment and the upper portion of the convex hull boundary segments. Models I_2 (PI_2 and DI_2) and O_2 (PO_2 and DO_2) allow decreasing returns to scale. A DMU is said to operate at decreasing returns to scale if a proportionate increase in all of its inputs results in a less than proportionate increase in its outputs. These models are hybrid models in which the efficient frontier consists of two parts, the lower portion of the surface of the convex hull and the upper portion of the ray segment. Thus, these hybrid models are completely determined by CCR (PI_0, POI_0, DI_0 and DO_0) and BCC (PI_3, PO_3, DI_3 and DO_3) models.

An inefficient DMU can be made more efficient by projection onto the frontier. Model orientation determines the direction of projection for inefficient DMUs. In an input orientation one improves efficiency through the proportional reduction of inputs, whereas an output orientation requires proportional augmentation of outputs. For the input models an inefficient DMU, (X_0, Y_0), is projected "back" to the boundary point $(\theta X_0, Y_0)$. θ is the contraction factor in DI_ρ models. For the output models an inefficient DMU, (X_0, Y_0), is projected "up" to the boundary point $(X_0, \phi Y_0)$. ϕ is the expansion factor in DO_ρ models. These boundary points may be efficient for one orientation and inefficient for the other.

Table 1: DEA models

Input oriented	
Miltiplier problem (PI_ρ)	Envelopment problem (DI_ρ)
$\max_{\mu,\nu} z = \mu^T Y_0 + u_*,$	$\min_{\theta,\lambda} \theta$
s.t. $\nu^T X_0 = 1,$	s.t. $Y\lambda \geq Y_0,$
$u_* e^T + \mu^T Y - \nu^T X \leq 0,$	$\theta X_0 - X\lambda \geq 0,$
$\mu^T \geq 0,$	θ free, $\lambda \geq 0.$
$\nu^T \geq 0,$	
where	
$u_* = 0$ in $PI_0,$	For DI_0 : append noting
≤ 0 in $PI_1,$	For DI_1 : append $e^T\lambda \leq 1$
≥ 0 in $PI_2,$	For DI_2 : append $e^T\lambda \geq 1$
free in $PI_3.$	For DI_3 : append $e^T\lambda = 1$

Output oriented	
Miltiplier problem (PO_ρ)	Envelopment problem (DO_ρ)
$\min_{\mu,\nu} q = \nu^T X_0 + v_*,$	$\max_{\theta,\lambda} \phi$
s.t. $\mu^T Y_0 = 1,$	s.t. $X\lambda \leq X_0,$
$v_* e^T - \mu^T Y + \nu^T X \geq 0,$	$\phi Y_0 - Y\lambda \leq 0,$
$\mu^T \geq 0,$	ϕ free, $\lambda \geq 0.$
$\nu^T \geq 0,$	
where	
$v_* = 0$ in $PO_0,$	For DO_0 : append noting
≤ 0 in $PO_1,$	For DO_1 : append $e^T\lambda \leq 1$
≥ 0 in $PO_2,$	For DO_2 : append $e^T\lambda \geq 1$
free in $PO_3.$	For DO_3 : append $e^T\lambda = 1$

The effect of model selection can be interpreted equivalently in terms of:

(i) a restriction on λ,

(ii) a restriction on the supporting hyperplane, or

(iii) a restriction on the returns to scale allowed.

These equivalencies are given in the next Table [10]:

	Restrictions		
Model	λ	Hyperplane	Returns to scale
$\rho = 0$	None	Passes through origin	Only constant allowed
$\rho = 1$	$\Sigma\lambda_i \leq 1$	$u_* \leq 0$ ($\nu_* \geq 0$)	Increasing is allowed
$\rho = 2$	$\Sigma\lambda_i \geq 1$	$u_* \geq 0$ ($\nu_* \leq 0$)	Decreasing is allowed
$\rho = 3$	$\Sigma\lambda_i = 1$	None	None

We have presented basic DEA models here. We can find in the literature numerous extensions of basic DEA models, see [5]. Some of these are:

- constraints are brought to weights for particular inputs and outputs,
- constraints are brought to amounts of particular virtual inputs and outputs,
- inputs and outputs that cannot be controlled are brought into the analysis,
- categorical variables are brought into the model,
- models for ranking relatively efficient DMUs are developed.

3. DEA MODEL UTILIZATION

This sections discusses how DEA models can be used to assess DMUs. A key stage in a DEA assessment is identification of the input/output variables pertaining to the units being assessed, see [2]. Since DEA is used to evaluate performances by directly considering input and output data, results will naturally depend exclusively on the input/output choice for the analysis and the number and homogeneity of the DMUs to be evaluated. In this stage it is important to consult the people working in the units which are to be evaluated so that major inputs and outputs can be identified properly. In principle, it is important to envelop in the analysis all important inputs, namely all the resources used and all important outputs, namely the products and services produced. However, a large number of inputs and outputs compared to the number of units to be evaluated may reduce the discriminating power of the method. The larger the number of inputs and outputs compared to the number of the units to be evaluated, the greater the chances that the units will allocate appropriate weights to a single subset of inputs and outputs that will make them appear efficient. In order to preserve the discriminating power of the method, the number of the units to be evaluated should be much larger than the number of inputs and outputs. Some authors suggest from experience that the number of DMUs should exceed the number of inputs and outputs by at least twice. Boussofiane et al. in [2] propose testing the correlation between inputs and outputs, as one of the possible ways to reduce their number. If a pair of inputs is positively correlated then they are multiples of each another, and one may be omitted without any implications upon the efficiency to be rated. The same applies to outputs. The availability of data may also affect the choice of inputs and outputs in practice. In case data on an input or output are not available then the possibility should be checked of using a substitute for which such data will be either available or can be relatively easily obtained.

DEA is a methodology of several different interactive approaches and models used to assess the relative efficiency of DMUs and to assess the efficiency frontier. It supplies information important for managing the operations of both efficient and nonefficient units. For each inefficient unit, DEA identifies a set of relatively efficient units making thus a peer group for the inefficient unit. The peer set for an inefficient unit consists of units having the same optimum weights as the inefficient unit, but having a relative efficiency rating of 1. Such peer units are identified rather easily by the fact that they all have a positive value for λ in the optimum solution to (M3) for an inefficient unit. The identification of peer groups should be very useful in practice. Peer units can be used to highlight the weak aspects of the performance of the corresponding inefficient unit. The input/output levels of a peer unit can also sometimes prove useful target levels for the inefficient unit.

The solution of any DEA model provides information as to how much relatively inefficient units should reduce their inputs, or increase their outputs, to become relatively efficient. For each inefficient DMU (one that lies below the frontier), DEA identifies the sources and level of inefficiency for each input and output. The level of inefficiency is determined by comparison to a single reference DMU or a convex combination of other relevant DMUs located on the efficient frontier that utilize the same level of inputs and produce the same or a higher level of outputs. We have seen in the previous section that we can get this using the optimal solution to model (M3) and relations (16) and (17). We can get to similar information by a sensitivity analysis of the optimal solution in model (M1). These results are very important to managers, because they pointing to the sources of inefficiency for relatively inefficient DMUs.

Efficiency improvement is not only inefficient but also efficient units can be attained by identifying what is efficient operating practice. It can be usually found in the relatively efficient units. However, among the relatively efficient units some are better than others to set a good example. The need to distinguish the relatively efficient units and find out what a good operating practice is, emerges from the essence of a DEA model that allows a unit to select the weights that will show it as having maximum efficiency. In this way the units may appear efficient because within their choice of weights all very small input subsets are ignored. Moreover, the inputs and outputs assigned larger weights could be given secondary importance while those that are ignored could be associated with the units' main functions.

To distinguish the relatively efficient DMUs Boussofiane et al. [2] suggested the following methods (or a combination of these):

- cross efficiency matrix,
- distribution of virtual inputs and outputs,
- weight restriction,
- frequency by which an efficient unit appears in the peer groups.

4. CONCLUSION

DEA is non-parameter methodology for evaluating the efficiency of non-profit DMUs. It consists of solving several mutually connected linear programming mathematical models for each DMU. While each of these models addresses managerial and economic issues and provides useful results, their orientations are different and, more importantly, they generalize and provide contact with these disciplines and concepts. Thus, models may focus on increasing, decreasing, or constant returns to scale as found in economics that are here generalized to the case of multiple outputs.

The extensive but probably incomplete bibliography [9] is intended to document the diffusion and growth of DEA. The bibliography is evidence of DEA applications involving a wide range of contexts, such as education (public schools and universities), health care (hospitals, clinics, physicians), banking, armed forces (recruiting, aircraft maintenance), auditing, sports, market research, mining, agriculture, retail outlets, organization effectiveness, transportation (ferries, highway maintenance), public housing, index number construction, benchmarking, etc. Our most importance experiences in DEA application are presented in [6], [7] and [8]. In [8], DEA is applied in evaluating the relative efficiency of 32 branches of "INVESTBANKA" bank in Belgrade. The results obtained in [7] show that DEA can be very successfully used in supporting the decision making process of investment banks. The DEA method is applied in assessing the relative efficiency of 20 investment programs in agriculture. In [6], it is shown how DEA can be used for the comparative analysis and ranking of 30 districts in Serbia.

REFERENCES

[1] Banker, R., Charnes A. and Cooper W., *Some models for estimating technical and scale inefficiencies in data envelopment analysis*, Management Science 30(1984), 1078-1092.

[2] Boussofiane, A., Dyson G. and Thanassoulis E., *Applied data envelopment analysis*, European Journal of Operational Research 52(1991), 1-15.

[3] Charnes, A., Cooper W., Lewin, A., Seiford, L. (eds), *Data Envelopment Analysis: Theory, Methodology and Application*, Kluwer, Boston, 1995.

[4] Charnes, A., Cooper W. and Rhodes E., *Measuring the efficiency of decision making units*, European Journal of Operational Research 2(1978), 429-444.

[5] Cooper, W., Tone K., *A Survey of Some Recent Developments in Data Envelopment Analysis*, Invited lecture presented at EURO XIV Conference Jerusalem, 1995, 149-166.

[6] Lečić, D., Martić M., Petrić J, *Comparative efficiency analysis of districts in Serbia using data envelopment analysis*, Proceedings of SymOrg '96, Vrnjačka Banja, 1996, 597-602. (in Serbian)

[7] Martić, M., Krčevinac S., Petrić J.,*An application of data envelopment analysis to assess the relative efficiency of investments in agriculture*, Studies in Regional & Urban Planning, 4(1996), 141 - 152.

[8] Martić, M., Krčevinac S., Žarkić N., Dragojević D., *An application of DEA for measuring the efficiency of bank branches in Yugoslavia*, Proceedings of 3rd Balkan Conference on OR, Thessaloniki, 1995.

[9] Seiford, L., *A Bibliography of data envelopment analysis (1978-1994)*, Working paper, The University of Massachusetts, Amherst, 1995.

[10] Seiford, L. and Thrall R., *Recent developments in DEA: the mathematical programming approach to frontier analysis*, Journal of Econometrics 46(1990), 7-38.

[11] Wong, Y. and Beasley J., *Restricting weight flexibility in DEA*, Journal of the Operational Research Society 41(9) (1990), 829-835.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 213-220

ON A GENERALIZATION OF A FIRST ORDER STEP – SIZE ALGORITHM †

Nada I. ĐURANOVIĆ – MILIČIĆ
Department of Mathematics
Faculty of Technology and Metallurgy
University of Belgrade
Karnegijeva 4, 11 000 Belgrade, Yugoslavia

Abstract. In this paper we present a generalization of the Armijo step-size algorithm. This generalization of the Armijo algorithm is based on so-called "forcing functions". It is proved that this generalized algorithm is well-defined. Proof is given of the convergence of the obtained sequence of points to a first-order point of the problem of the unconstrained optimization, as well as an estimate of the rate of convergence.

Keywords: unconstrained optimization, forcing function, step-size algorithm

1. INTRODUCTION

We consider the problem of unconstrained optimization:

$$(1) \qquad \min\{\varphi(x) \mid x \in D\},$$

where $\varphi : D \subset R^n \to R$ is a continuously differentiable function on the open set D.

We consider iterative algorithms for finding an optimal solution to problem (1) generating sequences of points $\{x_k\}$ of the following form:

$$(2) \qquad x_{k+1} = x_k - \alpha_k s_k, \quad k = 0, 1, \ldots,$$

†This research was supported by the Science Fund of Serbia, grant number 04M03, through the Institute of Mathematics, Serbian Academy of Science and ART

where we suppose that the direction vector s_k (where s_k is a $n\times 1$ direction of search) satisfies the condition:

$$(3) \qquad s_k \neq 0, \quad \langle \nabla\varphi(x_k), s_k\rangle \geq 0,$$

and the step-size α_k is defined by a special step-size algorithm.

The original Armijo step-size algorithm [4] defines the step-size α_k for the sequence $\{x_k\}$ satisfying the relations (2) and (3) in the following way:

$$\alpha_k = 0 \qquad if \qquad \langle \nabla\varphi(x_k), s_k\rangle = 0;$$

otherwise, $\alpha_k > 0$ is a number satisfying

$$\alpha_k = 2^{-i(k)},$$

where $i(k)$ is the smallest integer from $i = 0, 1, \ldots,$ such that

$$x_k - 2^{-i}s_k \in D$$

and

$$\varphi(x_k) - \varphi(x_k - 2^{-i}s_k) \geq \gamma\cdot 2^{-i}\langle \nabla\varphi(x_k), s_k\rangle,$$

where $0 < \gamma < 1$ is a preassigned constant.

At first, we shall give some preliminaries which we need in the following text.

Definition 1. *(See [5]). A mapping* $\sigma\colon [0,\infty) \to [0,\infty)$ *is a forcing function if for any sequence* $\{t_k\} \subset [0,\infty)$

$$\lim_{k\to\infty} \sigma(t_k) = 0 \quad implies \quad \lim_{k\to\infty} t_k = 0$$

and $\sigma(t) > 0$ *for* $t > 0$.

(The concept of the forcing function was first introduced by Elkin in [3].)

Definition 2. *(See [5]). Let* $\{x_k\} \subset R^n$ *be any sequence converging to* $\overline{x}$. *Then the* R *– convergence factors are defined as follows:*

$$R_p\{x_k\} = \begin{cases} \lim\limits_{k\to\infty} \sup \|x_k - \overline{x}\|^{\frac{1}{k}} & if \quad p = 1, \\ \lim\limits_{k\to\infty} \sup \|x_k - \overline{x}\|^{\frac{1}{p^k}} & if \quad p > 1. \end{cases}$$

If $0 < R_1\{x_k\} < 1$, *the sequence* $\{x_k\}$ *is said to converge to* $\overline{x}$ *at least* R *– linearly.*

Lemma 1. *Let* $\varphi\colon D \subset R^n \to R$ *be continuously differentiable on an open set* $D_0 \subset D$ *and suppose that* $\{x_k\} \subset D_0$ *converges to* $x^* \in D_0$. *Assume that* $\nabla\varphi(x^*) = 0$, *that* φ *has a second derivative at* x^* *and the Hessian matrix* $H(x^*)$ *is invertible, and that there is an* $\eta > 0$ *and a* k_0 *for which*

$$\varphi(x_k) - \varphi(x_{k+1}) \geq \|\nabla\varphi(x_k)\|^2, \quad \forall k \geq k_0.$$

Then $R_1\{x_k\} < 1$.

Proof. See [5].

Lemma 2. (See [6]). *Let* $\varphi\colon D \subset R^n \to R$ *be twice continuously differentiable and let there exist an* m, $0 < m < \infty$ *such that*

$$m\|y\|^2 \leq \langle y, H(x)y\rangle \quad \textit{for all } x \in D,\ y \in R^n.$$

Then the function φ *is strongly convex and the set* $\{x \mid \varphi(x) \leq \varphi(x_0)\}$ *is bounded for any* $x_0 \in D$.

2. A GENERALIZATION OF THE FIRST ORDER ARMIJO STEP - SIZE ALGORITHM

We consider a sequence of points $\{x_k\}$ with properties (2) and (3) where the step α_k is defined in the following way:

$$\alpha_k = 0 \quad \text{if} \quad \langle\nabla\varphi(x_k), s_k\rangle = 0;$$

otherwise, $\alpha_k > 0$ is a number satisfying

$$\alpha_k = q^{-i(k)},\ q > 1,$$

where $i(k)$ is the smallest integer from $i = 0, 1, \ldots,$ such that

$$x_k - q^{-i}s_k \in D \tag{4A}$$

and

$$\varphi(x_k) - \varphi(x_k - \alpha_k s_k) \geq \alpha_k\sigma\big(\langle\nabla\varphi(x_k), s_k\rangle\big), \tag{4B}$$

where $\sigma : [0,\infty) \to [0,\infty)$ is a forcing function such that $\delta_1 t \leq \sigma(t) \leq \delta_2 t$ for every $t \geq 0$ and some $0 < \delta_1 < \delta_2 < 1$.

Since D is open, $x_k - q^{-i}s_k \in D$ for sufficiently large i. The existence of a finite $i(k)$ such that $\alpha_k = q^{-i(k)}$ satisfies (4) is proved in the following lemma.

Lemma 3. *Let* $\varphi\colon D \subset R^n \to R$ *be continuously the differentiable on an open set* $D_0 \subset D$ *and suppose that* x_k, $\alpha_k = q^{-i}$ *and* s_k *satisfy* $\langle \nabla\varphi(x_k), s_k\rangle > 0$, $[x_k, x_k - q^{-i}s_k] \subset D$ *and*

$$\varphi(x_k) - \varphi(x_k - q^{-i}s_k) < q^{-i}\sigma\big(\langle \nabla\varphi(x_k), s_k\rangle\big)$$

for some i $(= 0, 1, \ldots)$ *and* $q > 1$, *where* $\sigma\colon [0,\infty) \to [0,\infty)$ *is a forcing function such that* $\sigma(t) \leq \delta t$ *for every* $t \geq 0$ *and some* $0 < \delta < 1$.

Then there exists a finite $i^* > i$ such that (4) holds.

Proof. Define the function $F\colon [0,1] \to R$ in the following way:

$$F(\lambda) = \begin{cases} \dfrac{\varphi(x_k) - \varphi(x_k - \lambda\cdot q^{-i}s_k)}{\lambda q^{-i}\langle \nabla\varphi(x_k), s_k\rangle} & , \quad \lambda \in (0,1]; \\ 1 & , \quad \lambda = 0, \end{cases}$$

By L'Hospital's rule, we have $F(\lambda) \to 1$ as $\lambda \to 0$, so that F is continuous on $[0,1]$. Consequently, since $F(0) = 1$ and

$$F(1) < \frac{\sigma\big(\langle \nabla\varphi(x_k), s_k\rangle\big)}{\langle \nabla\varphi(x_k), s_k\rangle} \leq \delta < 1,$$

$f(\lambda)$ takes on all valus between $\frac{\sigma\left(\langle \nabla\varphi(x_k), s_k\rangle\right)}{\langle \nabla\varphi(x_k), s_k\rangle}$ and 1. Hence, there exists a finite $\overline{i}$, $\overline{\lambda} = q^{-\overline{i}}$, $0 < \overline{\lambda} < 1$, such that

$$\frac{\sigma\big(\langle \nabla\varphi(x_k), s_k\rangle\big)}{\langle \nabla\varphi(x_k), s_k\rangle} \leq F(\overline{\lambda}) < 1,$$

i.e. (4) will be satisfied for $\overline{\alpha}_k = q^{-\overline{i}}q^{-i}$, i.e. for $i^* = i + \overline{i}$.

Theorem 1. *Let* $\varphi\colon D \subset R^n \to R$ *be a continuously differentiable function on the open set* D. *Let the sequence* $\{x_k\}$ *be defined by relations (2), (3) and (4). Let* $\overline{x} \in D$ *be a point of accumulation of* $\{x_k\}$ *and* K_1 *a set of indices such that* $x_k \to \overline{x}$ *for* $k \in K_1$. *Assume that there exists a* $\beta > 0$ *such that*

$$\langle \nabla\varphi(x_k), s_k\rangle \geq \beta\|\nabla\varphi(x_k)\|\,\|s_k\| \quad \textit{for } k \in K_1 \textit{ and } \nabla\varphi(x_k) \neq 0,$$

that the sequence $\{s_k\}$ *is uniformly bounded* $(k \in K_1)$ *and*

$$\|s_k\| \geq \mu\big(\|\nabla\varphi(x_k)\|\big) \quad \textit{for all } k \in K_1,$$

where $\mu\colon [0,\infty) \to [0,\infty)$ *is a forcing function such that* $\mu(t) \geq mt$ *for some* $\mu > 0$. *Then* $\nabla\varphi(\overline{x}) = 0$.

Proof. There are two cases to consider.
a) The set of indices $\{i(k)\}$ for $k \in K_1$ is uniformly bounded above by a number I.

Since, by (4) the sequence $\{\varphi(x_k)\}$ is monotone decreasing and each $\langle \nabla\varphi(x_k), s_k\rangle \geq 0$ it follows that

$$\begin{aligned}
\varphi(x_0) - \varphi(\overline{x}) &= \sum_{k=0}^{\infty}\left[\varphi(x_k) - \varphi(x_{k+1})\right] \geq \sum_{k\in K_1}\left[\varphi(x_k) - \varphi(x_{k+1})\right] \geq \\
&\geq \sum_{k\in K_1} q^{-i(k)}\sigma\big(\langle\nabla\varphi(x_k), s_k\rangle\big) \geq q^{-I}\sum_{k\in K_1}\delta_1\langle\nabla\varphi(x_k), s_k\rangle \geq \\
&\geq \delta_1 q^{-I}\sum_{k\in K_1}\beta\|\nabla\varphi(x_k)\|\,\|s_k\| \geq \\
&\geq \delta_1\beta q^{-I}\sum_{k\in K_1}\|\nabla\varphi(x_k)\|\mu\left(\|\nabla\varphi(x_k)\|\right) \geq \\
&\geq m\delta_1\beta q^{-I}\sum_{k\in K_1}\|\nabla\varphi(x_k)\|^2.
\end{aligned}$$

Since $\varphi(\overline{x})$ is finite and since $\|\nabla\varphi(x_k)\|^2 \geq 0$, it follows that $\nabla\varphi(x_k) \to 0$ for $k \in K_1$. Hence, by continuity of $\nabla\varphi$, we have $\nabla\varphi(\overline{x}) = 0$. b) There is a subset of indices $K_2 \subset K_1$ such that $\lim\limits_{k\in K_2} i(k) = \infty$. Because of the definition of $i(k)$, then either

$$x_k + q^{-i(k)+1}s_k \notin D$$

or

$$\varphi(x_k) - \varphi(x_k - q^{-i(k)+1}s_k) < q^{-i(k)+1}\sigma\big(\langle\nabla\varphi(x_k), s_k\rangle\big). \tag{5}$$

If the cause for termination of iteration k were that $x_k + q^{-i(k)+1}s_k \notin D$ infinitely often, then, since $i(k) \to \infty$ for $k \in K_2$ and because the sequence $\{s_k\}$ is uniformly bounded, it follows that $\overline{x}$ is on the boundary of D. Since D is an open set, $\overline{x} \notin D$, which is a contradiction to the theorem assumption.

Thus, without loss of generality (5) can be consider to hold for all $k \in K_2$.

Because φ is by assumption continuously differentiable and $\{s_k\}$ is uniformly bounded, it follows that (5) can be written as:

$$\begin{aligned}
\varphi(x_k) - \varphi(x_k - q^{-i(k)+1}s_k) &= q^{-i(k)+1}\langle\nabla\varphi(x_k), s_k\rangle + o\big(q^{-i(k)+1}\|s_k\|\big) < \\
&< q^{-i(k)+1}\sigma\big(\langle\nabla\varphi(x_k), s_k\rangle\big) \\
&< q^{-i(k)+1}\delta_2\langle\nabla\varphi(x_k), s_k\rangle.
\end{aligned}$$

Hence,

$$q^{-i(k)+1}\langle\nabla\varphi(x_k), s_k\rangle(1-\delta_2) < o\big(q^{-i(k)+1}\|s_k\|\big).$$

Dividing by $2^{-i(k)+1}\|s_k\|$ yields

$$\frac{o\big(q^{-i(k)+1}\|s_k\|\big)}{q^{-i(k)+1}\|s_k\|} > \frac{\langle\nabla\varphi(x_k), s_k\rangle(1-\delta_2)}{\|s_k\|} \geq \beta(1-\delta_2)\|\nabla\varphi(x_k)\|$$

Because $\{s_k\}$ is uniformly bounded, taking the limit as $k \to \infty$ for $k \in K_2$ yields, by continuity of $\nabla\varphi$:

$$\beta(1-\delta_2)\|\nabla\varphi(\overline{x})\| \leq 0.$$

Since $\beta > 0$, $1-\delta_2 > 0$, it follows $\nabla\varphi(\overline{x}) = 0$.

Theorem 2. *Let the assumptions of Theorem 1 be satisfied. Let additionally, the function $\varphi\colon D \subset R^n \to R$ be twice continuously differentiable and such that there exists an $l > 0$ satisfying*

$$l\|y\|^2 \leq \langle y, H(x)y\rangle \quad \textit{for all } x \in D,\ y \in R^n. \tag{6}$$

Then the sequence $\{x_k\}$ generated by the generalized Armijo algorithm converges to $\overline{x}$, where $\overline{x}$ is the unique optimal solution to problem (1), at least R - linearly.

Proof. From condition (6) it follows that the function φ is, by Lemma 2, strongly convex and that the level set $N = \big\{x \in D \mid \varphi(x_k) \leq \varphi(x_0)\big\}$ for some $x_0 \in D$ is convex and compact. Furthemore, since φ is twice continuously differentiable, from (6) it also follows that there exists an $L > 0$. $L \geq l$ such that

$$l\|y\|^2 \leq \langle y, H(x)y\rangle \leq L\|y\|^2 \quad \text{for all } x \in N,\ y \in \mathbf{R}^n. \tag{7}$$

From relation (4) we have that $\varphi(x_k) > \varphi(x_{k+1})$; hence $x_{k+1} \in N$ if $x_k \in N$.

By Theorem 1 we have that $\|\nabla\varphi(\overline{x})\| = 0$.

By strong convexity of φ it follows that $\overline{x}$ is the unique optimal solution to problem (1).

Denote by F the following function:

$$F(x, s, \alpha) = \varphi(x) - \varphi(x - \alpha s).$$

By Taylor's theorem we have

$$F(x_k, s_k, \alpha) = \alpha\langle\nabla\varphi(x_k), s_k\rangle - \alpha^2\int_0^1 (1-t)\langle s_k, H(x_k - t\alpha s_k)s_k\rangle dt.$$

From this equation, in view of (7), it follows that

$$F(x_k, s_k, \alpha) \geq \alpha\langle \nabla\varphi(x_k), s_k\rangle - \frac{1}{2}\alpha^2 L\|s_k\|^2.$$

Using the function F, because of the definition of $i\,(k)$, we have:

$$F(x_k, s_k, q^{-i\,(k)-1}) < q^{-i\,(k)-1}\sigma\big(\langle \nabla\varphi(x_k), s_k\rangle\big),$$

i.e., by introducing the function $\overline{F}$:

$$\begin{aligned}
\overline{F}(x_k, s_k, q^{-i\,(k)-1}) &= F(x_k, s_k, q^{-i\,(k)-1}) - q^{-i\,(k)-1}\sigma\big(\langle \nabla\varphi(x_k), s_k\rangle\big) < 0 \\
\overline{F}(x_k, s_k, q^{-i\,(k)-1}) &\geq q^{-i\,(k)-1}\langle \nabla\varphi(x_k), s_k\rangle - \frac{1}{2}q^{-2i\,(k)-2}L\|s_k\|^2 - \\
&- q^{-i\,(k)-1}\sigma\big(\langle \nabla\varphi(x_k), s_k\rangle\big) \geq \\
&\geq q^{-i\,(k)-1}\cdot\beta\|\nabla\varphi(x_k)\|\,\|s_k\| \\
&- q^{-i\,(k)-1}\delta_2\beta\|\nabla\varphi(x_k)\|\,\|s_k\| - \frac{1}{2}q^{-2i\,(k)-2}L\|s_k\|^2 = \\
&= q^{-i\,(k)-1}\beta(1-\delta_2)\|\nabla\varphi(x_k)\|\,\|s_k\| \\
&- \frac{1}{2}q^{-2i\,(k)-2}L\|s_k\|^2 < 0.
\end{aligned}$$

From the last inequality it follows that

$$\begin{aligned}
q^{-i\,(k)-1} &\geq \frac{2\beta(1-\delta_2)\|\nabla\varphi(x_k)\|}{L\|s_k\|}, \text{ i.e.} \\
q^{-i\,(k)} &\geq \frac{q\cdot 2\cdot\beta(1-\delta_2)\|\nabla\varphi(x_k)\|}{L\cdot\|s_k\|}.
\end{aligned}$$

Finally, from $(4B)$ it follows:

$$\begin{aligned}
\varphi(x_k) - \varphi(x_k - q^{-i\,(k)}s_k) &\geq q^{-i\,(k)}\sigma\big(\langle \nabla\varphi(x_k), s_k\rangle\big) \geq \\
&\geq \frac{2\cdot q\cdot\beta^2\cdot(1-\delta_2)\delta_1}{L}\|\nabla\varphi(x_k)\|^2 \\
&= \eta\|\nabla\varphi(x_k)\|^2,
\end{aligned}$$

where $\eta = \dfrac{2q\beta^2\delta_1(1-\delta_2)}{L} > 0$.

From the last inequality, by Lemma 3 it follows that the sequence $\{x_k\}$ converges to $\overline{x}$ at least R – linearly (where $\overline{x}$ is, as we have already proved, the unique optimal solution to problem (1)).

3. CONCLUSION

Because of general assumptions about the objective function φ, the generalized algorithm can be used to solve a wide class of unconstrained optimization problems. Also, the choice of forcing functions $\sigma(t)$ with the property $\delta_1 t \leq \sigma(t) \leq \delta_2 t$ $0 < \delta_1 < \delta_2 < 1$, is wide.

Finally, this generalized algorithm can be used to solve constrained optimization problems (see [1], [2]) when constraints are adequately considered.

REFERENCES

[1] Djuranović – Miličić N., *An Algorithm in Constrained Optimization,* in: M. Thoma, A. Wyner, (eds.), *Lecture Notes in Control and Information Sciences*, Springer–Verlag, Berlin 1986, 203–208.

[2] Djuranović – Miličić N., *A Generalization of the Céa–Goldstein Step-size Algorithm,* European Journal of Operational Research 63 (1992), 1–9.

[3] Elkin R., *Convergence theorems for Gauss–Siedel and other minimization algorithms,* Doctoral Thesis, University of Maryland, College Park, 1968.

[4] Mc Cormick G. P., *Nonlinear Programming, Theory, Algorithms and Applications,* Wiley, New York, 1983.

[5] Ortega J., Rheinboldt W., *Iterative Solution of Nonlinear Equations in Several Variables,* Academic Press, New York, 1970.

[6] Polak E., *Computational Methods in Optimization, A Unified Approach,* Academic Press, New York, 1971.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 221-228

ON THE PROBLEM OF DISCRETE OPTIMIZATION WITH PHASE CONSTRAINTS

Slobodan V. DAJOVIĆ
Faculty of Organizational Sciences, University of Belgrade
Jove Ilića 154, 11040 Belgrade, Yugoslavia

Abstract. We consider the problem of discrete optimization with phase constraints, using Clarke's tangent cone for an extension of the Milyutin-Dubovitskii approach, based on the Intersection Principle as an abstract condition which satisfy nonconvex sets approximated by convex cones, to nonsmooth optimization problems.

Keywords: Intersection Principle, Clarke's tangent cone, nondifferentiability

1. INTRODUCTION

Milyutin and Dubovitskii formulated a general approach for a broad class of extremal problems with differentiable constraints, based on the extension of the Kuhn-Tucker theorem to infinite-dimensional spaces and an analysis of Pontrjagin' s Maximum Principle [7], called the Intersection Principle.

Let $\Omega_0, \Omega_1, \ldots, \Omega_n$ be simultaneosly the subsets of the linear topological locally convex space R^n and the convex cones. Suppose that Ω_0 is a subspace and $\Omega_i \equiv int\Omega_i, i = 1, \ldots, n$. Then

$$\bigcap_{i=0}^{n} \Omega_i = \emptyset \iff \sum_{i=0}^{n} \lambda_i = 0 \text{ (at least one } \lambda_j \neq 0),$$

$$\lambda_i \in \Omega_i^*, \lambda_i(x) \geq 0, \forall x \in \Omega_i, i = 0, 1, \ldots, n.$$

This approach has the disadvantage that all but one of the convex sets must be open. Boltyanskii resolved this problem in the finite-dimensional case using the definition of the tent (tangential convex cone) as convex approximants generated with constraints. His "method of tents" is applicable to a broad class of optimization problems with differentiable data.

Definition 1. *Let $C \subset R^n$ and $x_0 \in C$. A convex cone K is a (smooth) tent for C at x_0 if there exist a neighoorhood Σ of x_0 and a (smooth) continuous map $\Psi : \Sigma \to R^n$ satisfying conditions:*

$$\Psi(K \cap \Sigma) \subseteq C;$$

$$\Psi(x) = x + o(x - x_0), \forall x \in \Sigma.$$

Using the "method of tents", Boltyanskii formulated and proved the necessary conditions of optimality for discrete control problems with and without phase constraints in the form of Pontrjagin's Maximum Principle.

In Watkins' paper [11] it is shown that Clarke's tangent cone satisfies the intersection theorem and, as a corollary, the Intersection Principle. The Milytin-Dubovitskii approach is thus extended to the nondifferentiable optimization theory through the use of Clarke's tangent cone and so the necessary conditions of optimality for discrete control problems could be extended to a nonsmooth problem with and without phase constraints.

2. CLARKE'S TANGENT CONE AND THE INTERSECTION PRINCIPLE

The following neighbourhood characterization of Clarke's tangent cone $T_C(x_0)$ of a closed set $C \subseteq R^n$ at $x_0 \in C$ is equivalent to Clarke's definition in [4].

Definition 2. *Let $C \subseteq R^n$ be closed and $x_0 \in C$.*

$$v \in T_C(x_0) \text{ if and only if } \forall \varepsilon > 0, \exists \delta, \lambda > 0 \text{ such that}$$

$$x' + t(v + \varepsilon B) \cap C \neq \emptyset,\ \forall t \in [0, \lambda], \forall x' \in C \cap (x_0 + \delta B)$$

for all v in any compact subset $S \subset T_C(x_0)$.

Watkins [8] proved that a solid Clarke's tangent cone ($intT_C(x_0) \neq \emptyset$) is a tent for C at x_0 and that a family of Clarke's tangent cones satisfy:

The Intersection Principle: *A system of tangential approximants satisfies the Intersection Principle if, whenever K_i, $i = 1, \dots, m$ are tangential approximants at x_0 for sets C_i ($C_i \subset R^n, i = 1, \dots, m$), and at least one cone K_i is not a subspace, and the family of cones is inseparable, then the set $\bigcap_{i=1}^{m} C_i$ contains points arbitrarily close to, but distinct from x_0.*

In the same paper Watkins proved that a family of Clarke's tangent cones also satisfies the following intersection theorem:

Theorem 1. *Let $C_i, i = 1, \dots, m$ be closed subsets of R^n with $x_0 \in C_{m+1} = \bigcap_{i=1}^{m} C_i$, and assume that Clarke's tangent cones $T_{C_i}(x_0), i = 1, \dots m$ are inseparable. Then*

$$\bigcap_{i=1}^{m} T_{C_i}(x_0) \subset T_{C_{m+1}}(x_0).$$

Theorem 2. *Clarke's tangent cone satisfies the Intersection Principle.*

3. DISCRETE OPTIMIZATION PROBLEMS

We consider the control problem in which the trajectories and controls are defined with sequences

$$x(0), x(1), \dots, x(N); \;\; u(0), u(1), \dots, u(N), \tag{1}$$

for each $t = 0, 1, \dots, N$.

Here $x(t) \in R^n, u(t) \in R^m$ and for each $x(t)$ there exists $u(t) \in U_t$, $t = 0, 1, \dots, N-1$; thus the phase state can be uniquely defined:

$$x(t+1) = f_t(x(t), u(t)), \;\; t = 0, 1, \dots, N-1 \tag{2}$$

We shall now discuss the following discrete control problem:

(P_1) : For given initial phase state $x(0)$ define discrete process

$$u(0), u(1), \dots, u(N-1) \tag{3}$$

(4) $$x(0), x(1), \ldots, x(N)$$

satisfying functional relation (2) and minimizing functional

(5) $$J = \sum_{k=0}^{N-1} f_k^0(x(k), u(k))$$

If we suppose that for each $t = 0, 1, \ldots, N$ the phase constraints

(6) $$x(t) \in M_t \subset R^n,$$

are satisfied then we can consider the following discrete control problem with phase constraints:

(P_2) For initial phase state $x(0)$ define discrete process (3), (4), satisfying functional relation (2) and phase constraints (6), and minimizing functional (5).

If

$$M_t \equiv R^n, t = 0, 1, \ldots, N,$$

we have the discrete control problem without constraints, denoted by (P_1).

4. MATHEMATICAL PROGRAMMING PROBLEM

Formulated discrete control problems (P_1) and (P_2) could be considered as mathematical programming problems.

Let us denote

(7) $$x_t^i = x^i(t),\ u_\tau^j = u^j(\tau),$$

$i = 0, 1, \ldots, m;\ j = 1, \ldots, r;\ t = 1, \ldots, N;\ \tau = 0, 1, \ldots, N-1,$

and consider x_t^i, u_τ^j as coordinates of point $z = (x_t^i, u_\tau^j) \in R^n, n = (m+r)N$. Thus, every discrete process (3), (4) corresponds to a point $z \in R^n$, satisfying (2), i.e.

(8) $$-x_{t+1}^i + f_t^i(x_t, u_t) = 0;\ i = 1, \ldots, m;\ t = 0, 1, \ldots, N-1,$$

where, for $t = 0$, $x_0 = (x_0^1, \ldots, x_0^m) = x(0)$.

If we denote

$$\Sigma_\tau = \{z = (x^i, u^j_\tau) \in R^n : u_\theta \in U_\theta\}, \tag{9}$$

where the coordinates $x^i_t, u^j_\tau (\tau \neq \theta)$ are arbitrary, we have that

$$z = (x^i_t, u^j_\tau) \in \Sigma_\tau,\ \tau = 0, 1, \ldots, N-1, \tag{10}$$

and on the contrary, if $z = (x^i_t, u^j_\tau)$ satisfies (8), (9), then discrete process (3), (4), defined by (7) will satisfy condition (2) and $u(\tau) \in U_\tau$.

Thus, if we denote by Σ the set of all points $z \in R^n$ satisfying (8), (10), we have a one-to-one correspondance between discrete process (2), (3), (4) and the points of the set Σ.

Consider the functions

$$F^i_t(z) = -x^i_{t+1} + f^i_t(x_t, u_t), i = 1, \ldots, m; t = 0, 1, \ldots, N-1, \tag{11}$$

and denote with Ω^i_t the set of all the points $z \in R^n$ satisfying the conditions $F^i_t(z) = 0$, and with

$$\Omega^* = \bigcap_{i=1}^{m} \bigcap_{i=0}^{N-1} \Omega^i_t,$$

the set of all the points $z \in R^n$, satisfying the equations

$$F^i_t(z) = 0, i = 1, \ldots, m; t = 0, 1, \ldots, N-1.$$

Then

$$\Sigma = \Omega^* \cap \Sigma_0 \cap \cdots \cap \Sigma_{N-1}.$$

If we denote

$$F^0(z) = \sum_{t=0}^{N-1} f^0_t(x_t, u_t),$$

then it is evident that the value of functional (5) is $J = F^0(z)$, and that the discrete problem (P_1) is equivalent to the problem of determining the minimum of function $F^0(z)$ on set Σ.

Definition 3. *The bounded function $f : R^n \to R$ is a locally Lipschitz function if for any set $B \subset R^n$ there exists a constant $K > 0$:*

$$|f(z_1) - f(z_2)| \leq K|z_1 - z_2|, \forall z_1, z_2 \in B.$$

Definition 4. *The epigraph of a bounded locally Lipschitz function* $f : R^n \to R$, *denoted by* $epi(f)$, *is the set*

$$epi(f) = \{(z, s) \in R^n \times R : f(z) \leq s\}.$$

Definition 5. *The generalized gradient of a bounded locally Lipschitz function* $f : R^n \to R$ *at point* $z_0 \in R^n$ *is the set*

$$\partial f(z_0) = \{z \in R^n : (z, -1) \in N_E(z_0, f(z_0))\},$$

where $N_E(z_0, f(z_0))$ *is the polar cone of closed set* E, *defined by the epigraph of function* f *at the point* z_0.

Remark 1. The connection between the generalized gradient of a locally Lipschitz function at a point z_0 and Clarke's tangent cone to the epigraph of f at $(z_0, f(z_0))$ is expressed by the polar relation

$$N_E(z_0) = (T_E(z_0))^o = \{z \in R^n : y \in T_E(z_0), < y, z > \leq 0\}.$$

Clarke exploited these notions in [4], [5] in deriving the necessary optimality conditions for nonconvex nonsmooth variational problems.

Assume that $F^0(z), g_i(z), h_j(z) (i = 1, \ldots, n; j = 1, \ldots, m)$ are bounded locally Lipschitz functions mapping the Banach space $X \to R$, and given closed subset $\Sigma \subset X$. Then the following could be proved:

Theorem 3. (Clarke) *If* $F^0(z_0) = \min_{z \in \Sigma} F^0(z)$ *and the constraints*

$$g_i(z) \leq 0, \ i = 1, \ldots, n; \tag{12}$$

$$h_j(z) = 0, \ j = 1, \ldots, m, \tag{13}$$

hold, then there exist the numbers $r_0, r_i, s_j (i = 1, \ldots, n; j = 1, \ldots, m)$, *at least one of them is not* 0, *and the point* $z^* \in X^*$, *satisfying the following conditions:*

1. $r_0 \geq 0, \ r_i \geq 0, \ i = 1, \ldots, n;$
2. $r_i g_i(z_0) = 0, \quad i = 1, \ldots, n;$
3. $z^* \in r_0 \partial F^0(z_0) + \sum_{i=1}^{n} r_i \partial g_i(z_0) + \sum_{j=1}^{m} s_j \partial h_j(z_0),$
4. $- z^* \in N_\Sigma(z_0).$

This theorem can be proved by the same method Boltyanskii used to demonstrate a similar proposition in [2]. This treatment implies the construction of Clarke's tangential cones for the set Σ at the point z_0, and the sets defined by the inequalities (12) and the equalities (13). After that, the Intersection Principle is applied to the considered family of Clarke's tangential cones.

The same analogy can be developed for the problem (P_2) with phase constraints if we assume that the set Σ includes the sets defined by the inequalities and the equalities corresponding to the phase constraints.

Theorem 4. *If* $F^0(z_0) = \min_{z\in\Sigma} F^0(z)$ *and the constraints*

$$\begin{aligned} g_i(z) &\le 0, \quad i = 1,\ldots,n; \\ h_j(z) &= 0, \quad j = 1,\ldots,m; \\ p_l(z) &\le 0, \quad l = 1,\ldots,n_s; \\ q_k(z) &= 0, \quad k = 1,\ldots,m_s, \end{aligned}$$

hold, then there exist the numbers r_0, r_i, s_j, t_l, w_k $(i = 1,\ldots,n;\ j = 1,\ldots,m;$ $l = 1,\ldots,n_s;\ k = 1,\ldots,m_s)$ *at least one of them is not* 0, *and the point* $z^* \in X^*$ (X^* *conjugate space of* X), *satisfying the following conditions:*

1. $r_0 \ge 0,\ r_i \ge 0\,(i = 1,\ldots,n),\ p_l \ge 0\,(l = 1,\ldots,n_s)$;

2. $r_i g_i(z_0) = 0\,(i = 1,\ldots,n),\ t_l p_l(z_0) = 0\,(l = 1,\ldots,n_s)$;

3. $$z^* \in r_0\partial F^0(z_0) + \sum_{i=1}^{n} r_i\partial g_i(z_0) + \sum_{j=1}^{m} s_j\partial h_j(z_0) + \sum_{l=1}^{n_s} t_l\partial p_l(z_0) + \sum_{k=1}^{m_s} w_k\partial q_k(z_0);$$

4. $-z^* \in N_\Sigma(z_0)$.

REFERENCES

[1] Boltyanskii V.G., *Svoistvo otdelimosti sistemy vypuklyh konusov*, Izv. AN Arm. Matematika 7(4)(1972).

[2] Boltyanskii V.G., *Neobhodimye uslovia v zadachah matematicheskogo programmirovania*, Kishinev, Matem. issl. 4(26)(1972).

[3] Boltyanskii V.G., *Metod shatrov v teorii ekstremalnyh zadach*, Uspehi Matem. nauk, T.30, N°3(1975).

[4] Clarke F.H., *Generalized gradients and applications*, Trans. of the Amer. Math. Society, 205(1975).

[5] Clarke F.H., *A new approach to Lagrange multipliers*, Math.of Op. R., Vol 1, N°2(1976).

[6] Clarke F.H., *Optimization and Nonsmooth Analysis*, Wiley Publ., New York, 1983.

[7] Dubovitskii A.,Ya, Milyutin A.A., *Zadachi na ekstremum pri nalicii ogranichenii*, DAN SSSR T.149, N°4(1963).

[8] Gardner R.J., Martin D.H., Watkins G.G., *Indicating cones and the intersection principle for tangential approximants in abstract multiplier rules*, JOTA, Vol. 33, N°4(1981).

[9] Milytin A.A., *Obscie shemy poluchenia neobhodimyh uslovii ekstremuma i zadachi optimalnogo upravlenia*, Uspehi Matem.nauk, T.25, N°5(1970).

[10] Neustadt L.W., *An abstract variational theory with applications to a broad class of optimization problems*, SIAM J. of Control 5(1)(1967).

[11] Watkins G.G., *Nonsmooth Milyutin-Dubovitskii theory and Clarke's tangent cone*, Math. of Op. R. 11, N°1 (1986).

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 229-236

DISCRETIZATION AND THE NUMERICAL SOLUTION OF NONLINEAR INTEGRAL EQUATIONS

Dušan TOŠIĆ
Faculty of Mathematics, University of Belgrade, Studentski trg 16,

Nikola PERIN
Faculty of Civil Engineering, University of Belgrade,
Bulevar revolucije 73,
11000 Belgrade, Yugoslavia

Abstract. The problem of finding a numerical solution for the nonlinear integral equations

$$y(x) - \int_a^x K[x, s, y(s)]ds = f(x) \tag{1}$$

and

$$y(x) - \int_a^x K(x, s)F[y(s)]ds = f(x) \tag{2}$$

is considered. Two methods for the discretization of (1) and (2) are proposed. The first method is based on numerical integration and interpolation. In the second method, the Euler-McLaurin formula and numerical derivation are used. The numerical results obtained using these methods are presented and compared.

Keywords: discretization, numerical solution, nonlinear, integral equations.

1. INTRODUCTION

Integral equations have a significant role in practice. A lot of problems from mechanics, physics, aerodynamics, astronomy, etc. (see [1] and [4]) can be resolved by solving the corresponding integral equations.

We consider nonlinear integral equations in the following forms

$$y(x) = f(x) + \int_a^x K[x, s, y(s)]ds \tag{3}$$

$$y(x) = f(x) + \int_a^x K(x, s)F[y(s)]ds \tag{4}$$

and look for the solution $y(x)$ on the interval:

$$a \le x \le b, \quad (a, b \in R).$$

We assume that all functions in (3) and (4) are continuous in the corresponding domains. Our consideration could be applied to other forms of integral equations, especially, if it is possible to express $y(x)$ as in the following example:

$$y(x) = \Phi\left(x, \int_a^x K(x, s)y(s)ds\right),$$

but in these cases an additional examination about the existence and uniqueness of solution $y(x)$ is necessary.

For integral equations (3) and (4) we assume that the following conditions hold

$$\begin{aligned} &|K(x, s, z_1) - K(x, s, z_2)| \le m(x, s)|z_1 - z_2|, \\ &\left|\int_a^x K(x, s, f(s))ds\right| \le n(x) \\ &\int_a^x n^2(s)ds \le N^2, \quad \int_a^b dx \int_a^x m^2(x, s)ds \le M^2, \quad a \le s \le x \le b \end{aligned} \tag{5}$$

where

- N and M are real constants, and
- $n(x)$ and $m(x, s)$ are the integrable functions with their squares.

In [4] it is quoted that if the conditions (5) are satisfied, then the existence and uniqueness of solution $y(x)$ are provided for equation (3) (but we will not consider these problems). In addition, conditions (5) are sufficient for the convergence (see [4]) of the iterative processes described by the formulas

$$(6) \qquad \begin{aligned} y_{k+1}(x) &= f(x) + \int_a^x K(x,s,y_k(s))ds, \\ y_{k+1}(x) &= f(x) + \int_a^x K(x,s)F(y_k(s))ds. \end{aligned} \qquad k = 0, 1, 2, \ldots$$

if the initial approximation $y_0(x)$ is known. (The initial approximation is an arbitrary continuous function, but usually $y_0(x) = f(x)$ is taken.)

Of course, all our considerations hold for the most popular integral equations - linear integral equations. If the function $F(y(s))$ in (4) is $y(s)$, we get a linear integral equation.

2. DISCRETIZATION

In most cases, it is not possible to find the solution of (3) and (4) in an analytical form. Because of that, different kinds of numerical methods are used.

The iterative process described by (6) ensures a way to find approximate solutions. Moreover, if this process is used in the analytical form, even during the calculation of the first or the second iteration, the analytical expressions obtained by (6) usually become too complex. Therefore, discretization is necessary.

The first step in discretization is to discrete the problem. Therefore, we are looking for the approximate solution $y(x)$ of (3) and (4) in the discrete point set:

$$G_n = \{x_i \mid a = x_0 < x_1 < \cdots < x_n = b\}$$

We use the iterative process (6) to find an approximate solution and the main problem encountered is to select a way of discretization for this process. The key action in discretization is calculating the values of function $W(x, s, y(s))$ and its integral:

$$(7) \qquad \int_a^x W(x,s,y(s))ds$$

The form of function $W(x, s, y(s))$ depends on the equation used in (6). We propose two approaches in the calculation of (7) and we practically create two methods for the discretization of (6). In (7) x is a parameter and in the process of calculating, x always has concrete value.

Approach I. This approach is based on numerical integration and interpolation. Approach I is firstly applied in [3] for the discretization of Chaplygin's method. Later we recognize that this approach is more general and could be used for the discretization of different analytical methods.

In each iteration $k(k = 0, 1, 2, ...)$ we do the following steps:

Step 1. $y_k(x)$ is replaced by an associated interpolation function $S_{k,n}(x)$.

Step 2. Discretization of $W(x, s, y_k(s))$ is made by calculating its values on G_n.

Step 3. Integrand $W(x, s, y_k(s))$ is replaced with $W_n(x, s, S_{k,n}(s))$ using the same interpolation function as in Step 1.

Step 4. Some of the formulas for numerical integration are applied in the computing of (7).

After the computation of (7) it is easy to find the approximate value for y(x).

The variety of numerical integration formulas used in Step 4 and interpolation formulas used in Step 1 generates different schemes for the solution of (4) and (5).

Approach II. This approach is based on the Euler-McLaurin formula and numerical derivations. It is firstly applied for the discretization of Chaplygins method too ([2]), and later in other analytical methods. An equidistant grid $G_n = G_h$ (the steplength h is constant, $h = \dfrac{(b-a)}{n}$, $x_i = a + ih$, $i = 0, ..., n$) is assumed.

According to the Euler-McLaurin formula we have:

$$
\begin{aligned}
(8) \qquad \int_a^x W(x, s, y(s))ds \approx \frac{1}{2}[W_0(x) + 2\sum_{j=1}^{i-1} W_j(x) + W_i(x)] - \\
- \sum_{j=1}^{m} \frac{B_{2j}}{(2j)!} h^{2j} (W^{(2j-1)}(x, x_i, y(x_i)) - \\
- W^{(2j-1)}(a, a, y(a)))
\end{aligned}
$$

where $W_i(x) = W(x, x_i, y(x_i))$, $i = 0, 1, ...n$. B_k are Bernoulli's numbers, implicitly defined by

$$
(9) \qquad \frac{t}{e^t - 1} = \sum_{k=0}^{\infty} B_k \frac{t^k}{k!}
$$

and from (9) we have:

$$B_0 = 1,\ B_1 = -\frac{1}{2},\ B_2 = \frac{1}{6},\ B_4 = -\frac{1}{30},\ B_6 = \frac{1}{42}, B_8 = -\frac{1}{30}, \ldots,\ B_{2k+1} = 0,\ k \geq 1.$$

To apply formula (8), it is necessary to know the values of derivatives of the function $W(x, s, y(s))$ on G_h. The derivatives are approximated by finite differences using the values of $W(x, s, y(s))$ on G_h. A global description of the algorithm for this discretization approach is shown by the following steps:

Step 1. Compute the approximate value of integral (7) using the trapezoidal rule.

Step 2. Compute the values of $W(x, s, y(s))$ on G_h.

Step 3. Approximate the derivatives of $W(x, s, y(s))$ using finite differences.

Step 4. Improve the value of integral (7) using formula (8).

In the process of programming, different modifications of this algorithm are possible. For example, using the derivatives of different orders in (9), we could approximately estimate the error in the obtained values on G_h.

3. NUMERICAL EXAMPLES

We will now present the numerical results obtained using both of the described approaches to discretization. Special software, developed in the programming language Pascal, is applied for this purpose. The program stops when two consequtive iterates have the same value. All examples of the integral equations considered here are taken from [4].

Example 1. Let us have the integral equation:

$$y(x) = \int_0^x \frac{1 + y^2(s)}{1 + s^2} ds \ \text{ with the exact solution } y^E(x) = x$$

If we apply Approach I, utilizing the Gaussian integration method with 7 nodes and cubic spline interpolation, for $h = 1.0$, $n = 10$, $y_0 = 0$, after 24 iterations we have:

x	$y(x)$	$y^E(x)$
1.00	1.0000000010	1.0000000000
2.00	2.0000000032	2.0000000000
3.00	3.0000000044	3.0000000000
4.00	4.0000000117	4.0000000000
5.00	5.0000000226	5.0000000000
6.00	6.0000000223	6.0000000000
7.00	6.9999999901	7.0000000000
8.00	7.9999999040	8.0000000000
9.00	8.9999997447	9.0000000000
10.00	9.9999994972	10.0000000000

Example 2. For the integral equation:

$$y(x) = 1 + 2\int_1^x \sqrt{y(s)}ds \quad \text{with the exact solution} \quad y^E(x) = x^2$$

we will also apply Approach I. Using the Lagrange interpolation and the Simpson rule for $h = 0.1$, $n = 10$, $y_0 = 1$, after 10 iterations we get:

x	$y(x)$	$y^E(x)$
1.10	1.2100000000	1.2100000000
1.20	1.4400000000	1.4400000000
1.30	1.6900000000	1.6900000000
1.40	1.9600000000	1.9600000000
1.50	2.2500000000	2.2500000000
1.60	2.5600000000	2.5600000000
1.70	2.8900000000	2.8900000000
1.80	3.2399999998	3.2400000000
1.90	3.6099999996	3.6100000000
2.00	3.9999999989	4.0000000000

Example 3. For the same problem as in example 2, Approach II is applied, using finite differences of order 5 for the approximation of necessary derivatives ($m = 5$). The following numerical results are obtained:

x	$y(x)$	$y^E(x)$
1.10	1.2100000000	1.2100000000
1.20	1.4400000000	1.4400000000
1.30	1.6900000000	1.6900000000
1.40	1.9600000000	1.9600000000
1.50	2.2500000000	2.2500000000
1.60	2.5600000000	2.5600000000
1.70	2.8899999999	2.8900000000
1.80	3.2399999998	3.2400000000
1.90	3.6099999995	3.6100000000
2.00	3.9999999989	4.0000000000

Example 4. The integral equation

$y(x) = \dfrac{1}{\cos x}\displaystyle\int_0^x \cos^2 s y^2(s) ds + \dfrac{1}{\cos x}$ has the exact solution $y^E(x) =$

$$-\frac{1}{((x-1)\cos x)}.$$

Using Approach II, (using the same finite differences as in the previous example) for $h = 0.05, n = 16$, after 15 iterations, we have:

x	$y(x)$	$yE(x)$
0.05	1.0539512026	1.0539487404
0.10	1.1166926334	1.1166899093
0.15	1.1898341759	1.1898311090
0.20	1.2754270512	1.2754235562
0.25	1.3761173876	1.3761133653
0.30	1.4953641150	1.4953594308
0.35	1.6377596540	1.6377541283
0.40	1.8095140011	1.8095073806
0.45	2.0192069735	2.0191989218
0.50	2.2789979735	2.2789878547
0.55	2.6066502542	2.6066373045
0.60	3.0290930323	3.0290707863
0.65	3.5890144004	3.5889976185
0.70	4.3582432279	4.3581975324
0.75	5.4668205170	5.4668044987
0.80	7.1768332073	7.1766209984

4. CONCLUSION

Analyzing the obtained numerical values we can conclude that both approaches to discretization give quite good results. Moreover, further theoretical study of the error estimation is necessary. Some considerations related to error estimation for the iterative process (6), especially for the linear case, can be found in [1] and [4]. Additional analysis is essential for successful application of the presented discretization approaches.

REFERENCES

[1] Noble B., *The numerical solution of integral equation*, in: D.A.H. Jacobs (ed.) *The State of the Art, in: Numerical Analysis*, Academic Press, 1977, 959–966.

[2] Perin N., Tošić D., *A new approach to discretization of Chaplygin's method*, Acta Math. Hungarica, 74 (1-2) (1996), 31-40 (forthcoming).

[3] Tošić D., *One way of discretization of Chaplygin's method*, in: G. Milovanović (ed.), *Proc. of Numerical methods and approximation theory*, Niš, 1984, 149–153.

[4] Verlan A.F., Sizikov V.S., *Metodi rešenija integralnih uravnenii programami dlja EVM*, Kiev, Naukova dumka, 1978.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 237-247

ENUMERATION OF LABELLED INITIALLY-FINALLY CONNECTED DIGRAPHS

V. JOVOVIĆ, G. KILIBARDA
Faculty of Technology and Metallurgy, University of Belgrade, Karnegijeva 4, P.O. Box 494, 11001 Belgrade, Yugoslavia

Abstract. The class of all labelled initially-finally connected digraphs with a fixed number of points (and arcs) is counted and corresponding numerical data are given.

Keywords: Enumeration, labelled digraphs, initially-finally connected digraphs

1. INTRODUCTION

In the literature, for example in [1], several classes of connected digraphs are defined. A digraph is *weakly connected* if, by ignoring all orientations of its arcs, we get a connected multigraph. A digraph is called *strongly connected* iff every pair of its points is mutually accessible (point v is *accessible* from point u if there exists a directed path from point u to point v). A digraph is *unilaterally connected* iff for every two of its points at least one is accessible from the other. For *initially connected* digraphs there is a point (at least one) called *a source,* such that every point is accessible from this point. Similarly, a digraph is *finally connected* if there exists a point (at least one) called a *sink*, which is accessible from every point. The class of *initially-finally connected* digraphs is naturally defined, too, namely, as a class of digraphs with at least one source and one sink. In the literature we can find the term "digraph with a source (sink)" instead of the term "initially (finally) connected digraph", also we can say a "digraph with a source and a sink" instead of an "initially-finally connected digraph". Here we follow the notation of Liskovec (for example [6]).

In his list of unsolved graph enumeration problems Harary [1] mentioned the problem of enumerating the above-defined classes of connected digraphs with n points, both for labelled and unlabelled digraphs. Note that Harary did not mention (we don't know why) certain results of Liskovec referring to these classes. Liskovec [2] counted strongly connected digraphs, and his formulas were simplified by Wright [3]. The case of unlabelled strongly connected digraphs, and the case of unlabelled rooted digraphs with a source were also solved by Liskovec [4–7]. In [8] Robinson gave a combinatorial proof of the simplified Liskovec-Wright formulas and counted labelled initially and unilaterally connected digraphs. Announcing the results of his research Robinson [9] outlined a way to count the class of all unlabelled strongly connected, initially and unilaterally connected digraphs with n points.

In this paper the problem of counting initially-finally connected labelled digraphs is solved, and the already known results for the number of strongly connected and the number of initially connected digraphs were obtained as intermediate results.

2. THE MAIN NOTIONS AND AUXILLIARY RESULTS

In what follows all digraphs will be labelled digraphs, if not stated otherwise.

Let $\mathcal{G}_0, \mathcal{G}_1, \ldots, \mathcal{G}_n$ be some finite classes of digraphs. We say that $\mathcal{G}_0$ is of the type $\{\mathcal{G}_1, \mathcal{G}_2, \ldots \mathcal{G}_n\}$ if:

1) every digraph G_0 from $\mathcal{G}_0$ can be represented in the form $G_1 \cup G_2 \cup \ldots \cup G_n$, where $G_i \in \mathcal{G}_i$ for each $i = 1, 2, \ldots, n$;
2) for each collection of digraphs $G_i \in \mathcal{G}_i$, $i = 1, 2, \ldots, n$, the digraph $G_1 \cup G_2 \cup \ldots \cup G_n$, belongs to class $\mathcal{G}_0$;
3) for two different collections of digraphs, $G_i \in \mathcal{G}_i$ and $G_i' \in \mathcal{G}_i$, $i = 1, 2, \ldots, n$, the digraphs $G_1 \cup G_2 \cup \ldots \cup G_n$ and $G_1' \cup G_2' \cup \ldots \cup G_n'$ are different.

(Here the operation $\cup$ is defined in such a way that for every two digraphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ the digraph $G_1 \cup G_2$ is a digraph with the set of points $V_1 \cup V_2$ and the set of arcs $E_1 \cup E_2$). Then it is clear that the equation holds:

$$|\mathcal{G}_0| = |\mathcal{G}_1| \cdot |\mathcal{G}_2| \cdot \ldots \cdot |\mathcal{G}_n|.$$

Let us denote by O_n an empty digraph of the order n (without arcs and with n points). We denote by $\mathcal{B}_{m,n}$ the class of all bipartite digraphs whose parts contain m and n points respectively, and every arc of these digraphs has its beginning in the first and its end in the second part. We take that $O_{m+n} \in \mathcal{B}_{m,n}$. Let us designate by $\mathcal{D}_n$ the class of all digraphs with n points.

Let $G = (V, E)$ be a digraph and $U, W \subseteq V$. Denote by $\mathbf{Acc}(U, W; G)$ the maximum subset of the set W, such that for every $v \in \mathbf{Acc}(U, W; G)$ there is $u \in U$, such that in digraph G there exists a directed path from u to v, i.e., point v is *accessible* from point u or, which is the same, point u is *antiaccessible* from point v. Let us assume here that $U \cap W \subseteq \mathbf{Acc}(U, W; G)$ (we also take into account the paths of length 0). Also, designate by $\overline{\mathbf{Acc}}(U, W; G)$ the set of all points from W antiaccessible from some point from U.

Let arbitrary sets V_1 and V_2, $V_1 \cap V_2 = \emptyset$, be given, where $|V_1| = m$ and $|V_2| = n$, $m \in \mathbf{N}$, $n \in \mathbf{N} \cup \{0\}$. Let $G = (V, E)$, where $V = V_1 \cup V_2$, be an arbitrary digraph such that $E(V_1) = (V_1 \times V_1) \cap E = \emptyset$, $\mathbf{Acc}(V_1, V_2; G) = V_2$ and $(V_2 \times V_1) \cap E = \emptyset$. We denote the set of all such digraphs by $\Lambda_{V_1}(V_2)$; if it is clear from the context which sets V_1 and V_2 are meant, then instead of $\Lambda_{V_1}(V_2)$ we use the notation $\Lambda_m(n)$. It is clear that the number of elements in $\Lambda_{V_1}(V_2)$ does not depend on the sets V_1 and V_2, but only on their cardinality. We denote $\lambda_m(n) = |\Lambda_m(n)|$.

Lemma 1. *For every $m \in \mathbf{N}$ the recurrent formula*

$$\lambda_m(n) = 2^{n(n+m-1)} - \sum_{k=0}^{n-1} C_n^k \, 2^{(n-k)(n-1)} \, \lambda_m(k)$$

holds, with the initial condition $\lambda_m(0) = 1$.

Proof. Let us fix the sets V_1 and V_2, $V_1 \cap V_2 = \emptyset$, so that $|V_1| = m$ and $|V_2| = n$, $m \in \mathbf{N}$, $n \in \mathbf{N} \cup \{0\}$. Let $\tilde{\Lambda}_m(n)$ be the set of all digraphs $G = (V, E)$, where $V = V_1 \cup V_2$, such that $E(V_1) = \emptyset$ and $(V_2 \times V_1) \cap E = \emptyset$. Denote $\tilde{\lambda}_m(n) = |\tilde{\Lambda}_m(n)|$. Also let $V_3(G) = \mathbf{Acc}(V_1, V_2; G)$ for arbitrary $G \in \tilde{\Lambda}_m(n)$. Denote by $\tilde{\Lambda}_m(n, k)$ the class of all G from $\tilde{\Lambda}_m(n)$, such that $|V_3(G)| = k$, and let $\tilde{\lambda}_m(n, k) = |\tilde{\Lambda}_m(n, k)|$. Then

$$\lambda_m(n) = \tilde{\lambda}_m(n) - \sum_{k=0}^{n-1} C_n^k \, \tilde{\lambda}_m(n; k). \tag{1}$$

Let us calculate the number $\tilde{\lambda}_m(n, k)$. Let $E_1 = [(V_2 \backslash V_3) \times V_3] \cap E$. It is obvious that $[V_3 \times (V_2 \backslash V_3)] \cap E = \emptyset$ and E_1 can be an arbitrary subset of the set $(V_2 \backslash V_3) \times V_3$. It is easy to see that the digraphs $G_1 = (V_1 \cup V_3, E(V_1 \cup V_3))$, $G_2 = (V_2 \backslash V_3, E(V_2 \backslash V_3))$ and $G_3 = (V_2, E_1)$ are from the classes $\Lambda_m(k)$, $\mathcal{D}_{n-k}$ and $\mathcal{B}_{n-k,k}$ respectively. It is clear that class $\tilde{\Lambda}_m(n, k)$ is of the type $\{\Lambda_m(k), \mathcal{D}_{n-k}, \mathcal{B}_{n-k,k}\}$. This means that

$$\tilde{\lambda}_m(n, k) = \lambda_m(k) \, 2^{(n-k)(n-k-1)} \, 2^{(n-k)k} = 2^{(n-k)(n-1)} \, \lambda_m(k). \tag{2}$$

As it is obvious that

$$\tilde{\lambda}_m(n) = 2^{n(n-1)}\, 2^{mn} = 2^{n(n+m-1)},$$

then the statement of the lemma follows from (1) and (2).

Let arbitrary disjoint sets V_1, V_2 and V_3 be given, where $|V_1| = k$, $|V_2| = m$ and $|V_3| = n$, $k, n \in \mathbf{N}$, $m \in \mathbf{N} \cup \{0\}$. Let $\Lambda_{V_1,V_3}(V_2)$ be a class of all digraphs $G = (V, E)$, where $V = V_1 \cup V_2 \cup V_3$, such that $E(V_1) = E(V_3) = \emptyset$, every point from the set V_2 is accessible from some point from the set V_1 and antiaccessible from some point from the set V_3, and $[(V_2 \cup V_3) \times V_1] \cap E = (V_3 \times V_2) \cap E = (V_1 \times V_3) \cap E = \emptyset$. If it is clear from the context what sets V_1, V_2 and V_3 are meant, then instead of the notation $\Lambda_{V_1,V_3}(V_2)$ we use the notation $\Lambda_{k,n}(m)$. It is clear that the number of elements of the set $\Lambda_{V_1,V_3}(V_2)$ depends only on the cardinality of the sets V_1, V_2 and V_3. Therefore we denote $\lambda_{k,n}(m) = |\Lambda_{k,n}(m)|$.

Lemma 2. *For every $k, n \in \mathbf{N}$, the recurrent formula*

$$\lambda_{k,n}(m) = \lambda_k(m)\, 2^{mn} - \sum_{i=0}^{m-1} C_m^i\, \lambda_{k+i}(m-i)\, \lambda_{k,n}(i)$$

holds, with the initial condition $\lambda_{k,n}(0) = 1$.

Proof. Let us fix some disjoint finite sets V_1, V_2 and V_3, where $|V_1| = k$, $|V_2| = m$ and $|V_3| = n$, $k, n \in \mathbf{N}$, $m \in \mathbf{N} \cup \{0\}$. We denote by $\tilde{\Lambda}_{k,n}(m)$ the set of all digraphs $G = (V, E)$, $V = V_1 \cup V_2 \cup V_3$, such that $E(V_1) = E(V_3) = \emptyset$, $V_2 = \mathbf{Acc}(V_1, V_2; G)$ and $[(V_2 \cup V_3) \times V_1] \cap E = (V_3 \times V_2) \cap E = (V_1 \times V_3) \cap E = \emptyset$. Also let $V_2'(G) = V_2 \cap \overline{\mathbf{Acc}}(V_3, V_2; G)$ for every $G \in \tilde{\Lambda}_{k,n}(m)$. We denote by $\tilde{\Lambda}_{k,n}(m; i)$ the class of all G from $\tilde{\Lambda}_{k,n}(m)$, such that $|V_2'(G)| = i$. It is clear that $|\tilde{\Lambda}_{k,n}(m)| = \lambda_k(m)\, 2^{mn}$. Then

$$\lambda_{k,n}(m) = |\tilde{\Lambda}_{k,n}(m)| - \sum_{i=0}^{m-1} C_m^i\, |\tilde{\Lambda}_{k,n}(m; i)| = \tag{3}$$

$$= \lambda_k(m)\, 2^{mn} - \sum_{i=0}^{m-1} C_m^i\, |\tilde{\Lambda}_{k,n}(m; i)|.$$

Let us calculate the number $|\tilde{\Lambda}_{k,n}(m; i)|$. It is clear that $[(V_2 \backslash V_2') \times V_2'] \cap E = \emptyset$. Let us consider the digraphs $G_1 = (V_1 \cup V_2, E(V_1 \cup V_2) \backslash E(V_1 \cup V_2'))$ and $G_2 = (V_1 \cup V_2 \cup V_3, E \backslash [E(V_1 \cup V_2) \backslash E(V_1 \cup V_2')])$. It is obvious that the digraphs G_1 and G_2 belong to classes $\Lambda_{k+i}(m-i)$ and $\Lambda_{k,n}(i)$ respectively, and that the class of digraphs $\tilde{\Lambda}_{k,n}(m; i)$ is of the type $\{\Lambda_{k+i}(m-i), \Lambda_{k,n}(i)\}$. This means that $|\tilde{\Lambda}_{k,n}(m; i)| = \lambda_{k+i}(m-i)\, \lambda_{k,n}(i)$. Then the statement of the lemma follows from (3).

3. INITIALLY-FINALLY CONNECTED DIGRAPHS

Let $G = (V, E)$ be some digraph. We say that G is an *initially connected* digraph if there is $u \in V$, such that $\mathbf{Acc}(\{u\}, V; G) = V$; such point u is called a *source*. We say that G is a *strongly connected* digraph, if for every $(u, v) \in V^2$ there is a directed path in G from u to v, that is all the points in the digraph are sources. We denote by $\tilde{I}(n)$ the class of all initially connected digraphs with n points and with fixed sources, and let $\tilde{i}(n) = |\tilde{I}(n)|$. Also let us denote by $Q(n)$ the class of all strongly connected digraphs with n points, and let $q(n) = |Q(n)|$. From the above it is easy to see that $\tilde{i}(n) = 2^{n-1}\,\lambda_1(n-1)$, that is, that the recurrent formula

$$\tilde{i}(n) = 2^{n(n-1)} - \sum_{j=1}^{n-1} C_{n-1}^{j-1}\, 2^{(n-j)(n-1)}\, \tilde{i}(j). \tag{4}$$

holds.

The result for the class of strongly connected digraphs, which was first proved in [2] immediately follows from the above given lemmas and from (4).

Theorem 1. *For the number $q(n)$ of strongly connected digraphs with n points the following recurrent formula*

$$q(n) = \tilde{i}(n) - \sum_{j=1}^{n-1} C_{n-1}^{j-1}\, \lambda_j(n-j)\, q(j)$$

holds, with the initial condition $q(0) = 1$.

Proof. It is obvious that $q(n) = \lambda_{1,1}(n-1)$ (we introduce one virtual point). By lemma 2 we have that

$$\lambda_{1,1}(n-1) = 2^{n-1}\,\lambda_1(n-1) - \sum_{i=0}^{n-2} C_{n-1}^{i}\, \lambda_{i+1}(n-i-1)\, \lambda_{1,1}(i).$$

Using the relation $\tilde{i}(n) = 2^{n-1}\,\lambda_1(n-1)$ we get the statement of the theorem.

Let $G = (V, E)$ be some digraph, and $l \le |V|$. We say that G is *l-initially (l-finally) connected* if:

1) there exists $V_1 \subseteq V$, $|V_1| = l$, such that for every $v \in V$ and for every $u \in V_1$ there is a directed path in G from u to v (a path from v to u), that is $\mathbf{Acc}(\{u\}, V; G) = V$ ($\overline{\mathbf{Acc}}(\{u\}, V; G) = V$) for every $u \in V_1$;
2) there does not exist a subset of V which has more than l elements and has the property of the set V_1 from 1.

We call the points from the set V_1 *sources* (*sinks*). In other words, the digraph G is *l-initially* (*l-finally*) connected if it has exactly l sources (sinks). We denote the class of all l-initially (l-finally) connected digraphs with n points by $I_l(n)$ ($F_l(n)$). Digraphs with n points which contain at least one source (sink) constitute class $I(n)$ ($F(n)$) of all initially (finally) connected digraphs. Let us also consider the class $IF_{k,l}(n)$, $k, l \in \mathrm{N}$, $k + l \leq n$, of all digraphs with n points, which have exactly k sources and exactly l sinks. It is clear that if the intersection of the set of all sources and of the set of all sinks in a digraph is not empty, then it is strongly connected. We denote by $IF(n)$ the class of all digraphs with n points which have at least one source and at least one sink. Let $i_l(n) = |I_l(n)|$, $i(n) = |I(n)|$, $f_l(n) = |F_l(n)|$, $f(n) = |F(n)|$, $if_{k,l}(n) = |IF_{k,l}(n)|$ and $if(n) = |IF(n)|$. It is obvious that the following formulas hold:

$$i(n) = \sum_{j=1}^{n} i_j(n),$$

$$f(n) = \sum_{j=1}^{n} f_j(n),$$

$$if(n) = q(n) + \sum_{i=1}^{n-1} \sum_{j=1}^{n-i} if_{i,j}(n).$$

Let

$$\overline{\lambda}_{i,j}(k) = \begin{cases} 2^{ij} \lambda_{i,j}(k), & k \neq 0, \\ 2^{ij} - 1, & k = 0, \end{cases} \tag{5}$$

for abitrary $i, j \in \mathrm{N}$, $k \in \mathrm{N} \cup \{0\}$. Let us prove the following statement.

Theorem 2. *For the number $i_k(n)$ ($f_k(n)$) of all l-initially* (l-finally) *connected digraphs with n points and the number $if_{k,l}(n)$ of all initially-finally connected digraphs with n points which have exactly k sources and exactly l sinks, the following formulas hold:*

$$i_k(n) = f_k(n) = C_n^k \, \lambda_k(n-k) \, q(k),$$
$$if_{k,l}(n) = C_n^k C_{n-k}^l \, q(k) \, q(l) \, \overline{\lambda}_{k,l}(n-k-l).$$

for abitrary $k, l, n \in \mathrm{N}$, $k, k + l \leq n$,

Proof. We shall prove the second formula (the first formula is similarly obtained). Firstly, let us note that the set of all sources (sinks) in an arbitrary digraph determines a digraph that is strongly connected, that is, the set of all sources (sinks) is its component of strong connectedness.

Now, let $G = (V, E)$ be a digraph from class $IF_{k,l}(n)$. Let us denote the sets of all its k sources and l sinks by V_1 and V_3 respectively. It is clear that $V_1 \cap V_2 = \emptyset$ (otherwise G would be a strongly connected digraph). Let $V_2 = V \backslash (V_1 \cup V_3)$. We can assume that $V_2 \neq \emptyset$. It is clear that

$$
\begin{aligned}
&G_1 = (V_1, (V_1 \times V_1) \cap E) \in Q(k), \\
&G_2 = (V_3, (V_3 \times V_3) \cap E) \in Q(l), \\
&G_3 = (V, E \backslash [(V_1 \times V_1) \cup (V_3 \times V_3) \cup (V_1 \times V_3)]) \in \Lambda_{k,l}(n-k-l), \\
&G_4 = (V_1 \cup V_3, (V_1 \times V_3) \cap E) \in \mathcal{B}_{k,l},
\end{aligned}
$$

and that the class $IF_{k,l}(n)$ is of the type $\{Q(k), Q(l), \Lambda_{k,l}(n-k-l), \mathcal{B}_{k,l}\}$. So, the number of all digraphs from $IF_{k,l}(n)$ with fixed sets of sources V_1 and sinks V_3 is $q(k)\, q(l)\, \overline{\lambda}_{k,l}(n-k-l)$ (if $V_2 = \emptyset$, as $(V_1 \times V_3) \cap E \neq \emptyset$, then by (5) we get that the same relation holds). As the choice of disjoint subsets V_1 and V_3 of the set V is arbitrary, we have the statement of the theorem.

4. TABLES

Using the obtained recurrent formulas, the cardinalities of all the above-mentioned classes of connected digraphs with n points, $n \leq 30$, were calculated on a computer. Here we give only the number of all initially-finally connected digraphs with n points, $n \leq 20$. Our calculations do not agree with those in [8] for the cases of strongly connected, unilaterally connected and initially connected digraphs for $n \geq 7$.

From the above-given formulas it is easy to obtain the formulas for the number of all initially-finally connected digraphs with a fixed number of arcs.

The third table gives the numbers of all nonisomorphic initially-finally connected digraphs (the problem of their enumeration has not yet been solved).

The number of all initially-finally connected digraphs with $n \leq 20$ points

1	1
2	3
3	48
4	3424
5	962020
6	1037312116
7	4344821892264
8	71771421308713624
9	4716467927380427847264
10	1237465168798883061207535456
11	1297923989772809185944542332007104
12	5444330658513426322624322033259452670016
13	91342931436147421630261703458729460990513248512
14	6129964251882515261054472090251920309790698538573860096
15	1645503177733288384586116507920619235642546108359341100411780608
16	176684664347284973763111082742401944150484468391798627637019210845 9880960
17	758854984960124422441257112611398737819231881444704038505491778861 2189218403938304
18	130370300026622248599703135686218619461581085131225526813568259651 926175542583586025687248896
19	895897892164657352217146580173273344229353550422221277586744161882 7352599603435356742676315502186192896
20	246262538369102439628535244312743522999462081332830098063375330217 7597034302014936247187932268733739585811672375296

The number of all initially-finally connected digraphs with l arcs and with n points, $n \leq 7$

$l \backslash n$	1	2	3	4	5	6	7
0	1	0	0	0	0	0	0
1		2	0	0	0	0	0
2		1	6	0	0	0	0
3			20	24	0	0	0
4			15	234	120	0	0
5			6	672	2544	720	0
6			1	908	16880	28200	5040
7				792	55000	351600	328320
8				495	111225	2185860	6830460
9				220	161660	8410500	71142540
10				66	183006	22861044	457789542
11				12	167660	47827680	2063736444
12				1	125945	81394740	7080369065
13					77520	116692320	19567339470
14					38760	143934660	45325126575
15					15504	154538696	90499079832
16					4845	145245435	158841949245
17					1140	119718630	248399944842
18					190	86486385	349230046305
19					20	54626580	443990373120
20					1	30044979	512277729264
21						14307150	537543858312
22						5852925	513501619764
23						2035800	446674955160
24						593775	353667944665
25						142506	254654937306
26						27405	166508378463
27						4060	98672228676
28						435	52860207765
29						30	25518729810
30						1	11058116839
31							4280561376
32							1471442973
33							445891810
34							118030185
35							26978328
36							5245786
37							850668
38							111930
39							11480
40							861
41							42
42							1
Σ	1	3	48	3424	962020	1037312116	4344821892264

The number of all nonisomorphic initially-finally connected digraphs with l arcs and with n points, $n \leq 5$

$l \backslash n$	1	2	3	4	5
0	1	0	0	0	0
1		1	0	0	0
2		1	1	0	0
3			4	1	0
4			4	11	1
5			1	31	23
6			1	45	152
7				38	486
8				27	992
9				13	1419
10				5	1641
11				1	1485
12				1	1152
13					707
14					379
15					154
16					61
17					16
18					5
19					1
20					1
Σ	1	2	11	173	8675

REFERENCES

[1] Harary, F., and Palmer, E.M., *Graphical Enumeration*, Academic Press, New York and London, 1973.

[2] Liskovec, V.A., "On a recurrent method for the enumeration of graphs with labelled vertices", *Soviet Math. Dokl.* 10 (1969) 242–246 [*Dokl. Akad. Nauk SSSR* 184 (1969) 1284–1287].

[3] Wright, E.M., "The number of strong digraphs", *Bull. London Math. Soc.* 3 (1971) 348–350.

[4] Liskovec, V.A., "Enumeration of rooted digraphs with a source" *Vesci Akad. Navuk B. SSR Ser. Fiz-Math. Navuk* (1969) 23–32 (in Russian).

[5] Liskovec, V.A., "The number of strongly connected oriented graphs" *Mat. Zametki* 8 (1970) 721–732 (in Russian).

[6] Liskovec, V.A., "The number of initially connected digraphs" *Dokl. Akad. Nauk. B. SSR* 15 (1971) 293–294 (in Russian).

[7] Liskovec, V.A., "Enumeration of rooted digraphs with a source II. The Redfield-Robinson method" *Vesci Akad. Navuk B. SSR Ser. Fiz-Mat. Navuk* (1973) 35–44, 136 (in Russian).

[8] Robinson, R.W., "Counting labeled acyclic digraphs", *New Directions in the Theory of Graphs*, Academic Press, New York, 1973, 239–273.

[9] Robinson, R.W., "Counting strong digraphs" (research announcement), *Journal of Graph Theory* 1 (1977) 189–190.

SCIENTIFIC REVIEW (1996), Number 19-20, pp. 249-264

CALCULATION OF THE FOURIER TRANSFORM ON FINITE NON-ABELIAN GROUPS THROUGH DECISION DIAGRAMS

Radomir S. STANKOVIĆ, Milena STANKOVIĆ, Dragan JANKOVIĆ
Faculty of Electrical Engineering, Beogradska 14,
18 000 Niš, Yugoslavia

Abstract. The paper presents a calculation procedure for the Fourier transform on finite non-Abelian groups through decision diagram representations of discrete functions. The procedure permits the processing of functions on large groups and thus removes the complexity of calculation as the chief limiting factor in applications of the Fourier transform on finite non-Abelian groups.

1. INTRODUCTION

The concept reported in mathematics as the Kurepa tree was introduced by Djuro R. Kurepa in his Ph.D. dissertation in 1935 [12]. In computer and information sciences, this concept can be recognized as a particular data structure denoted as the decision tree. After the publication of Bryant's paper [1], decision diagrams (DDs) derived by the reduction of the decision trees have been widely used for representations, manipulations and calculations with discrete functions. Reduction is performed by deleting the redundant nodes and sharing the equivalent sub-trees in the decision tree associated to a given function f.

Various DDs for switching [16], and discrete functions [2], [13], [17], [19], [20], are defined and applications are found, for example, in signal processing including logic design, linear programming, function decomposition, etc. DDs have proved very useful in spectral techniques considered to be part of abstract harmonic analysis devoted to applications mainly in electrical engineering. Thanks to DDs, applications of spectral techniques have been considerably extended to these areas where the complexity of calculating discrete transforms was the chief limiting factor. Regarding digital signals[1], calculation algorithms for discrete transforms though DDs proposed first in [3] have an impact comparable to that of the fast Fourier transform (FFT) [4] in the processing of discrete signals.

The algorithm suggested in [3] for the Walsh and Reed-Muller transforms was extended to various discrete transforms [13], [14], [22]. However, all these results concern signals described by functions on finite Abelian groups.

Among different discrete transforms, the Fourier transform on finite non-Abelian groups is recommended as the best choice in some particular applications [10], [9], [11], [24]. Therefore, in this paper we extend the theory of calculation algorithms for discrete transforms through DDs to functions on finite non-Abelian groups.

In Sections 2 and 3 we explain and illustrate by example the representation of discrete functions by DDs and briefly repeat some basic definitions of the Fourier transform on finite non-Abelian groups. Using the matrix interpretation of this transform in Section 4, we define the calculation procedure for the Fourier transform on finite non-Abelian groups in Section 5.

2. DECISION DIAGRAMS

Let G be a finite, not necessarily Abelian, group G of order g. We associate permanently and bijectively with each group element a non-negative integer from the set $\{0, 1, \ldots, g-1\}$, providing that 0 is associated with the group identity. In what follows, each group element will be identified with the fixed non-negative integer associated with it and with no other element. We assume that G can be represented as a direct product of subgroups $G_1, \ldots, G_n$ of orders $g_1, \ldots, g_n$, respectively, i.e.,

$$G = \times_{i=1}^{n} G_i, \quad g = \Pi_{i=1}^{n} g_i, \quad g_1 \leq g_2 \leq \ldots \leq g_n. \tag{1}$$

The convention adopted above for the denotation of group elements applies to the subgroups G_i as well. Provided that the notational bijections of the subgroups G_i and of G are consistently chosen, each $x \in G$ can be uniquely represented as

$$x = \sum_{i=1}^{n} a_i x_i, \quad x_i \in G_i, \quad x \in G, \tag{2}$$

[1]Digital signals are discrete signals whose amplitudes take values in finite sets.

with

$$a_i = \begin{cases} \Pi_{j=i+1}^{n} g_j, & i = 1, \ldots, n-1, \\ 1, & i = n, \end{cases}$$

where g_j is the order of G_j.

The group operation $\circ$ of G can be expressed in terms of the group operations $\dot{i}$ of the subgroups G_i, $i = 1, \ldots, n$ by:

$$x \circ y = (x_1 \,\dot{1}\, y_1,\ x_2 \,\dot{2}\, y_2,\ \ldots\ x_n \,\dot{n}\, y_n),\ \ x, y \in G,\ \ x_i, y_i \in G_i.$$

Therefore, a function $f(x)$ on G can be alternatively considered as an n-variable function $f(x_1, \ldots, x_n)$, $x_i \in G_i$.

A given f on the decomposable group G of order g can be conveniently represented by a decision diagram with n levels consisting of nodes at the level i with g_i output edges. The nodes corresponding to the same variable x_i form the i-th level in the DD. The point where an edge connecting non-successive levels, i.e., the edge longer than 1, crosses a level is denoted as a cross point in the DD [20]. Such a DD is a straightforward generalization of the binary decision diagrams (BDDs) [16] and multi-terminal binary decision diagrams (MTBDDs) [3] defined originally on finite dyadic groups and multiple-place decision diagrams (MDDs) [17] on p-adic groups to functions on arbitrary finite groups.

We will explain and illustrate the decision diagram representation of discrete functions by the following example.

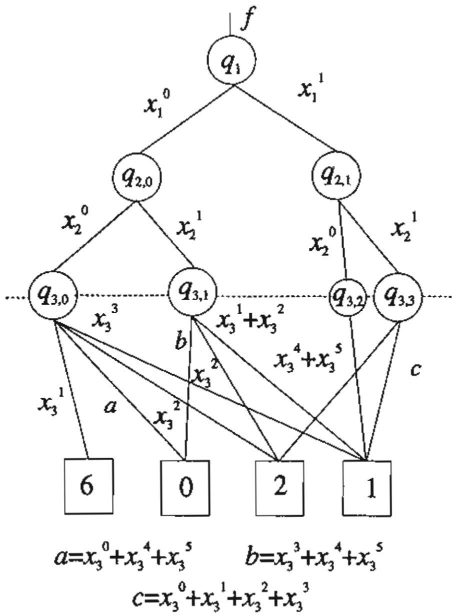
Example 1. *Let $G_{24} = C_2 \times C_2 \times S_3$, where $C_2 = (\{0,1\}, \oplus)$ be the basic cyclic group of order 2 with $\oplus$ denoting the modulo 2 addition and $S_3 = (0, (132), (123), (12), (13), (23), \circ)$ the symmetric group of permutations of order 3. Group elements of S_3 will be denoted by 0,1,2,3,4,5, respectively. Using this notation the group operation of S_3 is shown in Table 1. A function f on G_{24} given by the truth-vector*

$$\mathbf{f} = [0,6,2,1,0,0,2,1,1,0,0,0,1,1,1,1,1,1,1,1,1,1,2,2]^T,$$

can be represented by the generalized multi-terminal decision diagrams shown in Fig. 1. In this figure, x_i^j denotes that the variable x_i takes the value j.

Table 1: Group operation of S_3

$\circ$	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	0	5	3	4
2	2	0	1	4	5	3
3	3	4	5	0	1	2
4	4	5	3	2	0	1
5	5	3	4	1	2	0



DD representation of f in Example 1.

3. FOURIER TRANSFORM ON FINITE NON-ABELIAN GROUPS

Basic concepts of the Fourier transform on finite non-Abelian groups can be summarized as follows.

Any finite group is a compact group, so that all the results of Fourier analysis on compact groups are true for finite groups. In the case of finite groups

1. Every irreducible representation of a finite group G is equivalent to some unitary representation.

2. Every irreducible representation is finite dimensional.

3. The number of non-equivalent irreducible representations R_w of a finite non-Abelian group G of order g is equal to the number of equivalence classes of the dual object Γ of G. Denoting this number by K, it follows

$$\sum_{w=0}^{K-1} r_w^2 = g.$$

We will use the following notation to discuss the definition and properties of the Fourier transform on finite non-Abelian groups.

Let us denote by P the complex field or a finite field. Henceforth it will be assumed that:

1. char $P = 0$, or char P does not divide g.

2. P is a so-called splitting field for G.

Recall that a complex field is a splitting field for any finite group. We denote by $P(G)$ the space of functions f mapping G into P, i.e., $f : G \to P$.

Let K be the number of equivalence classes of irreducible representations of G over P. Each such equivalence class contains just one unitary representation. We shall denote the K unitary irreducible representations of G in some fixed order by $\mathbf{R}_0, \mathbf{R}_1, \ldots, \mathbf{R}_{K-1}$. We denote by $\mathbf{R}_w(x)$ the value of $\mathbf{R}_w$ at $x \in G$.

Note that $\mathbf{R}_w(x)$ stands for a non-singular r_w by r_w matrix, with elements $R_w^{(i,j)}(x)$, $i, j = 1, 2, \ldots, r_w$.

If the group G is representable in the form (1), then its unitary irreducible representations can be obtained as the Kronecker product of the unitary irreducible representations of subgroups G_i, $i = 1, \ldots, n$. Therefore, the number K of unitary irreducible representations of G can be expressed as:

$$K = \prod_{i=1}^{n} K_i, \tag{3}$$

where K_i is the number of unitary irreducible representations of the subgroup G_i.

Now, for a given group G of the form (1), the index w of each unitary irreducible representation $\mathbf{R}_w$ can be written as:

$$w = \sum_{i=1}^{n} b_i w_i, \quad w_i \in \{0, 1, \ldots, K_i - 1\}, \quad w \in \{0, 1, \ldots, K - 1\},$$

with

$$b_i = \begin{cases} \prod_{j=i+1}^{n} K_j, & i = 1, \ldots, n-1, \\ 1, & i = n, \end{cases} \tag{4}$$

where K_j is the number of unitary irreducible representations of the subgroup G_j.

The functions $R_w^{(i,j)}(x), \quad w = 0, 1, \ldots, K-1, \quad i, j = 1, \ldots, r_w$ form an orthogonal system in the space $P(G)$. Therefore, the direct and the inverse Fourier transform of a function $f \in P(G)$ are defined respectively by,

$$\mathbf{S}_f(w) = r_w g^{-1} \sum_{u=0}^{g-1} f(u) \mathbf{R}_w(u^{-1}), \tag{5}$$

$$f(x) = \sum_{w=0}^{K-1} Tr(\mathbf{S}_f(w) \mathbf{R}_w(x)). \tag{6}$$

Here and in the sequel we shall assume, without explicitly saying so, that all arithmetical operations are carried out in field P.

3.1. Matrix interpretation of the Fourier transform on finite non-Abelian groups

To derive the matrix interpretation of the Fourier transform on finite non-Abelian groups we need the generalized matrix multiplications defined as follows [18].

Definition 1. *Let* $\mathbf{A}$ *be an* $(m \times n)$ *matrix with elements* $a_{ij} \in P, \quad i \in \{0, 1, \ldots, m-1\}, \quad j \in \{0, 1, \ldots, n-1\}$. *Let* $[\mathbf{B}]$ *be an* $(n \times r)$ *matrix whose elements* $\mathbf{b}_{jk}, \quad j \in \{0, \ldots, n-1\}, \quad k \in \{0, 1, \ldots, r-1\}$ *are* $(p \times p)$ *matrices of not necessarily mutually equal orders with elements in* P. *We define the product* $\mathbf{A} \odot [\mathbf{B}]$ *as an* $(m \times r)$ *matrix* $[\mathbf{Y}]$ *whose elements* $y_{ik}, \quad i \in \{0, 1, \ldots, m-1\}, \quad k \in \{0, 1, \ldots, r-1\}$ *are* $(p \times q)$ *matrices with elements in* P *given by*

$$y_{ik} = \sum_{j=0}^{n-1} a_{ij} \mathbf{b}_{jk}.$$

The product $[\mathbf{B}] \odot \mathbf{A}$ is defined similarly.

Definition 2. *Let* $[\mathbf{Z}]$ *be an* $(m \times n)$ *matrix whose elements* $\mathbf{z}_{ij}$ $i \in \{0,1,\ldots,m-1\}$, $j \in \{0,1,\ldots,n-1\}$ *are the square matrices of not necessarily mutually equal orders with elements in* P. *Let* $[\mathbf{B}]$ *be an* $(n \times r)$ *matrix whose elements* $\mathbf{b}_{jk}$, $j \in \{0,1,\ldots,n-1\}$, $k \in \{0,1,\ldots,r-1\}$ *are square matrices of not necessarily mutually equal orders with elements in* P. *Under the condition that the matrices* $\mathbf{z}_{ij}$ *and* $\mathbf{b}_{jk}$ *are of the same order or, if not, that one of them is of the order 1, the product of matrices* $[\mathbf{Z}]$ *and* $[\mathbf{B}]$ *is defined as an* $(m \times r)$ *matrix* $\mathbf{Y} = [\mathbf{Z}] \circ [\mathbf{B}]$ *whose elements* $y_{ik} \in P$ *are given by*

$$y_{ik} = \sum_{j=0}^{n-1} Tr(\mathbf{z}_{ij}\mathbf{b}_{jk}).$$

Definition 3. *Let* $[\mathbf{Z}]$ *be an* $(m \times n)$ *matrix whose elements* $\mathbf{z}_{ij}$, $i \in \{0,1,\ldots,m-1\}$, $j \in \{0,1,\ldots,n-1\}$ *are* $(p \times q)$ *matrices of not necessarily mutually equal orders with elements in* P. *Let* $[\mathbf{B}]$ *be an* $(n \times r)$ *matrix whose elements* $\mathbf{b}_{jk}$, $j \in \{0,1,\ldots,n-1\}$, $k \in \{0,1,\ldots,r-1\}$ *are* $(s \times t)$ *matrices of not necessarily mutually equal orders with elements in* P. *The element-wise Kronecker product of matrices* $[\mathbf{Z}]$ *and* $[\mathbf{B}]$ *is defined as an* $(m \times r)$ *matrix* $[\mathbf{V}] = [\mathbf{Z}] \dot{\otimes} [\mathbf{B}]$ *whose elements* $\mathbf{v}_{ik}$ *are given by*

$$\mathbf{v}_{ik} = \sum_{j=0}^{n-1} \mathbf{z}_{ij} \otimes \mathbf{b}_{jk},$$

where $\otimes$ *denotes the ordinary Kronecker product.*

By using the matrix operations thus defined, the Fourier transform pair defined by (5) and (6) can be expressed as follows.

Definition 4. *Let* $f \in P(G)$ *be given as a vector* $\mathbf{f} = [f(0),\ldots,f(g-1)]^T$. *Then its Fourier transform is given by*

$$[\mathbf{S}_f] = g^{-1}[\mathbf{R}]^{-1} \odot \mathbf{f},$$

where $[\mathbf{S}_f] = [\mathbf{S}_f(0),\ldots,\mathbf{S}_f(K-1)]^T$, *and* $[\mathbf{R}]^{-1} = [\mathbf{b}_{sq}]$ *with* $\mathbf{b}_{sq} = r_w\mathbf{R}_s^{-1}(q)$, $s \in \{0,1,\ldots,K-1\}$, $q \in \{0,1,\ldots,g-1\}$.

The inverse Fourier transform is given by

$$\mathbf{f} = [\mathbf{R}] \circ [\mathbf{S}_f],$$

Table 2: Unitary irreducible representations of S_3 over $GF(11)$

x	$\mathbf{R}_0$	$\mathbf{R}_1$	$\mathbf{R}_2$
0	1	1	I
1	1	1	A
2	1	1	B
3	1	10	C
4	1	10	D
5	1	10	E

$$\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 5 & 8 \\ 3 & 5 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 5 & 3 \\ 8 & 5 \end{bmatrix},$$
$$\mathbf{C} = \begin{bmatrix} 1 & 0 \\ 0 & 10 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 5 & 8 \\ 8 & 6 \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} 5 & 3 \\ 3 & 6 \end{bmatrix}$$

where $[\mathbf{R}] = [\mathbf{a}_{ij}]$ *with* $\mathbf{a}_{ij} = \mathbf{R}_j(i)$, $i \in \{0, 1, \ldots, g-1\}$, $j \in \{0, 1, \ldots, K-1\}$.

Example 2. *The unitary irreducible representations of* C_2 *in Example 1 reduce to the group characters and considered over the Galois field* $GF(11)$ *are given by the Walsh transform matrix in* $GF(11)$

$$\mathbf{W} = 6 \begin{bmatrix} 1 & 1 \\ 1 & 10 \end{bmatrix}.$$

The unitary irreducible representations of S_3 *over the Galois field* $GF(11)$ *are shown in Table 2. Therefore, the Fourier transform on* G_{24} *in Example 1 is defined by the transform matrix*

$$[\mathbf{R}]^{-1} = \mathbf{W} \otimes \mathbf{W} \otimes [\mathbf{S}_3] = 6 \cdot 6 \cdot 2 \cdot \left(\begin{bmatrix} 1 & 1 \\ 1 & 10 \end{bmatrix} \otimes \begin{bmatrix} 1 & 1 \\ 1 & 10 \end{bmatrix} \otimes \right.$$
$$\left. \otimes \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 10 & 10 & 10 \\ 2\mathbf{I} & 2\mathbf{B} & 2\mathbf{A} & 2\mathbf{C} & 2\mathbf{D} & 2\mathbf{E} \end{bmatrix} \right),$$

with all the calculations performed in $GF(11)$.

For function f *in Example 1, the Fourier spectrum is given by the matrix-valued vector*

$$[\mathbf{S}_f] = 6[5, 9, \mathbf{S}_f(2), 3, 5, \mathbf{S}_f(5), 6, 7, \mathbf{S}_f(8), 7, 1, \mathbf{S}_f(11)]^T,$$

where

$$\mathbf{S}_f(2) = \begin{bmatrix} 5 & 2 \\ 9 & 5 \end{bmatrix}, \quad \mathbf{S}_f(5) = \begin{bmatrix} 5 & 2 \\ 9 & 8 \end{bmatrix},$$
$$\mathbf{S}_f(8) = \begin{bmatrix} 9 & 2 \\ 9 & 1 \end{bmatrix}, \quad \mathbf{S}_f(11) = \begin{bmatrix} 1 & 2 \\ 9 & 1 \end{bmatrix}.$$

4. MATRIX INTERPRETATION OF THE FAST FOURIER TRANSFORM ON FINITE NON-ABELIAN GROUPS

A fast algorithm for implementation of the Fourier transform on finite non-Abelian groups based on the classical Coley-Tukey FFT, is formulated in an analytical form in [8]. It seems that an earlier relevant result on this subject can be found in [5]. The matrix interpretation of FFT on finite non-Abelian groups we will discuss was given in [18].

To obtain a fast algorithm for the computation of the Fourier transform on finite non-Abelian groups we use the Good-Thomas factorization as in the case of the FFT on finite Abelian groups [6], [7], [23].

The matrix $[\mathbf{R}]$ in the definition of the Fourier transform on finite non-Abelian groups is the matrix of unitary irreducible representations of G over P. Since G is representable in the form (1), the matrix $[\mathbf{R}]$ can be generated as the Kronecker product of $(K_i \times g_i)$ matrices $[\mathbf{R}_i]$ of unitary irreducible representations of subgroups G, $i \in \{1, \ldots, n\}$, i.e.,

$$[\mathbf{R}] = \bigotimes_{i=1}^{n} [\mathbf{R}_i],$$

where $\otimes$ denotes the Kronecker product.

Owing to the well-known properties of the Kronecker product, the same applies to the matrix $[\mathbf{R}]^{-1}$, i.e., for this matrix the following holds

$$[\mathbf{R}]^{-1} = \bigotimes_{i=1}^{n} [\mathbf{R}_i]^{-1}.$$

This matrix can be factorized further into the element-wise Kronecker product of n sparse factors $[\mathbf{C}^i]$, $i \in \{1, \ldots, n\}$ as

$$[\mathbf{C}^i] = \bigotimes_{j=1}^{n} [\mathbf{S}_j^i], \quad i = 1, \ldots, n,$$

where

$$[\mathbf{S}_j^i] = \begin{cases} \mathbf{I}_{(g_j \times g_j)}, & j < i \\ [\mathbf{R}_j]^{-1}, & j = i \\ \mathbf{I}_{(K_j \times K_j)}, & j > i \end{cases} \tag{7}$$

where $\mathbf{I}_{a \times a}$ is an $(a \times a)$ identity matrix.

Each matrix $[\mathbf{C}^i]$ uniquely describes one step of the fast Fourier transform performed in n steps. The algorithm is best represented and performed through a flow-graph consisting of nodes connected with branches to which some weights are associated [18].

The matrix representation and the corresponding fast algorithm obtained in such a way is similar to the FFT on finite Abelian groups, but some important differences appear here. See [18] for a detailed discussion.

5. COMPUTATION OF THE FOURIER TRANSFORM ON FINITE NON-ABELIAN GROUPS THROUGH DDs

FFT algorithms on Abelian and non-Abelian groups are based upon the vector representations of discrete functions. It follows from the definition of FFT and their matrix description that the space complexity of FFT on a decomposable group G of order g approximates $O(g)$. The time complexity is $O(ng)$. Thus, the application of FFT is restricted to groups of relatively small orders. To overcome this restriction, the calculation procedures based on the decision diagram representation of discrete functions are proposed for various discrete transforms on Abelian groups [3], [13], [14], [22]. That approach can be extended to finite non-Abelian groups owing to the matrix interpretation in Section 4. For this purpose we need the following definition.

Definition 5. *The operation of concatenation, denoted by $\diamond$, over an ordered set of n vectors*
$\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ *of order m is the operation producing a vector $\mathbf{S}$ of order nm consisting of n successive subvectors $\mathbf{A}_1, \ldots, \mathbf{A}_n$.*

Example 3. *Applying the operation of concatenation to the set of three vectors $\mathbf{A} = [a_1 a_2 a_3]^T$, $\mathbf{B} = [b_1 b_2 b_3]^T$, $\mathbf{C} = [c_1 c_2 c_3]^T$ produces the vector* $\mathbf{D} = \mathbf{A} \diamond \mathbf{B} \diamond \mathbf{C} = [a_1 a_2 a_3 b_1 b_2 b_3 c_1 c_2 c_3]^T$.

From the theory of Good-Thomas FFT, calculation of the Fourier transform on a decomposable group G of order g, can be performed through n Fourier transforms on the constituent subgroups G_i of orders g_i. It follows that the calculation of the Fourier coefficients of f on a finite decomposable

begin {*procedure*}
    for $i = n$ to 1
        for $k = 0$ to Q_i do
        Determine $q_{i,k}$ by using the rule (8).
    Store $[\mathbf{S}_f] = q_1$.
end{*procedure*}

Figure 1: Calculation procedure for the Fourier transform.

group G of order g given by the DD can be carried out through the following procedure.

5.1. Calculation procedure

Given a function f on the decomposable group G of the form (1):

1. Represent f by the generalized MTBDD. Denote by Q_i the number of non-terminal nodes at the i-th level, i.e., the level corresponding to the variable x_i of f, in this DD.

2. Descend the DD in a recursive way level by level starting from the constant nodes at level $(n+1)$ up to the root node at level 1.

3. For $i = n$ to 1, process the nodes and cross points in the DD by using the rule

$$q_{i,k}(w_i) = r_{w_i} g_i^{-1} \sum_{j=0}^{g_i-1} q_{i+1,j} \mathbf{R}_{w_i}(x_j^{-1}), \quad k = 0, \ldots, Q_i - 1, \qquad w_i = 0, \ldots, K_i - 1, \tag{8}$$

   easily derived form the matrix factorization of $[\mathbf{R}]^{-1}$.

End of procedure.

Note that for $i = n$, $q_{i+1,j}$ takes the value of the constant node where the j-th outgoing edge of the corresponding node points to. Expressed in programming pseudo code, the procedure is shown in Fig. 2.

It should be pointed out that there are no matrix computations. Moreover, all calculations are preformed as vector operations in the computation of Fourier coefficients through this procedure, which ensures its efficiency. The matrix-valued vector determined in q_1 is the Fourier spectrum of f.

The procedure is probably best explained through an example using the matrix notation.

Example 4. *The Fourier spectrum of f in Example 1 is calculated using the proposed procedure as follows.*

1. *The non-terminal nodes $q_{3,0}, q_{3,1}, q_{3,3}$ and the cross point $q_{3,2}$ are processed first using matrix $[\mathbf{S}_3]$ in $[\mathbf{R}]^{-1}$. The input data for the procedure are the values of constant nodes. Thus, the following is determined*

$$q_{3,0} = \begin{bmatrix} 7 \\ 3 \\ \begin{bmatrix} 10 & 4 \\ 7 & 2 \end{bmatrix} \end{bmatrix}, \quad q_{3,1} = \begin{bmatrix} 8 \\ 8 \\ \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \end{bmatrix},$$

$$q_{3,2} = \begin{bmatrix} 1 \\ 0 \\ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{bmatrix}, \quad q_{3,3} = \begin{bmatrix} 5 \\ 7 \\ \begin{bmatrix} 7 & 0 \\ 0 & 4 \end{bmatrix} \end{bmatrix}.$$

2. *The non-terminal nodes $q_{2,0}, q_{2,1}$ are processed using $\mathbf{W}$. The following is determined*

$$q_{2,0} = 6\left(\begin{bmatrix} 7 \\ 3 \\ \begin{bmatrix} 10 & 4 \\ 7 & 2 \end{bmatrix} \end{bmatrix} + \begin{bmatrix} 8 \\ 8 \\ \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \end{bmatrix}\right) \diamond$$
$$\diamond\, 6\left(\begin{bmatrix} 7 \\ 3 \\ \begin{bmatrix} 10 & 4 \\ 7 & 2 \end{bmatrix} \end{bmatrix} + 10\begin{bmatrix} 8 \\ 8 \\ \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \end{bmatrix}\right),$$

$$q_{2,1} = 6\left(\begin{bmatrix} 1 \\ 0 \\ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{bmatrix} + \begin{bmatrix} 5 \\ 7 \\ \begin{bmatrix} 7 & 0 \\ 0 & 4 \end{bmatrix} \end{bmatrix}\right) \diamond$$
$$\diamond\, 6\left(\begin{bmatrix} 1 \\ 0 \\ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{bmatrix} + 10\begin{bmatrix} 5 \\ 7 \\ \begin{bmatrix} 7 & 0 \\ 0 & 4 \end{bmatrix} \end{bmatrix}\right).$$

3. The root node is processed using **W**. *The following is determined*

$$
q_1 = 6\left(\begin{bmatrix} 2 \\ 0 \\ \begin{bmatrix} 7 & 2 \\ 9 & 3 \end{bmatrix} 5 \\ 3 \\ \begin{bmatrix} 3 & 2 \\ 9 & 10 \end{bmatrix} \end{bmatrix} + \begin{bmatrix} 3 \\ 9 \\ \begin{bmatrix} 9 & 0 \\ 0 & 2 \end{bmatrix} 9 \\ 2 \\ \begin{bmatrix} 2 & 0 \\ 0 & 9 \end{bmatrix} \end{bmatrix}\right) \diamond
$$

$$
\diamond\, 6\left(\begin{bmatrix} 2 \\ 0 \\ \begin{bmatrix} 7 & 2 \\ 9 & 3 \end{bmatrix} 5 \\ 3 \\ \begin{bmatrix} 3 & 2 \\ 9 & 10 \end{bmatrix} \end{bmatrix} + 10 \begin{bmatrix} 3 \\ 9 \\ \begin{bmatrix} 9 & 0 \\ 0 & 2 \end{bmatrix} 9 \\ 2 \\ \begin{bmatrix} 2 & 0 \\ 0 & 9 \end{bmatrix} \end{bmatrix}\right).
$$

Thus, the Fourier spectrum of f is equal to the matrix-valued vector determined in q_1 and it is equal to that calculated by definition in Example 1.

Thanks to its recursive structure, the calculation procedure as well as the calculated Fourier spectrum can be represented by decision diagrams. The corresponding DDs differ from these used to represent f in the same way that the FFT algorithms on Abelian groups differ from FFT on non-Abelian groups [18]. DDs representing the Fourier spectrum of functions on finite non-Abelian groups are matrix-valued, since the values of constant nodes are Fourier coefficients. The number of outgoing edges of nodes at the i-th level are determined by the cardinality K_i of the dual object Γ_i of G_i.

The efficiency of the DD representation of f depends on the number of different values f can take. In the same way, the efficiency of DDs representation of the Fourier spectrum of f depends on the number of different Fourier coefficients. In this way, it depends indirectly on the number of different values of f.

6. CLOSING REMARKS

There are at least two reasons to extend the fast calculation algorithms for the Fourier transform to functions on finite non-Abelian groups:

1. There are signals in reality that are naturally modelled by functions on finite non-Abelian groups. Examples of such problems are reviewed in [9], [10], [24].

2. In some applications, the use of non-Abelian groups offers considerable advantages compared to the Abelian groups [11], [15].

Unlike FFT, the procedure presented in this paper does not require storage of the complete vector representing values of f and, thus, permits the Fourier processing of functions on large finite non-Abelian groups. In that way, it removes the complexity of calculation as a limiting factor for applications of the Fourier transform on finite non-Abelian groups.

Acknowledgments

The authors thank the reviewer for comments that improved the presentation in this paper.

REFERENCES

[1] Bryant, R.E. *Graph-based algorithms for Boolean functions manipulation,* IEEE Trans. Comput., Vol.C-35, No.8 (1986), 667-691.

[2] Bryant, R.E., Chen, Y-A., *Verification of arithmetic functions with binary moment decision diagrams*, unpublished paper, May 31, 1994, CMU-CS-94-160.

[3] Clarke, E.M., McMillan, K.L., Zhao, X., Fujita, M., *Spectral transforms for extremely large Boolean functions,* in: Kebschull, U., Schubert, E., Rosentiel, W., (eds), *Proc. IFIP WG 10.5 Workshop on Applications of the Reed-Muller Expansion in Circuit Design*, 16-17.9.1993, Hamburg, Germany, 86-90.

[4] Cooley, J.W., Tukey, J.W., *An algorithm for the machine calculation of complex Fourier series,* Math. Computation, 19 (1965), 297-301.

[5] Depeyrot, M., *Fondements algebriques de la transformation de Fourier rapide,* 1E/11, Centre d'Automatique, Ecole des Mines, Fontainebleau, France, 1969.

[6] Good, I.J., *The interaction algorithm and practical Fourier analysis,* J. Roy. Statist. Soc., ser. B, Vol.20 (1960), 361-372, Addendum, Vol.22 (1960), 372-375.

[7] Good, I.J., *The relationship between two fast Fourier transforms,* IEEE Trans. Comput., Vol.C-20 (1971), No.3, 310-317.

[8] Karpovsky, M.G., *Fast Fourier transforms on finite non-Abelian groups,* IEEE Trans. Comput., Vol.C-26, No.10 (1977), 1028-1030.

[9] Karpovsky, M.G., Trachtenberg, E.A., *Some optimization problems over finite groups,* Inform. Control. 34 (1977), 1-22.

[10] Karpovsky, M.G., Trachtenberg, E.A., *Fourier transform over finite groups for error detection and error correction in computation channels,* Inform. Control Vol.40, No.3 (1979), 335-358.

[11] Karpovsky, M.G., Trachtenberg, E.A., *Statistical and computational performance of a class of generalized Wiener filters,* IEEE Trans. Vol.IT-32 (1986).

[12] Kurepa, Dj.R., *Ensembles ordonnés et ramifiès* (Thèse), Paris, 1935, 6+138+2p.

[13] Lai, Y.F., Pedram, M., Vrudhula, S.B.K., *EVBDD-based algorithms for integer linear programming, spectral transformation, and functional decomposition,* IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems Vol.13, No.8, (1994), 959-975.

[14] Miller, D.M., *Spectral transformation of multiple-valued decision diagrams*, in: *Proc. 24th Int. Symp. on Multiple-valued Logic*, Boston, Massachusetts, USA, 22.-25.5.1994, 89-96.

[15] Roziner, T.D., Karpovsky, M.G., Trachtenberg, L.A., *Fast Fourier transforms over finite groups by multiprocessor system,* IEEE Trans., Vol.ASSP-38, No.2 (1990), 226-240.

[16] Sasao, T., *Representations of logic functions by using EXOR operators,* in: *Proc. IFIP WG 10.5 Workshop on Application of the Reed-Muller Expansion in Circuit Design*, 27-29.8.1995, Makuhari, Chiba, Japan, 11-20.

[17] A. Srinivasan, T. Kam, Sh. Malik, R.K. Brayant, *Algorithms for discrete function manipulation*, in: *Proc. Inf. Conf. on CAD*, 1990, 92-95.

[18] Stanković, R.S., *Matrix-interpretation of fast Fourier transforms on finite non-Abelian groups*, in: *Proc. Int. Conf. on Signal Processing, Beijing/90*, 22.-26.10.1990, Beijing, China, 1187-1190.

[19] Stanković, R.S., *Functional decision diagrams for multiple-valued functions*, in: *Proc. 25th Int. Symp. on Multiple-Valued Logic*, Bloomington, Indiana, U.S.A., 22.-26.5.1995, 284-289.

[20] Stanković, R.S., Sasao, T., Moraga, C., *Spectral transform decision diagrams*, in: *Proc. IFIP WG 10.5 Workshop on Application of the Reed-Muller Expansion in Circuit Design*, 27-29.8.1995, Makuhari, Chiba, Japan, 46-53.

[21] Stanković, R.S., Moraga, C., *Edge-valued decision diagrams for switching functions based upon the partial Reed-Muller transforms*, in: *Proc. 2nd Int. Conf. Application of Computer Systems*, Szczecin, Poland, 8-9.1995, 7-16.

[22] Stanković, R.S., Stanković, M., Moraga, C., Sasao, T., *Calculation of Vilenkin-Chrestenson transform coefficients of multiple-valued functions through multiple-place decision diagrams*, in: *Proc. 5th Int. Workshop on Spectral Techniques*, March 15-17, 1994, Beijing, China, 107-116.

[23] Thomas, L.H., *Using a computer to solve problems in physics*, In: *Application of Digital Computers*, Boston, Mass., Ginn, 1963.

[24] Trachtenberg, E.A., Karpovsky, M.G., *Filtering in a communication channel by Fourier transforms over finite groups*, in: M.G. Karpovsky (ed.), *Spectral Techniques and Fault Detection*, Academic Press, Orlando, Florida, 1985.

Number 19-20 (1996) YU ISSN 0350-2910

SCIENTIFIC REVIEW

CONTENTS

Number 19-20 (1996) YU ISSN 0350-2910

SCIENTIFIC REVIEW

CONTENTS (continued)

Printed by "Science and Society", Božidara Adžije 11a, Beograd