SCIENTIFIC ISSUES
JAN DŁUGOSZ UNIVERSITY
IN CZĘSTOCHOWA

MATHEMATICS XV

2010

Scientific Editor
Yuriy POVSTENKO

Proofreader
Urszula BRZOZOWSKA

Computer Typesetting and Making-up
Urszula BRZOZOWSKA

Graphical Project of a Cover
Sławomir SADOWSKI

ISBN 978-83-7455-161-8
ISSN 1896-0286

Publishing House of Jan Długosz University in Częstochowa
ul. Waszyngtona 4/8, tel. (0-34) 378-43-29, fax (0-34) 378-43-19
e-mail: wydawnictwo@ajd.czest.pl
www.ajd.czest.pl

PRACE NAUKOWE
AKADEMIA IM. JANA DŁUGOSZA
W CZĘSTOCHOWIE

MATEMATYKA XV

2010

Redaktor naukowy
Yuriy POVSTENKO

Korektor
Urszula BRZOZOWSKA

Skład i łamanie
Urszula BRZOZOWSKA

Projekt graficzny okładki
Sławomir SADOWSKI

ISBN 978-83-7455-161-8
ISSN 1896-0286

Wydawnictwo Akademii im. Jana Długosza w Częstochowie
ul. Waszyngtona 4/8, tel. (0-34) 378-43-29, fax (0-34) 378-43-19.
e-mail: wydawnictwo@ajd.czest.pl
www.ajd.czest.pl

CONTENTS

Part I. Mathematics and Its Applications

Part II. Computer Science

PART I
MATHEMATICS AND ITS APPLICATIONS

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XV, Częstochowa 2010

INDECOMPOSABLE PROJECTIVE REPRESENTATIONS OF DIRECT PRODUCTS OF FINITE GROUPS OVER A RING OF FORMAL POWER SERIES

Leonid F. Barannyk, Dariusz Klein

Institute of Mathematics
Pomeranian University of Słupsk
Arciszewskiego 22b, 76-200 Słupsk, Poland
e-mail: barannyk@apsl.edu.pl, klein@apsl.edu.pl

Abstract. Let F be a field of characteristic $p > 0$, $S = F[[X]]$ the ring of formal power series in the indeterminate X with coefficients in the field F, F^* the multiplicative group of F, $G = G_p \times B$ a finite group, where G_p is a p-group and B is a p'-group. We give necessary and sufficient conditions for G and F under which there exists a cocycle $\lambda \in Z^2(G, F^*)$ such that every indecomposable projective S-representation of G with the cocycle λ is the outer tensor product of an indecomposable projective S-representation of G_p and an irreducible projective S-representation of B.

1. Introduction

Let F be a field of characteristic $p > 0$ and $G = G_p \times B$, where G_p is a Sylow p-subgroup. Blau [6] and Gudyvok [10, 11] proved that every finitely generated FG-module is the outer tensor product $V\#W$ of an indecomposable FG_p-module V and an irreducible FB-module W if and only if either G_p is cyclic or F is a splitting field for B. Gudyvok [12, 13] also investigated a similar problem for group rings KG, where K is a complete discrete valuation ring. In particular, he proved that if K is of characteristic $p > 0$ and T is the quotient field of K, then every indecomposable KG-module is of the form $V\#W$ if and only if either $|G_p| = 2$ or T is a splitting field for B. In the paper [2], the results of Blau and Gudyvok were generalized to the twisted group rings $S^\lambda G$, where $G = G_p \times B$, $S = F$ or S is a complete discrete valuation ring of characteristic $p > 0$.

In this paper we continue the study of indecomposable projective representations of $G = G_p \times B$ over the ring $S = F[[X]]$ as begun in [2].

Let us present the main results of the paper. We assume that F is a field of characteristic $p > 0$, S^* the unit group of S, $|G_p| \neq 1$, $|B| \neq 1$, and if G_p is non-Abelian, then F contains a primitive q^{th} root of 1 for every prime $q \mid |B|$ such that $p \mid (q-1)$. Given a cocycle $\lambda\colon G \times G \to S^*$ in $Z^2(G, S^*)$, we denote by $S^\lambda G$ the twisted group ring of the group G over the ring S with the 2-cocycle λ. By an $S^\lambda G$-module we mean a finitely generated left $S^\lambda G$-module which is S-free. Given $\mu \in Z^2(G_p, S^*)$, the kernel $\mathrm{Ker}(\mu)$ of μ is the union of all cyclic subgroups $\langle\, g \,\rangle$ of G_p such that the restriction of μ to $\langle\, g \,\rangle \times \langle\, g \,\rangle$ is a coboundary. We recall from [4, p. 268] that $G_p' \subset \mathrm{Ker}(\mu)$, $\mathrm{Ker}(\mu)$ is a normal subgroup of G_p and the restriction of μ to $\mathrm{Ker}(\mu) \times \mathrm{Ker}(\mu)$ is a coboundary (see also [3, p. 197] for a simple proof). Up to cohomology in $Z^2(G_p, S^*)$, we have $\mu_{g,a} = \mu_{a,g} = 1$ for all $g \in G_p$ and $a \in \mathrm{Ker}(\mu)$. In what follows, we assume that every cocycle $\mu \in Z^2(G_p, S^*)$ under consideration satisfies this condition. If H is a subgroup of G, then the restriction of $\lambda \in Z^2(G, S^*)$ to $H \times H$ will also be denoted by λ. In this case, $S^\lambda H$ is a subring of $S^\lambda G$. A group G is of symmetric type if it decomposes into a direct product of two isomorphic groups. Denote

$$i(F) = \begin{cases} t & \text{if } [F : F^p] = p^t, \\ \infty & \text{if } [F : F^p] = \infty. \end{cases}$$

Let $G = G_p \times B$, $\mu \in Z^2(G_p, S^*)$ and $\nu \in Z^2(B, S^*)$. Then the map $\mu \times \nu\colon G \times G \to S^*$ defined by

$$(\mu \times \nu)_{x_1 b_1, x_2 b_2} = \mu_{x_1, x_2} \cdot \nu_{b_1, b_2}$$

for all $x_1, x_2 \in G_p$, $b_1, b_2 \in B$ belongs to $Z^2(G, S^*)$. Every cocycle $\lambda \in Z^2(G, S^*)$ is cohomologous to $\mu \times \nu$, where μ is the restriction of λ to $G_p \times G_p$ and ν is the restriction of λ to $B \times B$. From now on, we suppose that each cocycle $\lambda \in Z^2(G, S^*)$ under consideration satisfies the condition $\lambda = \mu \times \nu$.

For any $\lambda = \mu \times \nu \in Z^2(G, S^*)$, we have $S^\lambda G \cong S^\mu G_p \otimes_S S^\nu B$. If every indecomposable $S^\lambda G$-module is isomorphic to the outer tensor product $V \# W$, where V is an indecomposable $S^\mu G_p$-module and W is an irreducible $S^\nu B$-module, then we will say that the ring $S^\lambda G$ is of OTP representation type.

Let Ω be a subgroup of S^*. We say that a group $G = G_p \times B$ is of OTP projective (S, Ω)-representation type if there exists a cocycle $\lambda \in Z^2(G, \Omega)$ such that the ring $S^\lambda G$ is of OTP representation type. A group $G = G_p \times B$ is defined to be of purely OTP projective (S, Ω)-representation type if $S^\lambda G$ is of OTP representation type for any $\lambda \in Z^2(G, \Omega)$. If $\Omega = S^*$, then instead of "(S, Ω)-representation type" we write "S-representation type".

In Section 3, we characterize twisted group rings of OTP representation type. Let $G = G_p \times B$, $\mu \in Z^2(G_p, S^*)$, $\nu \in Z^2(B, S^*)$, $\lambda = \mu \times \nu$ and $H = \mathrm{Ker}(\mu)$. In Theorem 1, we prove that if $|H| > 2$, then the ring $S^\lambda G$ is of OTP representation type if and only if F is a splitting field for the F-algebra $S^\nu B/XS^\nu B$. Assume that $|G_p'| \neq 2$, $\mu \in Z^2(G_p, F^*)$, $\nu \in Z^2(B, S^*)$ and $\lambda = \mu \times \nu$. In Proposition 3, we show that $S^\lambda G$ is of OTP representation type if and only if one of the following conditions is satisfied:

(i) $F^\mu G_p$ is a field;

(ii) $p = 2$, $|G_2'| = 1$ and $2\dim_F(F^\mu G_2/\operatorname{rad} F^\mu G_2) = |G_2|$;

(iii) F is a splitting field for the F-algebra $S^\nu B/XS^\nu B$.

In Section 4, we study the groups of OTP projective representation type. Let $G = G_p \times B$, $|G_p'| \neq 2$ and s be the number of invariants of G_p/G_p'. In Theorem 2, we prove that G is of OTP projective (S, F^*)-representation type if and only if one of the following conditions is satisfied:

(i) $|G_p'| = 1$ and $s \leq i(F)$;

(ii) $p = 2$, $|G_2'| = 1$, $s = i(F) + 1$ and G_2 has at least one invariant equal to 2;

(iii) F is a splitting field for $F^\sigma B$ for some $\sigma \in Z^2(B, F^*)$.

Let $G = G_p \times B$ be an Abelian group and s the number of invariants of G_p. In Proposition 5, we establish that G is of OTP projective (S, F^*)-representation type if and only if one of the following conditions is satisfied:

(i) $s \leq i(F)$;

(ii) $p = 2$, $s = i(F) + 1$ and G_2 has at least one invariant equal to 2;

(iii) B has a subgroup H such that B/H is of symmetric type and F contains a primitive m^{th} root of 1, where $m = \max\{\exp(B/H), \exp H\}$.

In Section 5, we show in Theorem 3 that $G = G_p \times B$ is of purely OTP projective S-representation type if and only if $|G_p| = 2$ or F is a splitting field for any $F^\nu B$. Corollary to Theorem 3 asserts that if G is a nilpotent group, then G is of purely OTP projective S-representation type if and only if one of the following conditions is satisfied:

(i) $|G_p| = 2$;

(ii) $F = F^q$ and F contains a primitive q^{th} root of 1 for every prime $q \mid |B|$.

2. Preliminaries

Throughout this paper, we use the following notations: $p \geq 2$ is a prime; F is a field of characteristic $p > 0$; $S = F[[X]]$ is the ring of formal power series in the indeterminate X with coefficients in the field F; $P = XS$ is unique maximal ideal of S; F^* is the multiplicative group of F; $F^q = \{\alpha^q \colon \alpha \in F\}$; S^* is the unit group of S; $G = G_p \times B$ is a finite group, where G_p is a p-group and B is a p'-group; H' is the commutant of a group H, e is the identity

element of H, $|h|$ is the order of $h \in H$; $\operatorname{soc} A$ is the socle of an Abelian group A and $\exp A$ is the exponent of A. We suppose that $|G_p| > 1$ and $|B| > 1$. Given a subgroup Ω of S^*, we denote by $Z^2(H, \Omega)$ the group of all Ω-valued normalized 2-cocycles of the group H, where we assume that H acts trivially on Ω. An S-basis $\{u_h \colon h \in H\}$ of $S^\lambda H$ satisfying $u_a u_b = \lambda_{a,b} u_{ab}$ for all $a, b \in H$ is called natural (corresponding to $\lambda \in Z^2(H, S^*)$). Given an $S^\lambda H$-module V, we write $\operatorname{End}_{S^\lambda H}(V)$ for the ring of all $S^\lambda H$-endomorphisms of V, $\operatorname{rad}\operatorname{End}_{S^\lambda H}(V)$ for the Jacobson radical of $\operatorname{End}_{S^\lambda H}(V)$ and $\overline{\operatorname{End}_{S^\lambda H}(V)}$ for the quotient ring

$$\operatorname{End}_{S^\lambda H}(V)/\operatorname{rad}\operatorname{End}_{S^\lambda H}(V).$$

Moreover, we denote by $\widetilde{S^\lambda H}$ the F-algebra $S^\lambda H/XS^\lambda H$ and by $\widetilde{V}$ the factor module V/XV. Given $\lambda \in Z^2(H, F^*)$, $F^\lambda H$ denotes the twisted group algebra of H over F and $\overline{F^\lambda H}$ the quotient algebra of $F^\lambda H$ by the radical $\operatorname{rad} F^\lambda H$. We identify an element $a + P$, $a \in F$, of the field $\bar{S} = S/P$ with the element a.

Lemma 1. [8, p.125] *Let H be a finite group, $\lambda \in Z^2(H, S^*)$ and V an $S^\lambda H$-module. Then V is indecomposable if and only if $\overline{\operatorname{End}_{S^\lambda H}(V)}$ is a skewfield.*

Lemma 2. *Let H be a finite p-group, D a subgroup of H, $\lambda \in Z^2(H, S^*)$ and M an indecomposable $S^\lambda D$-module. Assume that $\overline{\operatorname{End}_{S^\lambda D}(M)}$ is isomorphic to a field K, $K \supset F$ and one of the following conditions is satisfied:*

(i) H is Abelian;

(ii) $[s(K) : F]$ is not divisible by p, where $s(K)$ is the separable closure of F in K.

Then $M^H := S^\lambda H \otimes_{S^\lambda D} M$ is an indecomposable $S^\lambda H$-module and

$$\overline{\operatorname{End}_{S^\lambda H}(M^H)}$$

is isomorphic to a field that is a finite purely inseparable extension of the field K.

The proof is similar to that of Lemma 2.2 [2, p.540]. It uses the same idea as in Theorem 8 of [9].

Lemma 3. *Let K be a finite separable extension of the field F and H a finite p-group. If $|H| > 2$, then there exists an indecomposable SH-module V such that $\overline{\operatorname{End}_{SH}(V)}$ is isomorphic to K.*

P r o o f. Let $K = F(\theta)$, $f(t)$ be the monic minimal polynomial of θ over F and Γ the companion matrix of $f(t)$. Assume that either H is cyclic of order

$|H| > 2$ or H is a group of type $(2,2)$. Let $H = \langle a \rangle$ and V be the underlying SH-module of the representation

$$a \mapsto \begin{pmatrix} E & XE & \Gamma \\ 0 & E & XE \\ 0 & 0 & E \end{pmatrix}$$

of H, where E is the identity matrix of order $n = \deg f(t)$. Then, by [13, pp. 70–71], $\overline{\mathrm{End}_{SH}(V)} \cong K$. If $H = \langle a \rangle \times \langle b \rangle$ is a group of type $(2,2)$, then as V we take the underlying SH-module of the representation

$$a \mapsto \begin{pmatrix} E & E \\ 0 & E \end{pmatrix}, \quad b \mapsto \begin{pmatrix} E & \Gamma \\ 0 & E \end{pmatrix}.$$

By [13, p. 71], we have $\overline{\mathrm{End}_{SH}(V)} \cong K$. □

Lemma 4. *Let $p = 2$, $[F : F^2] = 2$, H be a 2-group such that $|H| \neq 8$ and $|H'| = 2$. Assume also that K is a finite separable extension of the field F and $[K : F]$ is not divisible by 2. Then, for any $\lambda \in Z^2(H, F^*)$, there exists an indecomposable $S^\lambda H$-module V such that $\overline{\mathrm{End}_{S^\lambda H}(V)}$ is isomorphic to a field that is a finite purely inseparable extension of the field K.*

P r o o f. Let $H' = \langle c \rangle$, s be the number of invariants of the Abelian group H/H', D the subgroup of H such that $H' \subset D$ and $D/H' = \mathrm{soc}(H/H')$. We have

$$S^\lambda D / S^\lambda D(u_c - u_e) \cong S^{\bar\lambda} \bar D,$$

where $\bar D = D/H'$ and $\bar\lambda_{xH',yH'} = \lambda_{x,y}$ for all $x, y \in D$. Assume $s > 2$. Since $i(F) = 1$,

$$F^{\bar\lambda}\bar D \cong F^{\bar\lambda}\bar D_1 \otimes_F F\bar D_2,$$

where $\bar D = \bar D_1 \times \bar D_2$ and $|\bar D_2| \geq 4$. It follows that $S^{\bar\lambda}\bar D \cong S^{\bar\lambda}\bar D_1 \otimes_S S\bar D_2$. By Lemmas 2 and 3, there exists an indecomposable $S^{\bar\lambda}\bar D$-module V such that

$$\overline{\mathrm{End}_{S^{\bar\lambda}\bar D}(V)}$$

is a finite purely inseparable extension of the field K. The module V is also an $S^\lambda D$-module. In view of Lemma 2, V^H is an indecomposable $S^\lambda H$-module and

$$\overline{\mathrm{End}_{S^\lambda H}(V^H)}$$

is a finite purely inseparable extension of K.

Now we consider the case $s = 2$. Since $|H| > 8$, then D is Abelian. Let $D = \langle a \rangle \times \langle b \rangle$, where $a^2 = c$ and $b^2 = e$. Then

$$S^\lambda D = \bigoplus_{i,j,k} S u_a^i u_b^j u_c^k,$$

where

$$u_a^2 = \alpha u_c, \quad u_b^2 = \beta u_e, \quad u_c^2 = u_e$$

and $\alpha, \beta \in F^*$. If $\alpha \in F^2$, then $S[u_a]$ is the group ring of the group $\langle a \rangle$ over the ring S. If $\beta \in F^2$ then $S^\lambda D$ contains the group ring SQ, where $Q = \langle c \rangle \times \langle b \rangle$. Assume that $\alpha \notin F^2$ and $\beta \notin F^2$. Since $i(F) = 1$, $\alpha^{-1} = \delta_0^2 + \delta_1^2 \beta$ for some $\delta_0, \delta_1 \in F$. Let $v = u_a(\delta_0 u_e + \delta_1 u_b)$. Then $v^2 = \alpha u_c \cdot \alpha^{-1} u_e = u_c$.

If $D = \langle a \rangle \times \langle b \rangle \times \langle c \rangle$ is of type $(2,2,2)$, then $S^\lambda D$ contains SQ, where Q is a group of type $(2,2)$.

Applying Lemmas 2 and 3, we finish the proof. □

Lemma 5. *Let $G = G_p \times B$ and $\lambda \in Z^2(G, S^*)$. The ring $S^\lambda G$ is of OTP representation type if and only if the outer tensor product of any indecomposable $S^\lambda G_p$-module and any irreducible $S^\lambda B$-module is an indecomposable $S^\lambda G$-module.*

The proof is similar to that of the corresponding fact for a group ring (see [6, p. 41], [13, p. 68]).

Let B be a finite p'-group and $\lambda \in Z^2(B, S^*)$. We denote by $\widetilde{S^\lambda B}$ the F-algebra $S^\lambda B / X S^\lambda B$. For $y \in S^\lambda B$, let $\tilde{y}$ denote $y + X S^\lambda B$. The F-algebra $\widetilde{S^\lambda B}$ is separable. By Theorem 6.8 [8, p. 124], if

$$\widetilde{S^\lambda B} = \widetilde{S^\lambda B}\varepsilon_1 \oplus \ldots \oplus \widetilde{S^\lambda B}\varepsilon_n$$

is a decomposition into minimal left ideals, then there exists a decomposition

$$S^\lambda B = S^\lambda B e_1 \oplus \ldots \oplus S^\lambda B e_n,$$

where ε_i is an idempotent of $\widetilde{S^\lambda B}$, e_i is an idempotent of $S^\lambda B$ and $\tilde{e_i} = \varepsilon_i$ for every $i \in \{1, \ldots, n\}$. Each ideal $S^\lambda B e_i$ is an irreducible $S^\lambda B$-module. By Theorem 76.8 [7, p. 532] and Corollary 76.15 [7, p. 536], any irreducible $S^\lambda B$-module is isomorphic to $S^\lambda B e_j$ for some $j \in \{1, \ldots, n\}$. Moreover, by Proposition 5.22 [8, p. 112] and Theorem 76.8 [7, p. 532],

$$\overline{\operatorname{End}_{S^\lambda B} S^\lambda B e_j} \cong \operatorname{End}_{S^\lambda B} S^\lambda B e_j / X \operatorname{End}_{S^\lambda B} S^\lambda B e_j \cong \operatorname{End}_{\widetilde{S^\lambda B}} \widetilde{S^\lambda B}\varepsilon_j.$$

Lemma 6. *Let $G = G_p \times B$ and $\lambda \in Z^2(G, S^*)$. If V is an indecomposable $S^\lambda G_p$-module and W is an irreducible $S^\lambda B$-module, then*

$$\overline{\operatorname{End}_{S^\lambda G}(V \# W)} \cong \overline{\operatorname{End}_{S^\lambda G_p}(V)} \otimes_F \overline{\operatorname{End}_{S^\lambda B}(W)}.$$

P r o o f. By Proposition 7.6 [14, p. 652],

$$\mathrm{End}_{S^\lambda G}(V\#W) \cong \mathrm{End}_{S^\lambda G_p}(V) \otimes_S \mathrm{End}_{S^\lambda B}(W).$$

Applying Proposition 2 [6, p. 39], we obtain

$$\overline{\mathrm{End}_{S^\lambda G}(V\#W)} \cong \left(\overline{\mathrm{End}_{S^\lambda G_p}(V)} \otimes_F \overline{\mathrm{End}_{S^\lambda B}(W)}\right)/R,$$

where $R := \mathrm{rad}\left(\overline{\mathrm{End}_{S^\lambda G}(V)} \otimes_F \overline{\mathrm{End}_{S^\lambda B}(W)}\right)$. Since $\overline{\mathrm{End}_{S^\lambda B}(W)}$ is a separable F-algebra, then

$$\overline{\mathrm{End}_{S^\lambda G_p}(V)} \otimes_F \overline{\mathrm{End}_{S^\lambda B}(W)}$$

is a semisimple algebra. Hence $R = 0$ and the result follows. □

Lemma 7. *Let $G = G_p \times B$ and $\lambda \in Z^2(G, S^*)$. If F is a splitting field for the algebra $\widetilde{S^\lambda B}$, then $S^\lambda G$ is of OTP representation type.*

P r o o f. Let W be an irreducible $S^\lambda B$-module. Then

$$\overline{\mathrm{End}_{S^\lambda B} W} \cong \mathrm{End}_{\widetilde{S^\lambda B}} \widetilde{W} \cong F,$$

where $\widetilde{W} = W/XW$. By Lemmas 1 and 6, $V\#W$ is an indecomposable $S^\lambda G$-module for every indecomposable $S^\lambda G_p$-module V. By Lemma 5, $S^\lambda G$ is of OTP representation type. □

Lemma 8. *Let B be a finite p'-group. Assume that F contains a primitive q^{th} root of 1 for every prime $q \mid |B|$ such that $p \mid (q-1)$. Then, for any F-algebra $\widetilde{S^\lambda B}$, there exists a splitting field K such that $[K : F]$ is not divisible by p.*

P r o o f. See [2, p. 548]. □

Proposition 1. *Let $S = F[[X]]$, T be the quotient field of S, B a finite p'-group and $\lambda \in Z^2(B, S^*)$. The field T is a splitting field for the algebra $T^\lambda B$ if and only if F is a splitting field for the F-algebra $\widetilde{S^\lambda B}$.*

P r o o f. Assume that T is a splitting field for $T^\lambda B$. Denote by W an irreducible $S^\lambda B$-module. Since $T \otimes_S W$ is an absolutely irreducible $T^\lambda B$-module, by Schur's Lemma, $\mathrm{End}_{S^\lambda B}(W) \cong S$. It follows that

$$\mathrm{End}_{\widetilde{S^\lambda B}}(\widetilde{W}) \cong F. \tag{1}$$

Hence F is a splitting field for $\widetilde{S^\lambda B}$.

Now suppose that F is a splitting field for $\widetilde{S^\lambda B} = S^\lambda B/XS^\lambda B$. Then there exists an isomorphism (1) for any irreducible $S^\lambda B$-module W. It follows, by Theorem 76.8 [7, p. 532] and Corollary 76.16 [7, p. 536], that $\mathrm{End}_{S^\lambda B}(W) \cong S$, therefore $\mathrm{End}_{T^\lambda B}(T \otimes_S W) \cong T$. Hence T is a splitting field for $T^\lambda B$. □

3. Twisted group rings of OTP representation type

In this Section, $S = F[[X]]$ and $G = G_p \times B$, where G_p is a Sylow p-subgroup of G, $|G_p| \neq 1$ and $|B| \neq 1$. We assume that if G_p is non-Abelian, then F contains a primitive q^{th} root of 1 for every prime $q \mid |B|$ such that $p \mid (q-1)$.

Theorem 1. *Let $G = G_p \times B$, $\mu \in Z^2(G_p, S^*)$, $\nu \in Z^2(B, S^*)$, $\lambda = \mu \times \nu$ and $H = \mathrm{Ker}(\mu)$. Assume that $|H| > 2$. The ring $S^\lambda G$ is of OTP representation type if and only if F is a splitting field for $\widetilde{S^\nu B}$.*

P r o o f. If F is a splitting field for $\widetilde{S^\nu B}$, then, by Lemma 7, the ring $S^\lambda G$ is of OTP representation type.

Assume now that F is not a splitting field for $\widetilde{S^\nu B}$. There exists an irreducible $S^\nu B$-module W such that $D := \overline{\mathrm{End}_{S^\lambda B}(W)}$ is a division F-algebra of dimension greater than one. By [4, p. 268], the restriction of μ to $H \times H$ is a coboundary and $G_p' \subset H$. Suppose that G_p is non-Abelian. Then, by Lemma 8, there exists a splitting field K for $\widetilde{S^\nu B}$, which is a finite separable extension of the field F and satisfies $[K : F] \not\equiv 0 (\mathrm{mod}\, p)$. In view of Lemma 3, there is an indecomposable SH-module M such that $\overline{\mathrm{End}_{SH}(M)}$ is isomorphic to K. According to Lemma 2, we conclude that M^{G_p} is an indecomposable $S^\mu G_p$-module and

$$\overline{\mathrm{End}_{S^\mu G_p}(M^{G_p})}$$

is isomorphic to a field L that is a finite purely inseparable extension of the field K. Since L is a splitting field for D, $L \otimes_F D$ is not a skewfield. Hence, by Lemmas 1 and 6, $M^{G_p} \# W$ is not an indecomposable $S^\lambda G$-module. In view of Lemma 5, $S^\lambda G$ is not of OTP representation type.

The case, when G_p is Abelian, is treated similarly. □

Corollary. [2, p. 553] *Let $G = G_p \times B$, $|G_p'| > 2$ and $\lambda \in Z^2(G, S^*)$. The ring $S^\lambda G$ is of OTP representation type if and only if F is a splitting field for $\widetilde{S^\lambda B}$.*

P r o o f. Let μ be the restriction of λ to $G_p \times G_p$. Since $G_p' \subset \mathrm{Ker}(\mu)$, we have $|\mathrm{Ker}(\mu)| > 2$. Next apply Theorem 1. □

Proposition 2. *Let B be a nilpotent p'-group.*

(i) If the field F does not contain a primitive q^{th} root of 1 for some prime $q \mid |B|$, then F is not a splitting field for each algebra $F^\lambda B$.

(ii) The field F is a splitting field for all twisted group algebras $F^\lambda B$ if and only if $F = F^q$ and F contains a primitive q^{th} root of 1 for every prime $q \mid |B|$.

P r o o f. (i) Assume that F does not contain a primitive q^{th} root of 1 for some prime $q \mid |B|$. The center of a Sylow q-subgroup B_q of B contains an element b of order q. If $\{u_g \colon g \in B\}$ is a natural F-basis of the algebra $F^\lambda B$, then u_b lies in the center of $F^\lambda B$. Let $u_b^q = \gamma u_e$, $\gamma \in F^*$, and let F be a splitting field for the algebra $F^\lambda B$. Denote by $f_1, \ldots, f_m$ a complete system of minimal pairwise orthogonal central idempotents of $F^\lambda B$. We have $u_b = \beta_1 f_1 + \ldots + \beta_m f_m$, where $\beta_j \in F$ for any $j \in \{1, \ldots, m\}$. Then $\gamma = \beta_j^q$ for every j. It follows that $\beta_1 = \ldots = \beta_m$, hence $u_b = \beta_1 u_e$. This contradiction proves that F is not a splitting field for the algebra $F^\lambda B$.

(ii) Suppose that F is a splitting field for $F^\lambda B$ for each $\lambda \in Z^2(B, F^*)$. Then every irreducible projective F-representation of the group B is absolutely irreducible. Let q be a prime divisor of $|B|$. There exists a normal subgroup D of B such that $|B/D| = q$. Denote by $\pi \colon B \to B/D$ the canonical group homomorphism and by V a finite-dimensional vector space over F. If $\bar{\Gamma} \colon B/D \to \mathrm{GL}(V)$ is an irreducible projective F-representation of B/D on V, then $\Gamma := \bar{\Gamma} \circ \pi$ is an irreducible projective F-representation of B on the space V and $D \subset \mathrm{Ker}(\Gamma)$. Assume that $B/D = \langle\, bD \,\rangle$ and $\bar{\Gamma}(bD)^q = \gamma\, \mathrm{id}_V$, $\gamma \in F^*$. Since every $\bar{\Gamma}$ is absolutely irreducible, $\gamma \in F^q$ and F contains a primitive q^{th} root of 1.

Assume now that the field F contains a primitive q^{th} root of 1 and $F = F^q$ for each prime $q \mid |B|$. Let $\lambda \in Z^2(B, F^*)$. Then $F^\lambda B = F^\mu B$, where $\mu_{x,y}^{|B|} = 1$ for all $x, y \in B$. There exists an F-algebra homomorphism of FH onto $F^\mu B$, where H is a central extension of a cyclic group of order $|B|$ by the group B. Since F contains a primitive $|H|^{\text{th}}$ root of 1, by Corollary 70.24 [7, p. 475], F is a splitting field for FH. Hence, F is a splitting field for $F^\lambda B$ for each $\lambda \in Z^2(B, F^*)$. □

Proposition 3. *Let $G = G_p \times B$, $|G_p'| \neq 2$, $\mu \in Z^2(G_p, F^*)$, $\nu \in Z^2(B, S^*)$ and $\lambda = \mu \times \nu$. The ring $S^\lambda G$ is of OTP representation type if and only if one of the following conditions is satisfied:*

(i) $F^\mu G_p$ is a field;

(ii) $p = 2$, $|G_2'| = 1$ and $2 \dim_F \overline{F^\mu G_2} = |G_2|$;

(iii) F is a splitting field for the F-algebra $\widetilde{S^\nu B}$.

P r o o f. If $|G_p'| > 2$ then, by Corollary to Theorem 1, the ring $S^\lambda G$ is of OTP representation type if and only if F is a splitting field for $\widetilde{S^\nu B}$. Let $|G_p'| = 1$ and $K = F^\mu G_p$. If K is a field, then $S^\mu G_p = K[[X]]$ is a principal ideal ring. Every indecomposable $S^\mu G_p$-module is isomorphic to $S^\mu G_p$. We have

$$\overline{\mathrm{End}_{S^\mu G_p}(S^\mu G_p)} \cong S^\mu G_p / X S^\mu G_p \cong K.$$

The field K is a finite purely inseparable extension of F. Let W be an irreducible $S^\nu B$-module and $D := \overline{\mathrm{End}_{S^\nu B}(W)}$. Then $D \cong \mathrm{End}_{\widetilde{S^\nu B}}(\widetilde{W})$. Since $\widetilde{S^\nu B}$ is a separable algebra, the center of the division F-algebra D is a separable extension of F [7, p. 485]. The index of D is not divisible by p [16]. It follows that $K \otimes_F D$ is a skewfield. Applying Lemmas 1 and 6, we conclude that $S^\mu G_p \# W$ is an indecomposable $S^\lambda G$-module. Hence, by Lemma 5, $S^\lambda G$ is of OTP representation type.

Assume that $p > 2$ and K is not a field. Let H be the socle of G_p. We have $F^\mu H \cong F^\mu H_1 \otimes_F F H_2$, where $|H_2| \geq p$. It follows that $S^\mu H \cong S^\mu H_1 \otimes_S S H_2$. By Lemmas 2 and 3, for any finite separable extension L of the field F, there exists an indecomposable $S^\mu G_p$-module V such that $\overline{\mathrm{End}_{S^\mu G_p}(V)}$ is a finite purely inseparable extension of L. Arguing as in the proof of Theorem 1, we conclude that $S^\lambda G$ is of OTP representation type if and only if F is a splitting field for the algebra $\widetilde{S^\nu B}$.

Suppose that $p = 2$ and K is not a field. If $4 \dim_F \overline{F^\mu G_2} \leq |G_2|$ then, as in the case $p > 2$, we prove that $S^\lambda G$ is of OTP representation type if and only if F is a splitting field for the algebra $\widetilde{S^\nu B}$. If $2 \dim_F \overline{F^\mu G_2} = |G_2|$ then, by Theorem 4.2 [2, p. 552], the ring $S^\lambda G$ is of OTP representation type. □

Corollary. *Let G_p be an Abelian p-group, B a nilpotent p'-group, $G = G_p \times B$, $\mu \in Z^2(G_p, F^*)$, $\nu \in Z^2(B, S^*)$ and $\lambda = \mu \times \nu$. Assume that the field F does not contain a primitive q^{th} root of 1 for some prime $q \mid |B|$. The ring $S^\lambda G$ is of OTP representation type if and only if one of the following conditions is satisfied:*

(i) $F^\mu G_p$ is a field;

(ii) $p = 2$ and $2 \dim_F \overline{F^\mu G_2} = |G_2|$.

P r o o f. Apply Propositions 2 and 3. □

Proposition 4. *Let $p = 2$, $G = G_2 \times B$, $\mu \in Z^2(G_2, F^*)$, $\nu \in Z^2(B, S^*)$ and $\lambda = \mu \times \nu$. Assume that $|G_2| \neq 8$, $|G_2'| = 2$ and $[F : F^2] \leq 2$. Then $S^\lambda G$ is of OTP representation type if and only if F is a splitting field for $\widetilde{S^\nu B}$.*

P r o o f. If F is a perfect field, then μ is a coboundary [15, p. 43]. In this case $S^\mu G_2$ is the group ring $S G_2$. Since $|G_2| > 8$, by Theorem 1, $S^\lambda G$ is of OTP representation type if and only if F is a splitting field for $\widetilde{S^\nu B}$. Assume

now that $[F : F^2] = 2$. Arguing as in the proof of Theorem 1, we deduce, by Lemmas 1, 4, 5, 6 and 7, that $S^\lambda G$ is of OTP representation type if and only if F is a splitting field for $\widetilde{S^\nu B}$. □

4. Groups of OTP projective representation type

We recall from [3, p. 200] that $i(F)$ is the supremum of the set that consists of 0 and all positive integers m such that an F-algebra of the form

$$F[t]/(t^p - \alpha_1) \otimes_F \ldots \otimes_F F[t]/(t^p - \alpha_m)$$

is a field for some $\alpha_1, \ldots, \alpha_m \in K$.

Theorem 2. *Let $G = G_p \times B$, $|G_p'| \neq 2$ and s be the number of invariants of G_p/G_p'. The group G is of OTP projective (S, F^*)-representation type if and only if one of the following conditions is satisfied:*

(i) $|G_p'| = 1$ and $s \leq i(F)$;

(ii) $p = 2$, $|G_2'| = 1$, $s = i(F) + 1$ and G_2 has at least one invariant equal to 2;

(iii) F is a splitting field for $F^\sigma B$ for some $\sigma \in Z^2(B, F^*)$.

P r o o f. Let $p = 2$ and G_2 be Abelian. If $s \geq i(F) + 2$, then $4 \dim_F \overline{F^\lambda G_2} \leq |G_2|$ for any $\lambda \in Z^2(G_2, F^*)$. In this case, by Proposition 3, G is of OTP projective (S, F^*)-representation type if and only if the condition (iii) is satisfied. Assume that $s = i(F) + 1$. If G_2 has at least one invariant equal to 2, then there exists a cocycle $\lambda \in Z^2(G_2, F^*)$ such that $2 \dim_F \overline{F^\lambda G_2} = |G_2|$. Hence, by Proposition 3, G is of OTP projective (S, F^*)-representation type. Suppose that every invariant of G_2 is greater than 2. Then $4 \dim_F \overline{F^\lambda G_2} \leq |G_2|$ for each $\lambda \in Z^2(G_2, F^*)$. By Proposition 3, G is of OTP projective (S, F^*)-representation type if and only if the condition (iii) is satisfied.

Let $p \geq 2$ and G_p be Abelian. There exists a cocycle $\mu \in Z^2(G_p, F^*)$ such that $F^\mu G_p$ is a field if and only if $s \leq i(F)$. For any $\nu \in Z^2(B, F^*)$, we have $\widetilde{S^\nu B} \cong F^\nu B$. Applying Proposition 3, we finish the proof. □

Corollary. *Let G_p be an Abelian p-group, s the number of invariants of G_p, B a nilpotent p'-group and $G = G_p \times B$. Assume that the field F does not contain a primitive q^{th} root of 1 for some prime $q \mid |B|$. The group G is of OTP projective (S, F^*)-representation type if and only if one of the following conditions is satisfied:*

(i) $s \leq i(F)$;

(ii) $p = 2$, $s = i(F) + 1$ and G_2 has at least one invariant equal to 2.

P r o o f. Apply Proposition 2 and Theorem 2. □

Lemma 9. *Let B be an Abelian p'-group. The field F is a splitting field for some algebra $F^\lambda B$ if and only if B has a subgroup H such that B/H is of symmetric type and F contains a primitive m^{th} root of 1, where $m = \max\{\exp(B/H), \exp H\}$.*

P r o o f. Let $\lambda \in Z^2(B, F^*)$, $\{u_b \colon b \in B\}$ be a natural F-basis of the algebra $F^\lambda B$, Z the center of $F^\lambda B$ and $H = \{g \in B \colon u_g \in Z\}$. Then H is a subgroup of B and $Z = F^\lambda H$. The algebra $F^\lambda B$ may be viewed as a twisted group ring of the group $\bar{B} := B/H$ over the ring Z. By Lemma 3 [1, p. 785],

$$F^\lambda B = Z^{\bar{\lambda}}\bar{B} \cong Z^{\bar{\lambda}}N_1 \otimes_Z \ldots \otimes_Z Z^{\bar{\lambda}}N_r,$$

where N_i is a group of type $(q_i^{n_i}, q_i^{n_i})$, q_i is a prime divisor of $|\bar{B}|$ and $Z^{\bar{\lambda}}N_i$ is a central Z-algebra, moreover

$$\gamma_{x,y} := \bar{\lambda}_{x,y} \cdot \bar{\lambda}_{y,x}^{-1} \in F$$

and

$$\gamma_{x,y}^{q_i^{n_i}} = 1$$

for all $x, y \in N_i$. It follows that F contains a primitive $(\exp \bar{B})^{\text{th}}$ root of 1.

If F is a splitting field for $F^\lambda B$, then F is a splitting field for the commutative F-algebra $Z = F^\lambda H$. Therefore F contains a primitive $(\exp H)^{\text{th}}$ root of 1. The group $\bar{B} = N_1 \times \ldots \times N_r$ is of symmetric type. This proves the necessity.

Let us prove the sufficiency. Denote by K a finite subfield of the field F which contains a primitive m^{th} root of 1, where $m = \max\{\exp(B/H), \exp H\}$. We may assume that B is an Abelian q-group, where $q \neq p$. Let

$$\bar{B} := B/H = \langle x_1H \rangle \times \langle y_1H \rangle \times \ldots \times \langle x_rH \rangle \times \langle y_rH \rangle,$$

where $|x_iH| = |y_iH| = q^{n_i}$ for each $i \in \{1, \ldots, r\}$. We have

$$x_i^{q^{n_i}} = h_i, \quad y_i^{q^{n_i}} = h_i^*,$$

where $h_i, h_i^* \in H$. Let $Z = KH$ with K-basis $\{u_h \colon h \in H\}$ and let $A = Z^\mu \bar{B}$ be the twisted group ring of $\bar{B}$ over Z with Z-basis $\{v_{bH} \colon b \in B\}$ satisfying the following conditions:

1) if $bH = (x_1H)^{i_1}(y_1H)^{j_1} \ldots (x_rH)^{i_r}(y_rH)^{j_r}$, where $0 \leq i_s, j_s < q^{n_s}$, then

$$v_{bH} = v_{x_1H}^{i_1} v_{y_1H}^{j_1} \ldots v_{x_rH}^{i_r} v_{y_rH}^{j_r};$$

2) $v_{x_sH}^{q^{n_s}} = u_{h_s}$, $v_{y_sH}^{q^{n_s}} = u_{h_s^*}$ for all $s \in \{1, \ldots, r\}$;

3) $v_{bH} \cdot v_{\bar{b}H} = \xi_1^{j_1\bar{i}_1} \ldots \xi_r^{j_r\bar{i}_r} v_{x_1H}^{i_1+\bar{i}_1} v_{y_1H}^{j_1+\bar{j}_1} \ldots v_{x_rH}^{i_r+\bar{i}_r} v_{y_rH}^{j_r+\bar{j}_r}$,

where ξ_s is a primitive $(q^{n_s})^{\text{th}}$ root of 1 for every $s \in \{1, \dots, r\}$. Then

$$A \cong Z^\mu N_1 \otimes_Z \dots \otimes_Z Z^\mu N_r,$$

where $Z^\mu N_s$ is a central twisted group ring of the group $N_s = \langle\, x_s H \,\rangle \times \langle\, y_s H \,\rangle$ over the ring Z.

Let g be an element of the group B. Then

$$g = x_1^{d_1} y_1^{t_1} \dots x_r^{d_r} y_r^{t_r} h,$$

where $0 \le d_s, t_s < q^{n_s}$ for every $s \in \{1, \dots, r\}$ and $h \in H$. We set

$$w_g = v_{x_1 H}^{d_1} v_{y_1 H}^{t_1} \dots v_{x_r H}^{d_r} v_{y_r H}^{t_r} u_h.$$

Then $\{w_g \colon g \in B\}$ is a K-basis of the algebra A and $w_{g_1} w_{g_2} = \lambda_{g_1, g_2} w_{g_1 g_2}$, where $\lambda_{g_1, g_2} \in K^*$ for all $g_1, g_2 \in B$. Hence $A = K^\lambda B$ and K is a splitting field for the algebra $K^\lambda B$. It follows that F is a splitting field for the algebra $F^\lambda B = F \otimes_K K^\lambda B$. □

Lemma 10. *Let B be an Abelian p'-group of symmetric type and $\exp B = q_1^{m_1} \dots q_t^{m_t}$, where $q_1, \dots, q_t$ are pairwise distinct prime numbers. The field F is a splitting field for certain algebra $F^\lambda B$ if and only if F contains a primitive n^{th} root of 1, where $n = q_1^{k_1} \dots q_t^{k_t}$ and $2k_j \ge m_j$ for every $j \in \{1, \dots, t\}$.*

P r o o f. Without loss of generality, we may assume that B is an Abelian q-group of exponent q^m. Let F contain a primitive $(q^l)^{\text{th}}$ root of 1 and F does not contain a primitive $(q^{l+1})^{\text{th}}$ root of 1. If $l \ge m$ then F is a splitting field for the group algebra FB. Let $\frac{m}{2} \le l < m$. The group B has a subgroup H of exponent q^{m-l} such that B/H is of symmetric type and $\exp(B/H) = q^l$. Since $m - l \le l$, by Lemma 9, F is a splitting field for certain algebra $F^\nu B$. Suppose now that $l < \frac{m}{2}$. Let $\lambda \in Z^2(B, F^*)$, Z be the center of $F^\lambda B$ and H a subgroup of B such that $Z = F^\lambda H$. Then $\exp H \ge q^{m-l}$. If F is a splitting field for $F^\lambda B$, then $\exp H \le q^l$. We have $q^{m-l} \le q^l$, whence $m - l \le l$. Hence $l \ge \frac{m}{2}$. This contradiction shows that F is not a splitting field for every algebra $F^\lambda B$. □

Proposition 5. *Let $G = G_p \times B$ be an Abelian group and s the number of invariants of G_p. The group G is of OTP projective (S, F^*)-representation type if and only if one of the following conditions is satisfied:*

(i) $s \le i(F)$;

(ii) $p = 2$, $s = i(F) + 1$ and G_2 has at least one invariant equal to 2;

(iii) B has a subgroup H such that B/H is of symmetric type and F contains a primitive m^{th} root of 1, where $m = \max\{\exp(B/H), \exp H\}$.

P r o o f. Apply Theorem 2 and Lemma 9. □

Proposition 6. *Let $G = G_p \times B$ be an Abelian group and s the number of invariants of G_p. Assume that B is of symmetric type and $\exp B = q_1^{m_1} \dots q_t^{m_t}$, where $q_1, \dots, q_t$ are pairwise distinct prime numbers. The group G is of OTP projective (S, F^*)-representation type if and only if one of the following conditions is satisfied:*

(i) $s \leq i(F)$;

(ii) $p = 2$, $s = i(F) + 1$ and G_2 has at least one invariant equal to 2;

(iii) F contains a primitive n^{th} root of 1, where $n = q_1^{k_1} \dots q_t^{k_t}$ and $2k_j \geq m_j$ for each $j \in \{1, \dots, t\}$.

P r o o f. Apply Theorem 2 and Lemma 10. □

5. Groups of purely OTP projective representation type

Lemma 11. [5, p. 322] *Let R be a Noetherian integral domain whose integral closure is a finitely generated R-module. Then every finitely generated torsion free R-module is a direct sum of ideals in R if and only if each ideal in R is generated by one or two elements.*

Theorem 3. *Let $G = G_p \times B$. The group G is of purely OTP projective S-representation type if and only if $|G_p| = 2$ or F is a splitting field for $F^\nu B$ for any $\nu \in Z^2(B, F^*)$.*

P r o o f. Assume that $|G_p| > 2$ and $\sigma \in Z^2(B, S^*)$. By Theorem 1, the ring $S^\lambda G = SG_p \otimes_S S^\sigma B$ is of OTP representation type if and only if F is a splitting field for $\widetilde{S^\sigma B}$. Hence, by Lemma 7, if $|G_p| > 2$ then G is of purely OTP projective S-representation type if and only if F is a splitting field for every algebra $F^\nu B$.

Let $p = 2$ and $G_2 = \langle a \rangle$ be the group of order 2. If V is an indecomposable SG_2-module then, by [13, p. 70], $\overline{\mathrm{End}_{SG_2}(V)} \cong F$. Hence, by Lemmas 1, 5 and 6, the ring $SG_2 \otimes_S S^\nu B$ is of OTP representation type for any $\nu \in Z^2(B, S^*)$. Suppose now that $\lambda \in Z^2(G, S^*)$ and $S^\lambda G_2$ is not a group ring. Then $S^\lambda G_2 = Su_e + Su_a$, where $u_a^2 = f(X)u_e$, $f(X) \in S^*$ and $f(X) \notin S^2$. Let $f(X) = a_0 + a_1 X + a_2 X^2 + \dots$, where $a_j \in F$ for every $j \in \{0, 1, 2, \dots\}$, θ be a root of the polynomial $t^2 - f(X)$ and $K = T(\theta)$, where T is the quotient field of S. We have $S^\lambda G_2 \cong S[\theta]$. Denote by L the integral closure of $S[\theta]$ in the field K. Then $L = S[\omega]$, where $\omega = \theta$ or $\omega = X^{-n}(b_0 + b_1 X + \dots + b_{n-1}X^{n-1} + \theta)$, moreover in the second case

$$f(X) = b_0^2 + b_1^2 X^2 + \dots + b_{n-1}^2 X^{2(n-1)} + \sum_{j \geq 2n} a_j X^j,$$

$n \geq 1$, $a_{2n} \notin F^2$ or $a_{2n+1} \neq 0$.

Every ideal of the ring $S[\theta]$ is generated by one or two elements. Let V be an indecomposable $S[\theta]$-module. If $z \in S[\theta]$, $v \in V$ and $zv = 0$, then $z^2 v = 0$. Since $z^2 \in S$ and V is a free S-module, $z^2 = 0$ or $v = 0$. Hence $z = 0$ or $v = 0$. This means that V is a torsion-free $S[\theta]$-module. By Lemma 11, V is isomorphic to an ideal J of the ring $S[\theta]$. The ideal J is a free S-module of rank 2. It follows that $T \otimes_S J$ is an indecomposable $T^\lambda G_2$-module. By Theorem 3.1 [2, p. 549], the algebra $T^\lambda G$ is of OTP representation type. Therefore, $(T \otimes_S J)\#(T \otimes_S W)$ is an indecomposable $T^\lambda G$-module for any irreducible $S^\lambda B$-module W. It follows that $J\#W$ is an indecomposable $S^\lambda G$-module. By Lemma 5, the ring $S^\lambda G$ is of OTP representation type. Hence, G is of purely OTP projective S-representation type. □

Corollary. *Let $G = G_p \times B$ be a nilpotent group. The group G is of purely OTP projective S-representation type if and only if one of the following conditions is satisfied:*

(i) $|G_p| = 2$;

(ii) $F = F^q$ and F contains a primitive q^{th} root of 1 for each prime $q \mid |B|$.

P r o o f. Apply Proposition 2 and Theorem 3. □

Proposition 7. *Let $G = G_p \times B$. Assume that $F = F^q$ and F contains a primitive q^{th} root of 1 for each prime $q \mid |B|$. Then G is of purely OTP projective S-representation type.*

P r o o f. The field F is a splitting field for any algebra $F^\nu B$. Hence, by Lemma 7, $S^\lambda G$ is of OTP representation type for every $\lambda \in Z^2(G, S^*)$. □

Corollary. *If F is a separably closed field, then every group $G = G_p \times B$ is of purely OTP projective S-representation type.*

References

[1] L.F. Barannyk. On the question of faithful projective representations of finite Abelian groups over arbitrary field. *Ukrain. Math. J.,* **26**, No. 6, 784–790, 1974.

[2] L.F. Barannyk. Modular projective representations of direct products of finite groups. *Publ. Math. Debrecen,* **63** (4), 537–554, 2003.

[3] L.F. Barannyk. On faithful irreducible projective representations of finite groups over a field of characteristic p. *J. Algebra,* **321**, 194–204, 2009.

[4] L.F. Barannyk, D. Klein. Twisted group rings of strongly unbounded representation type. *Colloq. Math.,* **100** (2), 265–287, 2004.

[5] H. Bass. Torsion free and projective modules. *Trans. Amer. Math. Soc.,* **102** (2), 319–327, 1962.

[6] H.I. Blau. Indecomposable modules for direct products of finite groups. *Pacific J. Math.,* **54** (1), 39–44, 1974.

[7] C.W. Curtis, I. Reiner. *Representation Theory of Finite Groups and Associative Algebras.* Interscience, New York, 1962.

[8] C.W. Curtis, I. Reiner. *Methods of Representation Theory with Applications to Finite Groups and Orders*, Vol. 1. Willey, New York, 1981.

[9] J.A. Green. On the indecomposable representations of a finite group. *Math. Z.,* **70**, 430–445, 1959.

[10] P.M. Gudyvok. On modular and integral representations of finite groups. *Dokl. Akad. Nauk SSSR,* **214**, No. 5, 993–996, 1974. (In Russian).

[11] P.M. Gudyvok. On modular and integral P-adic representations of direct product of groups. *Ukrain. Math. J.,* **29**, No. 5, 580-588, 1977. (In Russian).

[12] P.M. Gudyvok. On representations of direct product of groups over complete discrete valuation rings. *Dokl. Akad. Nauk SSSR,* **237**, No. 1, 25–27, 1977. (In Russian).

[13] P.M. Gudyvok. On representations of a direct product of finite groups over complete discrete valuation rings. *Ukrain. Math. Bull.,* **2**, No. 1, 67–75, 2005.

[14] G. Karpilovsky. *Group Representations*, Vol. 1. North-Holland Mathematics Studies 175, North-Holland, Amsterdam, 1992.

[15] G. Karpilovsky. *Group Representations*, Vol. 2. North-Holland Mathematics Studies 177, North-Holland, Amsterdam, 1993.

[16] H.N. Ng. Degrees of irreducible representations of finite groups. *J. London Math. Soc.,* (**2**)**10**, 379–384, 1975.

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XIV, Częstochowa 2010

A CHARACTERIZATION OF A HOMOGRAPHIC TYPE FUNCTION

Katarzyna Domańska

Institute of Matematics and Computer Science
Jan Długosz University in Częstochowa
Armii Krajowej 13/15, 42 – 201 *Częstochowa, Poland*
e-mail: k.domanska@ajd.czest.pl

Abstract. We deal with a functional equation of the form

$$f(x+y) = F(f(x), f(y))$$

(the so called addition formula) assuming that the given binary operation F is associative but its domain of definition is not necessarily connected. In the present paper we shall restrict our consideration to the case when

$$F(u,v) = \frac{u+v+2uv}{1-uv}.$$

These considerations may be viewed as counterparts of Losonczi's [7] and Domańska's [3] results on local solutions of the functional equation

$$f(F(x,y)) = f(x) + f(y)$$

with the same behaviour of the given associative operation F. In this paper we admit fairly general structure in the domain of the unknown function.

1. Introduction

If $(G, \star)$ is a group or a semigroup and F stands for an arbitrary binary operation in some set H, then a solution of the functional equation

$$f(x \star y) = F(f(x), f(y))$$

is called a homomorphism of structures $(G, \star)$ and (H, F). We consider here a rational function $F : \{(x,y) \in \mathbb{R} : \quad xy \neq 1\} \longrightarrow \mathbb{R}$ of the form

$$F(u,v) = \frac{u+v+2uv}{1-uv}.$$

This is a rational two-place real-valued function defined on a disconnected subset of the real plane $\mathbb{R}^2$, that satisfies the equation

$$F(F(x,y),z) = F(x,F(y,z))$$

for all $(x,y,z) \in \mathbb{R}^3$ such that products $xy, yz, F(x,y)z, xF(y,z)$ are not equal to 1. Rational functions with such or similar properties are termed associative operations. The class of the associative operations was described by Chéritat [2], and his work was followed by the author.

A homografic function $\varphi : \mathbb{R} \setminus \{1\} \longrightarrow \mathbb{R}$ given by the formula

$$\varphi(x) = \frac{x}{1-x}, \quad x \neq 1$$

satisfies the functional equation

$$f(x+y) = \frac{f(x)+f(y)+2f(x)f(y)}{1-f(x)f(y)}$$

for every pair $(x,y) \in \mathbb{R}^2 \setminus D$, where

$$D = \{(x,1-x) : \quad x \in \mathbb{R}\} \cup \{(x,1) : \quad x \in \mathbb{R}\} \cup \{(1,x) : \quad x \in \mathbb{R}\}.$$

We shall determine all functions $f : G \longrightarrow \mathbb{R}$, where $(G,\star)$ is a group, that satisfy the functional equation

$$f(x \star y) = \frac{f(x)+f(y)+2f(x)f(y)}{1-f(x)f(y)}. \tag{1}$$

A neutral element of a group $(G,\star)$ will be written as 0.

By a solution of the functional equation (1) we understand any function $f : G \longrightarrow \mathbb{R}$ that satisfies the equality (1) for every pair $(x,y) \in G^2$ such that $f(x)f(y) \neq 1$. Thus we deal with the following conditional functional equation:

$$f(x)f(y) \neq 1 \quad \text{implies} \quad f(x \star y) = \frac{f(x)+f(y)+2f(x)f(y)}{1-f(x)f(y)} \tag{E}$$

for all $x, y \in G$. Some results on addition formulas can be found for example in Aczél's monography [1] and in the work of Domańska and Ger [4].

The following lemma will be useful in the sequel (see Ger [6]).

Lemma (on a characterization on subgroups). *Let $(G,+)$ be a group. Then $(H,+)$ is a subgroup of group $(G,+)$ if and only if $G \supset H \neq \emptyset$ and*

$$H + H' \subset H',$$

where $H' := G \setminus H$.

2. Main result

We proceed with a description of solutions of (E).

Theorem. *Let $(G,\star)$ be a group. A function $f : G \longrightarrow \mathbb{R}$ yields a nonconstant solution to the functional equation*

$$f(x)f(y) \neq 1 \quad \textit{implies} \quad f(x\star y) = \frac{f(x)+f(y)+2f(x)f(y)}{1-f(x)f(y)} \tag{E}$$

for all $x, y \in G$ if and only if either

$$f(x) := \begin{cases} 1 & \textit{for } x \in H, \\ -1 & \textit{for } x \in G \setminus H \end{cases}$$

or

$$f(x) := \begin{cases} \dfrac{A(x)}{1-A(x)} & \textit{for } x \in \Gamma \\ -1 & \textit{for } x \in G \setminus \Gamma \end{cases}$$

or

$$f(x) := \begin{cases} 1 & \textit{for } x \in \Gamma \setminus Z \\ 0 & \textit{for } x \in Z \\ -1 & \textit{for } x \in G \setminus \Gamma, \end{cases}$$

where $(H,\star), (\Gamma,\star)$ are subgroups of the group $(G,\star)$, $(Z,\star)$ is a subgroup of the group $(\Gamma,\star)$, and $A : \Gamma \longrightarrow \mathbb{R}$ is a homomorphism such that $1 \notin A(\Gamma)$.

Proof. Assume that f is a nonconstant solution of equation (E). First we show that $f(0) \in \{-1, 0, 1\}$. Indeed, setting $x = y = 0$ in (E), we obtain

$$f^2(0) = 1 \quad \text{or} \quad f(0) = \frac{2f(0)+2f^2(0)}{1-f^2(0)}.$$

Put $c := f(0)$. By equality

$$c = 2c\,\frac{1+c}{1-c^2}$$

we have $c = 0 \quad \text{or} \quad 2(1+c) = 1-c^2$, whence $c \in \{0, -1\}$ which jointly with the equality $c^2 = 1$ implies $f(0) \in \{-1, 0, 1\}$, which was to be shown.

If $f(0) = -1$, then setting $y = 0$ in (E) we obtain

$$f(x) = -1 \quad \text{or} \quad f(x) = \frac{f(x)-1-2f(x)}{1+f(x)} = -1$$

for all $x \in G$, whence $f = -1$, a contradiction because we were assuming f to be nonconstant.

Now assume that $f(0) = 1$. We show that $f(G) \subset \{-1, 1\}$. Indeed, putting $y = 0$ in (E), we obtain

$$f(x) = 1 \quad \text{or} \quad f(x) = \frac{3f(x) + 1}{1 - f(x)}$$

for all $x \in G$ and by the equality

$$c = \frac{3c + 1}{1 - c}$$

we have $c = -1$, whence

$$f(x) = 1 \quad \text{or} \quad f(x) = -1$$

for all $x \in G$. By setting

$$H := \{x \in G : \quad f(x) = 1\},$$

we have

$$H' = \{x \in G : \quad f(x) = -1\}$$

and we show that $H \star H' \subset H'$, which implies that H is a subgroup of the group G (see Lemma). Fix arbitrarily elements $x \in H$ and $y \in H'$. Since $f(x)f(y) = -1$, we get by (E) $f(x \star y) = -1$, i.e. $x \star y \in H'$, which was to be shown. So, in this case we have

$$f(x) := \begin{cases} 1 & \text{for } x \in H, \\ -1 & \text{for } x \in G \setminus H. \end{cases}$$

Let now $f(0) = 0$. Put

$$\Gamma := \{x \in G : \quad f(x) \neq -1\}.$$

We are going to show that the complement Γ' of the set Γ enjoys the property $\Gamma \star \Gamma' \subset \Gamma'$, which implies (see Lemma) that Γ is a subgroup of the group G. Fix arbitrarily $x \in \Gamma$ and a $y \in \Gamma'$. Since $f(x)f(y) = -f(x) \neq 1$, we obtain by (E)

$$f(x \star y) = \frac{f(x) - 1 + 2f(x)(-1)}{1 - f(x)(-1)} = \frac{-1 - f(x)}{1 + f(x)} = -1,$$

i.e. $x \star y \in \Gamma'$, which was to be shown. Since $-1 \notin f(\Gamma)$ and $f|_\Gamma$ satisfies (E), a straightforward verification shows that

$$f(x)f(y) \neq 1 \quad \text{implies} \quad \frac{f(x \star y)}{1 + f(x \star y)} = \frac{f(x)}{1 + f(x)} + \frac{f(y)}{1 + f(y)}$$

for all $x, y \in \Gamma$, which jointly with

$$1 - \frac{f(x)}{1+f(x)} - \frac{f(y)}{1+f(y)} = 1 - \frac{f(x)+2f(x)f(y)+f(y)}{(1+f(x))(1+f(y))}$$

$$= \frac{1-f(x)f(y)}{(1+f(x))(1+f(y))},$$

i.e.

$$f(x)f(y) = 1 \Longleftrightarrow \frac{f(x)}{1+f(x)} + \frac{f(y)}{1+f(y)} = 1,$$

states that the function $A : \Gamma \longrightarrow \mathbb{R}$ of the form

$$A(x) := \frac{f(x)}{1+f(x)}, \quad x \in \Gamma$$

yields a solution of equation

$$A(x) + A(y) \neq 1 \quad \text{implies} \quad A(x+y) = A(x) + A(y) \tag{2}$$

for all $x, y \in \Gamma$. We show that $1 \notin A(\Gamma)$. To prove this, assume that $A(x_0) = 1$ for some $x_0 \in \Gamma$. Then we conclude that $f(x_0) = 1 + f(x_0)$, which is immposible. Since $f(0) = 0$, evidently $A(0) = 0$. From the theorem proved by Ger [5] (since $A(0) = 0$) we conclude that A yields a homomorphism of groups Γ and $\mathbb{R}$ or there exist a subgroup Z of a group Γ such that A is of the form

$$A(x) := \begin{cases} 0 & \text{for } x \in Z, \\ \frac{1}{2} & \text{for } x \in \Gamma \setminus Z, \end{cases}$$

whence

$$f(x) := \begin{cases} \dfrac{A(x)}{1-A(x)} & \text{for } x \in \Gamma, \\ -1 & \text{for } x \in G \setminus \Gamma \end{cases}$$

or

$$f(x) := \begin{cases} 0 & \text{for } x \in Z, \\ 1 & \text{for } x \in \Gamma \setminus Z, \\ -1 & \text{for } x \in G \setminus \Gamma. \end{cases}$$

It is easy to check that each of the functions above yields a solution to the equation (E). Thus the proof has been completed.

The following remark gives the form of a constant solutions of equation (E).

Remark. *Let $(G,\star)$ be a group. The only constant solutions of* (E) *are* $f=-1$, $f=0$, *and* $f=1$.

To check this, assume that $f=c$ fulfils (E). Then

$$c^2\neq 1 \Longrightarrow c=2c\frac{1+c}{1-c^2},$$

i.e.

$$c\in\{-1,1\} \quad \text{or} \quad c=0 \quad \text{or} \quad c=2\frac{1+c}{1-c^2},$$

whence

$$c\in\{-1,0,1\},$$

which was to be shown.

References

[1] J. Aczél. *Lectures on Functional Equations and Their Applications.* Academic Press, New York, 1966.

[2] A. Chéritat. Fractions rationnelles associatives et corps quadratiques, *Rev. Math. de l'Enseignement Supérieur*, **109**, 1025–1040, 1998-1999.

[3] K. Domańska. Cauchy type equations related to some singular associative operations. *Glasnik Matematički*, 31(51), 135–149, 1996.

[4] K. Domańska, R. Ger. Addition formulae with singularities. *Ann. Math. Silesianae*, **18**, 7–20, 2004.

[5] R. Ger. On some functional equations with a restricted domain, II. *Fund. Math.*, **98**, 249–272, 1978.

[6] R. Ger. O pewnych równaniach funkcyjnych z obciętą dziedziną. *Prace Naukowe Uniwersytetu Śląskiego*, Nr 132, Katowice, 1976.

[7] L. Losonczi. Local solutions of functional equations, *Glasnik Matematički*, 25(45), 57–67, 1990.

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XV, Częstochowa 2010

RIEMANN INTEGRABILITY AND QUASI-UNIFORM CONVERGENCE

Roman Ger[a], **Jacek Jędrzejewski**[b]

[a] *Institute of Mathematics*
Silesian University
ul. Bankowa 14, 40-007 Katowice, Poland
e-mail: romanger@us.edu.pl

[b] *Institute of Mathematics and Computer Science*
Jan Długosz University in Częstochowa
al. Armii Krajowej 13/15, 42-200 Częstochowa, Poland
e-mail: jacek.m.jedrzejewski@gmail.com

Abstract. We consider quasi-uniform convergence of sequences of functions in a context of Riemann integrability of its limit. Some generalizations are discussed as well.

Arzelá considered an additional regularity condition for a pointwise convergent sequence of functions. Its role is particularly interesting while dealing with convergence in the space of continuous functions. The precise definition reads as follows.

Definition 1. (Arzelá [1]. *A pointwise convergent sequence* $(f_n)_{n=1}^{\infty}$ *of real functions defined in a topological space* X *is called quasi-uniformly convergent to a function* $f : X \longrightarrow \mathbb{R}$ *if*

$$\forall_{\varepsilon>0}\forall_{n\in\mathbb{N}}\exists_{k_n}\exists_{p_1,\dots,p_{k_n}\geq n}\forall_{t\in X}\left(\min\left\{|f_{p_i}(t)-f(t)| : i\in\{1,\dots,k_n\}\right\}<\varepsilon\right). \quad (1)$$

The following two facts concerning this convergence are well known.

Fact 1. (Szökefalvi-Nagy [2]). *Assume that a sequence of continuous real functions defined in a topological space* X *is quasi-uniformly convergent to a function* $f : X \longrightarrow \mathbb{R}$. *Then* f *itself is continuous.*

Fact 2. *Let X be a compact topological space and let $(f_n)_{n=1}^{\infty}$ be a pointwise convergent sequence of continuous real functions defined in X. If the limit function $f : X \longrightarrow \mathbb{R}$ is continuous as well, then the sequence $(f_n)_{n=1}^{\infty}$ is quasi-uniformly convergent to f.*

In what follows, we shall consider a locally compact topological group $(G, +)$ with Haar measure h. It turns out that, in such circumstances, Fact 1 carries over to functions that are merely h-almost everywhere continuous. Namely, the following theorem holds true.

Theorem 1. *Let A stand for a Haar measurable subset of G with $h(A) > 0$. Assume that a sequence $(f_n)_{n=1}^{\infty}$ of h-almost everywhere continuous real functions defined in A is quasi-uniformly convergent to a function $f : A \longrightarrow \mathbb{R}$. Then f itself is h-almost everywhere continuous.*

Proof. Let E_n denote the set of all continuity points of the function f_n, $n \in \mathbb{N}$. Moreover, let

$$E = \bigcap_{n=1}^{\infty} E_n.$$

Since $h(E_n) = h(A)$ for all $n \in \mathbb{N}$, we also have $h(E) = h(A)$.

Fix arbitrarily x_0 from E. For any positive ε there exists a positive integer n_0 such that

$$|f_n(x_0) - f(x_0)| < \frac{\varepsilon}{3} \quad \text{provided that} \quad n \geq n_0.$$

In view of condition (1), we infer that there exist $n_1, \ldots, n_k$ such that

$$n_1 \geq n_0, \ldots, n_k > n_0 \quad \text{and} \quad |f_{n_1}(t) - f(t)| < \frac{\varepsilon}{3} \vee \ldots \vee |f_{n_k}(t) - f(t)| < \frac{\varepsilon}{3}$$

for all $t \in A$.

Each of the functions f_{n_i} is continuous at x_0; therefore there exists a neighborhood U_0 of x_0 such that

$$|f_{n_i}(t) - f_{n_i}(x_0)| < \frac{\varepsilon}{3}$$

for all $t \in U_0 \cap A$ and $i \in \{1, \ldots, k\}$.

Let $x \in U_0 \cap A$, and let n_{i_o} be such that

$$\left| f_{n_{i_0}}(x) - f(x) \right| < \frac{\varepsilon}{3}.$$

Then

$$|f(x) - f(x_0)| \leq$$

$$\leq \left|f(x) - f_{n_{i_0}}(x)\right| + \left|f_{n_{i_0}}(x) - f_{n_{i_0}}(x_0)\right| + \left|f_{n_{i_0}}(x_0) - f(x_0)\right| < \varepsilon,$$

which proves that f is continuous at every point x_0 from the set E, i.e. h-almost everywhere in A. Thus the proof has been completed.

Recall that a function $f : [a,b] \longrightarrow \mathbb{R}$ is Riemann integrable if and only if it is almost everywhere continuous with respect to the one-dimensional Lebesgue measure. Therefore, applying Theorem 1 for the group $(\mathbb{R}, +)$ and A being a compact interval in $\mathbb{R}$, we obtain immediately the following result.

Theorem 2. *Let $(f_n)_{n=1}^{\infty}$ be a sequence of Riemann integrable functions defined in $[a,b]$. If $f : [a,b] \longrightarrow \mathbb{R}$ is a quasi-uniform limit of the sequence $(f_n)_{n=1}^{\infty}$, then f is Riemann integrable as well.*

Plainly, in general, the Riemann integrability of f does not imply that its Riemann integral is the limit of the sequence of integrals of functions f_n, $n \in \mathbb{N}$. Nevertheless, we have the following *dominated convergence* result.

Theorem 3. *Let $f_n : [a,b] \longrightarrow \mathbb{R}$ be Riemann integrable functions. If $\lim\limits_{n\longrightarrow\infty} f_n = f$ quasi-uniformly and there exists a Riemann integrable function $g : [a,b] \longrightarrow \mathbb{R}$ such that for every positive integer n one has*

$$|f_n| \leq g,$$

then f is Riemann integrable and

$$\lim_{n\longrightarrow\infty} \int_a^b f_n(x)\,dx = \int_a^b \lim_{n\longrightarrow\infty} f_n(x)\,dx.$$

For the proof it suffices to apply Theorem 2 jointly with the classical Lebesgue theorem on majorized convergence.

A careful inspection of the proof of Theorem 1 shows that the group structure as well as the translation invariance of the measure in question are inessential. As a matter of fact, the following abstract setting will allow us to reproduce this proof with no essential changes. Namely, given a topological space X and a proper σ-ideal $\mathscr{I}$ of subsets of X, we say that a function f is $\mathscr{I}$-almost everywhere continuous in X whenever the set of all discontinuity points of the function f yields a member of $\mathscr{I}$. So, we terminate this paper with the following

Theorem 4. *Assume that a sequence $(f_n)_{n=1}^{\infty}$ of $\mathscr{I}$-almost everywhere continuous real functions defined on X is quasi-uniformly convergent to a function $f : X \longrightarrow \mathbb{R}$. Then f itself is $\mathscr{I}$-almost everywhere continuous.*

Now, Theorem 1 becomes a special case of the latter result on setting $X := A$ and $\mathscr{I} := \{F \subset X : h(F) = 0\}$ which, obviously, forms a proper σ-ideal $\mathscr{I}$ of subsets of X.

References

[1] C. Arzelá. Sulle serie di functioni. *Mem. R. Accad. Sci. Inst. Bologna*, ser. 5, (8), 130–186, 701–744, 1899–1900.

[2] B. Szökefalvi-Nagy. *Theory of Real Functions and Orthogonal Expansions*. Akademiai Kiado, Budapest, 1964.

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XV, Częstochowa 2010

THE LOGIC DUAL TO SOBOCIŃSKI'S n–VALUED LOGIC

Anetta Górnicka

Institute of Mathematics and Computer Science
Jan Długosz University of Częstochowa
al. Armii Krajowej 13/15, 42-200 Częstochowa, Poland
e-mail: a.gornicka@ajd.czest.pl

Abstract. In this paper, we describe the logic dual to n–valued Sobociński logic. According to the idea presented by Malinowski and Spasowski [1], we introduce the consequence dual to the consequence of n–valued Sobociński logic in two ways: by a logical matrix and by a set of rules of inference. Then we prove that both approaches are equivalent and the consequence is dual in Wójcicki sense (see [3]).

1. Introduction

By a *language of a propositional logic* (*propositional calculus*) we mean an absolutely free algebra $J = (S, \mathbb{F})$, where S is the set of all formulas built in the standard way on a countable set of propositional variables $p_1, p_2, \ldots$ using functors from the set $\mathbb{F}$.

Let $\mathbf{C}$ denote the family of all consequences in S and let $Cn \in \mathbf{C}$. The consequence dCn dual to the consequence Cn is defined as follows:

Definition 1.

$$\alpha \in dCn(X) \Leftrightarrow \exists_Y \Big(Y \subseteq X \wedge card(Y) < \aleph_0 \wedge \bigcap_{\beta \in Y} Cn(\{\beta\}) \subseteq Cn(\{\alpha\}) \Big)$$

for all formulas $\alpha, \beta \in S$ *and every* $X \subseteq S$.

The definition of a dual consequence applied here was given by Wójcicki [3].

Let $J = (S, \{\Rightarrow, \neg\})$ be the language of Sobociński's n–valued logic described in [2].

Definition 2. *n–valued implicational-negational Sobociński propositional calculus is determined by the following matrix:*

$$\mathfrak{M}_{Sob} = (\{0, 1, 2, \dots, n-1\}, \{1, 2, \dots, n-1\}, \{\Rightarrow, \neg\}), \quad n \geq 3.$$

Here the only nondesignated value is 0.
Functions $\Rightarrow, \neg$ are defined as follows:

$$x \Rightarrow y = \begin{cases} y & \text{if} \quad x \neq y, \\ n-1 & \text{if} \quad x = y, \end{cases}$$

$$\neg x = \begin{cases} x+1 & \text{if} \quad x < n-1, \\ 0 & \text{if} \quad x = n-1, \end{cases}$$

for any $x, y \in \{0, 1, \dots, n-1\}$.

Let us consider the following matrix, which will be called dual to the matrix $\mathfrak{M}_{Sob}$:

$$\mathfrak{M}^d_{Sob} = (\{0, 1, 2, \dots, n-1\}, \{0\}, \{\Rightarrow, \neg\}), \quad n \geq 3,$$

where functions $\Rightarrow$ and $\neg$ are defined in the same way as in the matrix $\mathfrak{M}_{Sob}$.

Definition 3.

1. $\neg^* \alpha \overset{df}{=} (\alpha \Rightarrow \neg(\alpha \Rightarrow \alpha))$.

2. $\alpha \vee^* \beta \overset{df}{=} (\neg^* \alpha \Rightarrow \beta)$.

We call the functors $\neg^*$ and $\vee^*$ the strong negation and the strong disjunction, respectively.

It is easy to observe that a function $\neg^*$ defined by

$$\neg^*(x) = \begin{cases} n-1 & \text{if} \quad x = 0, \\ 0, & \text{otherwise}, \end{cases}$$

corresponds in the matrix $\mathfrak{M}_{Sob}$ to the functor $\neg^*$.
Similarly, a function $\vee^*$ defined by

$$x \vee^* y = \begin{cases} y & \text{if} \quad y \geq 1, \\ 0 & \text{if} \quad x = 0 \text{ and } y = 0, \\ n-1 & \text{if} \quad x \geq 1 \text{ and } y = 0, \end{cases}$$

corresponds in the matrix $\mathfrak{M}_{Sob}$ to the functor $\vee^*$.

Lemma 1. *For arbitrary formulas $\alpha, \beta \in S$ and for any homomorphism $h : J \to (\{0,1,2,\dots,n-1\}, \{\Rightarrow, \neg^*, \vee^*\})$ the following statements are true:*

1. *if $h(\alpha \Rightarrow \beta), h(\alpha) \in \{1,2,\dots,n-1\}$, then $h(\beta) \in \{1,2,\dots,n-1\}$,*
2. *$h(\alpha \Rightarrow \beta) = 0$ iff $h(\alpha) \in \{1,2,\dots,n-1\}$ and $h(\beta) = 0$,*
3. *$h(\alpha) \in \{1,2,\dots,n-1\}$ iff $h(\neg^*\alpha) = 0$,*
4. *$h(\alpha \vee^* \beta) \in \{1,2,\dots,n-1\}$*
   iff $h(\alpha) \in \{1,2\dots,n-1\}$ or $h(\beta) \in \{1,2,\dots,n-1\}$.

Let us consider two inference rules:

$$r_{mp} : \frac{\alpha \Rightarrow \beta, \alpha}{\beta}, \qquad r^d_{mp} : \frac{\neg^*(\alpha \Rightarrow \beta), \beta}{\alpha}.$$

Let $R = \{r_{mp}\}, R^d = \{r^d_{mp}\}$.

Denote by Hom the set of all homomorphisms from $(S, \{\Rightarrow, \neg\})$ into $(\{0,1,\dots,n-1\}, \{\Rightarrow, \neg\})$ and let $X \subseteq S$. We define the matrix consequence $C_{\mathfrak{M}}(X)$, the content $E(\mathfrak{M})$ of the matrix $\mathfrak{M}$ and the consequence $C_R(X)$ based on inference rules from the set X in the standard way:

Definition 4.

1. $C_{\mathfrak{M}_{Sob}}(X) =$
   $= \{\alpha \in S : \forall_{h \in Hom}(h(X) \subseteq \{1,\dots,n-1\} \Rightarrow h(\alpha) \in \{1,\dots,n-1\})\}$.
2. $C_{\mathfrak{M}^d_{Sob}}(X) = \{\alpha \in S : \forall_{h \in Hom}(h(X) \subseteq \{0\} \Rightarrow h(\alpha) = 0)\}$.
3. $E(\mathfrak{M}_{Sob}) = \{\alpha \in S : \forall_{h \in Hom} h(\alpha) \in \{1,2,\dots,n-1\}\}$.
4. $E(\mathfrak{M}^d_{Sob}) = \{\alpha \in S : \forall_{h \in Hom} h(\alpha) = 0\}$.
5. *$C_R(X)$ is the least set Y, which is closed under the rule r_{mp} and which satisfies $E(\mathfrak{M}_{Sob}) \cup X \subseteq Y$.*
6. *$C_{R^d}(X)$ is the least set Y, which is closed under the rule r^d_{mp} and which satisfies $E(\mathfrak{M}^d_{Sob}) \cup X \subseteq Y$.*

2. Some properties of $C_{\mathfrak{M}_{Sob}}, C_{\mathfrak{M}^d_{Sob}}, C_R$ and C_{R^d}

Since modus ponens is the primitive rule of $C_R(X)$ and, as can be easily seen, $\alpha \Rightarrow \alpha, \alpha \Rightarrow (\beta \Rightarrow \alpha), (\alpha \Rightarrow (\beta \Rightarrow \gamma)) \Rightarrow ((\alpha \Rightarrow \beta) \Rightarrow (\alpha \Rightarrow \gamma)) \in E(\mathfrak{M}_{Sob})$, then the classical deduction theorem holds:

Lemma 2. *For arbitrary $\alpha, \beta \in S$ and $X \subseteq S$*

$$\beta \in C_R(X \cup \{\alpha\}) \text{ iff } \alpha \Rightarrow \beta \in C_R(X).$$

Proof. Let us assume that the sequence $\alpha_1, \ldots, \alpha_n$ is the proof based on the set $X \cup \{\alpha\}$ of a formula β. We prove, by induction, that for any $1 \leqslant k \leqslant n$ it holds

$$\alpha \Rightarrow \alpha_k \in C_R(X).$$

Let $k = 1$. Then $\alpha_1 = \alpha$ or $\alpha_1 \in X$.

If $\alpha_1 = \alpha$, then since $\alpha \Rightarrow \alpha \in E(\mathfrak{M}_{Sob})$, we get $\alpha \Rightarrow \alpha_1 \in C_R(X)$.

If $\alpha_1 \in X$, then noticing that $\alpha_1 \Rightarrow (\alpha \Rightarrow \alpha_1) \in E(\mathfrak{M}_{Sob})$, we can see that the sequence $\alpha_1 \Rightarrow (\alpha \Rightarrow \alpha_1), \alpha_1, \alpha \Rightarrow \alpha_1$ is the proof based on X of the formula $\alpha \Rightarrow \alpha_1$.

Assume now that $k > 1$ and for any $i < k, \alpha \Rightarrow \alpha_i \in C_R(X)$.

If $\alpha_k \in X \cup \{\alpha\}$, then the proof is analogous as in the case $k = 1$.

Thus, let α_k results by r_{mp} from α_i, α_j for some $i, j < k$.
Therefore $\alpha_j = \alpha_i \Rightarrow \alpha_k$ and $\alpha \Rightarrow \alpha_i$, $\alpha \Rightarrow (\alpha_i \Rightarrow \alpha_k) \in C_R(X)$. Suppose $\beta_0, \ldots, \beta_{n-1}, \alpha \Rightarrow \alpha_i$ and $\gamma_0, \ldots, \gamma_{m-1}, \alpha \Rightarrow (\alpha_i \Rightarrow \alpha_k)$ are proofs of $\alpha \Rightarrow \alpha_i$ and $\alpha \Rightarrow \alpha_j$, respectively. Then the sequence
$\beta_0, \ldots, \beta_{n-1}, \gamma_0, \ldots, \gamma_{m-1}, (\alpha \Rightarrow (\alpha_i \Rightarrow \alpha_k)) \Rightarrow ((\alpha \Rightarrow \alpha_i) \Rightarrow (\alpha \Rightarrow \alpha_k))$,
$\alpha \Rightarrow (\alpha_i \Rightarrow \alpha_k), (\alpha \Rightarrow \alpha_i) \Rightarrow (\alpha \Rightarrow \alpha_k), \alpha \Rightarrow \alpha_i, \alpha \Rightarrow \alpha_k$ is a proof of $\alpha \Rightarrow \alpha_k$, because $(\alpha \Rightarrow (\beta \Rightarrow \gamma)) \Rightarrow ((\alpha \Rightarrow \beta) \Rightarrow (\alpha \Rightarrow \gamma)) \in E(\mathfrak{M}_{Sob})$.

In the end, let us assume that the sequence $\alpha_1, \ldots, \alpha_n$ is the proof based on X of the formula $\alpha \Rightarrow \beta$. Then $\alpha_n = \alpha \Rightarrow \beta$. It is easy to observe that the sequence $\alpha_1, \ldots, \alpha_n, \alpha, \beta$ is the proof based on $X \cup \{\alpha\}$ of the formula β. □

The next Lemma follows directly from definitions and Lemma 1.

Lemma 3. *For arbitrary $\alpha, \beta \in S$ and $X \subseteq S$*

1. $\beta \in C_{\mathfrak{M}_{Sob}}(X \cup \{\alpha\})$ *iff* $\alpha \Rightarrow \beta \in C_{\mathfrak{M}_{Sob}}(X)$.
2. $\alpha \in C_{\mathfrak{M}_{Sob}}(\{\beta\})$ *iff* $\beta \in C_{\mathfrak{M}^d_{Sob}}(\{\alpha\})$.
3. $\alpha \in C_{\mathfrak{M}^d_{Sob}}(\{\beta\})$ *iff* $\neg^*(\alpha \Rightarrow \beta) \in C_{\mathfrak{M}^d_{Sob}}(\emptyset)$.
4. *The consequences $C_{\mathfrak{M}_{Sob}}, C_{\mathfrak{M}^d_{Sob}}, C_R$ and C_{R^d} are finitary.*

Lemma 4.

1. *The rule r_{mp} is an admissible rule of the consequence $C_{\mathfrak{M}_{Sob}}$.*
2. *The rule r^d_{mp} is an admissible rule of the consequence $C_{\mathfrak{M}^d_{Sob}}$.*

Proof.

1. By Lemma 1, for any homomorphism $h \in Hom$ such that $h(\alpha \Rightarrow \beta), h(\alpha) \in \{1, \dots, n-1\}$ we have $h(\beta) \in \{1, \dots, n-1\}$. This means that $\beta \in C_{\mathfrak{M}_{Sob}}(\{\alpha \Rightarrow \beta, \alpha\})$ and then modus ponens is an admissible rule in $C_{\mathfrak{M}_{Sob}}$.

2. The proof can be carried out on the basis of Definition 4 and Lemma 1.

□

Lemma 5.

1. $C_{\mathfrak{M}^d_{Sob}}(\emptyset) = C_{R^d}(\emptyset) = E(\mathfrak{M}^d_{Sob})$.

2. $C_{\mathfrak{M}_{Sob}}(\emptyset) = C_R(\emptyset) = E(\mathfrak{M}_{Sob})$.

3. $C_{\mathfrak{M}_{Sob}} = C_R$.

Proof. Equalities *1.* and *2.* follow directly from definitions. The proof of the equality *3.* runs as follows:

Let $X \subseteq S$. To prove the inclusion $C_{\mathfrak{M}_{Sob}}(X) \subseteq C_R(X)$ assume that $\alpha \in C_{\mathfrak{M}_{Sob}}(X)$. Due to the finitariness of the matrix consequence $C_{\mathfrak{M}_{Sob}}$ there exists a finite set $X_0 \subseteq X$ such that $\alpha \in C_{\mathfrak{M}_{Sob}}(X_0)$.

If $X_0 = \emptyset$, then using equality *2.*, we infer that $\alpha \in C_R(X_0)$ and therefore $\alpha \in C_R(X)$.

Let $X_0 = \{\alpha_1, \dots, \alpha_m\}$.
By Lemma 3, we get $\alpha_1 \Rightarrow (\dots \Rightarrow (\alpha_m \Rightarrow \alpha)\dots) \in C_{\mathfrak{M}_{Sob}}(\emptyset)$. Then, by equality *2.* and Lemma 2, we have that $\alpha \in C_R(\{\alpha_1, \dots, \alpha_m\})$. As $X_0 \subseteq X$, we see that $\alpha \in C_R(X)$.

To prove the inclusion $C_R(X) \subseteq C_{\mathfrak{M}_{Sob}}(X)$, we apply Lemma 2, Lemma 3 and the fact that C_R is finitary. □

Let us define recursively a generalized strong disjunction by

Definition 5.

1. $\vee^*(\alpha) = \alpha$,

2. $\vee^*(\alpha, \beta) = \alpha \vee^* \beta$,

3. $\vee^*(\alpha_1, \dots, \alpha_{n+1}) = \vee^*(\vee^*(\alpha_1, \dots, \alpha_n), \alpha_{n+1}), \quad n \geqslant 2$.

Lemma 6. *For any natural number $m \geqslant 1$:*

$$C_{\mathfrak{M}^d_{Sob}}(\{\vee^*(\alpha_1, \dots, \alpha_m)\}) = C_{\mathfrak{M}^d_{Sob}}(\{\alpha_1, \dots, \alpha_m\}).$$

Proof. We are going to show that for any formula $\alpha \in S$,

$$\alpha \in C_{\mathfrak{M}^d_{Sob}}(\{\vee^*(\alpha_1, \ldots, \alpha_m)\}) \text{ iff } \alpha \in C_{\mathfrak{M}^d_{Sob}}(\{\alpha_1, \ldots, \alpha_m\}).$$

By Lemma 3, we have the following chain of equivalent statements:

$$\begin{aligned}\alpha \in C_{\mathfrak{M}^d_{Sob}}(\{\vee^*(\alpha_1, \ldots, \alpha_m)\}) \text{ iff } & \vee^*(\alpha_1, \ldots, \alpha_m) \in C_{\mathfrak{M}_{Sob}}(\{\alpha\}) \\ & \text{iff } \alpha \Rightarrow \vee^*(\alpha_1, \ldots, \alpha_m) \in C_{\mathfrak{M}_{Sob}}(\emptyset).\end{aligned}$$

The equivalence $\alpha \Rightarrow \vee^*(\alpha_1, \ldots, \alpha_m) \in C_{\mathfrak{M}_{Sob}}(\emptyset)$ iff $\alpha \in C_{\mathfrak{M}^d_{Sob}}(\{\alpha_1, \ldots, \alpha_m\})$ can be justified in the following way:

„$\Rightarrow$". Suppose that there exists a homomorphism $h_0 \in Hom$ such that $h_0(\{\alpha_1, \ldots, \alpha_m\}) \subseteq \{0\}$ and $h_0(\alpha) \in \{1, \ldots, n-1\}$. Then, by Lemma 1, we get $h_0(\alpha \Rightarrow \vee^*(\alpha_1, \ldots, \alpha_m)) = 0$.

„$\Leftarrow$". Let $\alpha \in C_{\mathfrak{M}^d_{Sob}}(\{\alpha_1, \ldots, \alpha_m\})$ and let us suppose that there exists a homomorphism h_1 such that $h_1(\alpha \Rightarrow \vee^*(\alpha_1, \ldots, \alpha_m)) = 0$. By Lemma 1, we have $h_1(\alpha) \in \{1, \ldots, n-1\}$ and $h_1(\vee^*(\alpha_1, \ldots, \alpha_m)) = 0$. According to Lemma 1, we obtain $h_1(\{\alpha_1, \ldots, \alpha_m\}) \subseteq \{0\}$, so $h_1(\alpha) = 0$. This contradicts our assumption. □

Lemma 7. *For any natural number $m \geqslant 1$:*

$$C_{R^d}(\{\vee^*(\alpha_1, \ldots, \alpha_m)\}) \subseteq C_{R^d}(\{\alpha_1, \ldots, \alpha_m\}).$$

Proof. The proof is inductive on m.

Let us observe that $\neg^*(\neg^*(\alpha_1 \vee^* \alpha_2 \Rightarrow \alpha_1) \Rightarrow \alpha_2) \in E(\mathfrak{M}^d_{Sob})$. By Lemma 5 and Definition 4, we have $\alpha_1 \vee^* \alpha_2 \in C_{R^d}(\{\alpha_1, \alpha_2\})$.
Thus $C_{R^d}(\{\alpha_1 \vee^* \alpha_2\}) \subseteq C_{R^d}(\{\alpha_1, \alpha_2\})$.

Assume that $C_{R^d}(\{\vee^*(\alpha_1, \ldots, \alpha_k)\}) \subseteq C_{R^d}(\{\alpha_1, \ldots, \alpha_k\})$ for some $k \geq 2$. We show that

$$C_{R^d}(\{\vee^*(\alpha_1, \ldots, \alpha_{k+1})\}) \subseteq C_{R^d}(\{\alpha_1, \ldots, \alpha_{k+1}\}).$$

Indeed, $C_{R^d}(\{\vee^*(\alpha_1, \ldots, \alpha_{k+1})\}) = C_{R^d}(\{\vee^*(\vee^*(\alpha_1, \ldots, \alpha_k), \alpha_{k+1})\}) \subseteq$

$\subseteq C_{R^d}(\{\vee^*(\alpha_1, \ldots, \alpha_k), \alpha_{k+1}\}) = C_{R^d}(\{\vee^*(\alpha_1, \ldots, \alpha_k)\} \cup \{\alpha_{k+1}\}) =$

$= C_{R^d}(C_{R^d}(\{\vee^*(\alpha_1, \ldots, \alpha_k)\}) \cup \{\alpha_{k+1}\}) \subseteq C_{R^d}(C_{R^d}(\{\alpha_1, \ldots, \alpha_k\}) \cup \{\alpha_{k+1}\}) =$

$= C_{R^d}(\{\alpha_1, \ldots, \alpha_k\} \cup \{\alpha_{k+1}\}) = C_{R^d}(\{\alpha_1, \ldots, \alpha_{k+1}\})$. □

Lemma 8. *For arbitrary formulas $\alpha, \alpha_1, \ldots, \alpha_m \in S$*
$\alpha \in C_{\mathfrak{M}_{Sob}}(\{\vee^*(\alpha_1, \ldots, \alpha_m)\})$ *iff* $\alpha \in C_{\mathfrak{M}_{Sob}}(\{\alpha_1\}) \cap \ldots \cap C_{\mathfrak{M}_{Sob}}(\{\alpha_m\})$.

Proof. It is a direct consequence of Lemma 1 and Definition 4. □

Lemma 9.

1. $C_{\mathfrak{M}_{Sob}}(\{\alpha\}) = S \Leftrightarrow \alpha \in C_{\mathfrak{M}^d_{Sob}}(\emptyset)$.
2. $C_{\mathfrak{M}^d_{Sob}}(\{\alpha\}) = S \Leftrightarrow \alpha \in C_{\mathfrak{M}_{Sob}}(\emptyset)$.

Proof.

1. „⇒". Let us assume that $C_{\mathfrak{M}_{Sob}}(\{\alpha\}) = S$. Since $\neg^*(p \Rightarrow p) \in C_{\mathfrak{M}_{Sob}}(\{\alpha\})$, then applying Lemma 3, we get $\alpha \in C_{\mathfrak{M}^d_{Sob}}(\{\neg^*(p \Rightarrow p)\})$.
   But $\neg^*(p \Rightarrow p) \in C_{\mathfrak{M}^d_{Sob}}(\emptyset)$, so $\alpha \in C_{\mathfrak{M}^d_{Sob}}(\emptyset)$.

   „⇐". Let us assume that $\alpha \in C_{\mathfrak{M}^d_{Sob}}(\emptyset)$. By Lemma 1 and Definition 4, we get $h(\alpha \Rightarrow \gamma) \in \{1, \ldots, n-1\}$ for every homomorphism h and any formula $\gamma \in S$. By Definition 4 and Lemma 3, we obtain that $\gamma \in C_{\mathfrak{M}_{Sob}}(\{\alpha\})$ for any formula $\gamma \in S$, so $S \subseteq C_{\mathfrak{M}_{Sob}}(\{\alpha\})$. As the opposite inclusion trivially holds, we obtain $C_{\mathfrak{M}_{Sob}}(\{\alpha\}) = S$.
2. The proof is analogous as above. □

3. Main result

Now, we consider the consequences dual in the sense of Definition 1 to the consequences C_R and $C_{\mathfrak{M}_{Sob}}$ and their relation to $C_{\mathfrak{M}^d_{Sob}}$ and C_{R^d}.

Theorem 1.

$$C_{R^d} = C_{\mathfrak{M}^d_{Sob}} = dC_{\mathfrak{M}_{Sob}} = dC_R.$$

Proof. 1° $C_{R^d} = C_{\mathfrak{M}^d_{Sob}}$.

By Lemma 5, we know that $C_{R^d}(\emptyset) = C_{\mathfrak{M}^d_{Sob}}(\emptyset)$ and since, by Lemma 4, the rule r^d_{mp} is an admissible rule of the consequence $C_{\mathfrak{M}^d_{Sob}}$, we get $C_{R^d}(X) \subseteq C_{\mathfrak{M}^d_{Sob}}(X)$ for every $X \subseteq S$, which means that $C_{R^d} \leq C_{\mathfrak{M}^d_{Sob}}$.

Now, let $\alpha \in C_{\mathfrak{M}^d_{Sob}}(X)$. Since the consequence $C_{\mathfrak{M}^d_{Sob}}$ is finitary, there exists a finite set X_0 such that $X_0 \subseteq X$ and $\alpha \in C_{\mathfrak{M}^d_{Sob}}(X_0)$.

If $X_0 = \emptyset$, then by Lemma 5 we get $\alpha \in C_{R^d}(X)$.

Assume then that $X_0 = \{\alpha_1, \ldots, \alpha_m\}$.
Applying Lemma 6, we have $\alpha \in C_{\mathfrak{M}^d_{Sob}}(\{\vee^*(\alpha_1, \ldots, \alpha_m)\})$. In turn, Lemma 3 yields that $\neg^*(\alpha \Rightarrow \vee^*(\alpha_1, \ldots, \alpha_m)) \in C_{\mathfrak{M}^d_{Sob}}(\emptyset)$. Therefore, by Lemma 5, we obtain that $\neg^*(\alpha \Rightarrow \vee^*(\alpha_1, \ldots, \alpha_m)) \in C_{R^d}(\emptyset)$.

Hence, $\alpha \in C_{R^d}(\{\vee^*(\alpha_1, \ldots, \alpha_m)\}) \subseteq C_{R^d}(\{\alpha_1, \ldots, \alpha_m\})$ and then $\alpha \in C_{R^d}(X)$.

Thus we have shown that $C_{\mathfrak{M}^d_{Sob}} \leq C_{R^d}$.

2° $C_{\mathfrak{M}^d_{Sob}} = dC_{\mathfrak{M}_{Sob}}$.

Let $\alpha \in C_{\mathfrak{M}^d_{Sob}}(X)$. Then, by finitariness of $C_{\mathfrak{M}^d_{Sob}}$, we deduce that $\alpha \in C_{\mathfrak{M}^d_{Sob}}(X_1)$ for a finite set $X_1 \subseteq X$.

If $X_1 = \emptyset$, then by Lemma 9
$C_{\mathfrak{M}_{Sob}}(\{\alpha\}) = S$. Hence $\bigcap\limits_{\beta\in\emptyset} C_{\mathfrak{M}_{Sob}}(\{\beta\}) \subseteq C_{\mathfrak{M}_{Sob}}(\{\alpha\})$, i.e. $\alpha \in dC_{\mathfrak{M}_{Sob}}(X)$.

If $X_1 = \{\alpha_1, \ldots, \alpha_m\}$, then $\alpha \in C_{\mathfrak{M}^d_{Sob}}(\{\alpha_1, \ldots, \alpha_m\})$. Applying Lemmas 6 and 3, we obtain that $\vee^*(\alpha_1, \ldots, \alpha_m) \in C_{\mathfrak{M}_{Sob}}(\{\alpha\})$. From this and Lemma 8, we have $C_{\mathfrak{M}_{Sob}}(\{\alpha_1\}) \cap \ldots \cap C_{\mathfrak{M}_{Sob}}(\{\alpha_m\}) \subseteq C_{\mathfrak{M}_{Sob}}(\{\alpha\})$. Thus $\alpha \in dC_{\mathfrak{M}_{Sob}}(X)$ by Definition 1. We have just shown that $C_{\mathfrak{M}^d_{Sob}} \leq dC_{\mathfrak{M}_{Sob}}$.

Suppose now that $\alpha \in dC_{\mathfrak{M}_{Sob}}(X)$. By Definition 1, there exists a finite set $Y \subseteq X$ such that $\bigcap\limits_{\beta\in Y} C_{\mathfrak{M}_{Sob}}(\{\beta\}) \subseteq C_{\mathfrak{M}_{Sob}}(\{\alpha\})$.

If $Y = \emptyset$, then from the fact that $\bigcap\limits_{\beta\in\emptyset} C_{\mathfrak{M}_{Sob}}(\{\beta\}) = S$ and Lemma 9, we obtain that $\alpha \in C_{\mathfrak{M}^d_{Sob}}(X)$.

Therefore, let us assume that $Y = \{\beta_1, \ldots, \beta_m\}$.
Thus $C_{\mathfrak{M}_{Sob}}(\{\beta_1\}) \cap \ldots \cap C_{\mathfrak{M}_{Sob}}(\{\beta_m\}) \subseteq C_{\mathfrak{M}_{Sob}}(\{\alpha\})$. By Lemma 8, we have that $C_{\mathfrak{M}_{Sob}}(\{\vee^*(\beta_1, \ldots, \beta_m)\}) \subseteq C_{\mathfrak{M}_{Sob}}(\{\alpha\})$, i.e.,
$\vee^*(\beta_1, \ldots, \beta_m) \in C_{\mathfrak{M}_{Sob}}(\{\alpha\})$.

Applying Lemma 3, we conclude that $\alpha \in C_{\mathfrak{M}_{Sob}{}^d}(\{\vee^*(\beta_1, \ldots, \beta_m)\})$. Then, according to Lemma 6, we obtain that $\alpha \in C_{\mathfrak{M}^d_{Sob}}(\{\beta_1, \ldots, \beta_m\})$. Hence $\alpha \in C_{\mathfrak{M}^d_{Sob}}(X)$ because $Y = \{\beta_1, \ldots, \beta_m\} \subseteq X$. This proves that $dC_{\mathfrak{M}_{Sob}}(X) \subseteq C_{\mathfrak{M}^d_{Sob}}(X)$, so $dC_{\mathfrak{M}_{Sob}} \leq C_{\mathfrak{M}^d_{Sob}}$.

3° The equality $dC_{\mathfrak{M}_{Sob}} = dC_R$ follows directly from Lemma 5. □

Therefore, the sentential logic (S, C_{R^d}) can be regarded as a logic dual to the Sobociński's n–valued logic (S, C_R). Moreover, it is characterized by the matrix $\mathfrak{M}_{Sob}{}^d$.

References

[1] G. Malinowski, M. Spasowski. Dual counterparts of Łukasiewicz's sentential calculi. *Studia Logica*, **33** (2), 153–162, 1974.

[2] B. Sobociński. Axiomatization of certain many-valued systems of the theory of deduction. *Roczniki prac naukowych zrzeszenia asystentów Uniwersystetu Józefa Piłsudskiego w Warszawie*, No. 1, 399–419, 1936.

[3] R. Wójcicki. Dual counterparts of consequence operations. *Bull. Sect. Logic*, **2** (1), 201–214, 1973.

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XV, Częstochowa 2010

SOME COUNTING FORMULAS FOR FINITE DISTRIBUTIVE LATTICES

Joanna Grygiel

Institute of Mathematics and Computer Science
Jan Długosz University of Częstochowa
al. Armii Krajowej 13/15, 42-200 Częstochowa, Poland
e-mail: j.grygiel@ajd.czest.pl

Abstract. In the paper we show that the weighted double skeleton of a finite distributive lattice is a sufficient structure to characterize the lattice numerically. We prove some combinatorial formulas for the number of all elements of a finite distributive lattice with the given weighted double skeleton, all its elements with exactly k lower covers and all its covering pairs. Introducing some simple examples, we show how the formulas work.

1. Introduction

In the case of big finite lattices it is often impossible to represent them by diagrams. To simplify their description it is useful to introduce the method given by Herrmann in [6], called gluing of lattices, which in fact is a way of building a lattice by means of smaller structures. It is particularly useful in the case of a finite distributive lattice, which turns out to be glued from its maximal Boolean intervals according to some factor structure (being also a lattice) called its skeleton.

However, knowing only the skeleton and Boolean lattices – bricks from which an original distributive lattice $\mathcal{D}$ is built – does not mean that we know how the lattice $\mathcal{D}$ looks like. To make the description complete we introduced in [5] the notion of weighted double skeleton.

Here we are going to show how to compute some combinatorial values of a finite distributive lattice, whose weighted double skeleton is known.

Let us start with introducing some basic notions. It was proved in [2] that maximal Boolean intervals which constitute a finite distributive lattice are in fact blocks of the smallest glued tolerance relation of the lattice.

A tolerance relation on a lattice $\mathcal{L}$ is a reflexive and symmetric binary relation on $\mathcal{L}$ compatible with lattice operations. A block of a tolerance relation Θ is a maximal subset B of L such that every pair of elements of B belongs to Θ. In the case of finite lattices the blocks of any tolerance relation Θ on a lattice $\mathcal{L}$ are intervals and by introducing an order of blocks compatible with the order of their largest elements we get a lattice called the factor lattice $\mathcal{L}/\Theta$.

It is clear that a congruence relation is a special case of a tolerance relation. However, while dealing with congruences we get a partition of the underlying set, here we are rather concerned with overlapping subsets determined by so called glued tolerances. A tolerance relation on $\mathcal{L}$ is called glued if its transitive closure is the total relation on $\mathcal{L}$. It can be proved (see [3]) that blocks of the smallest glued tolerance relation $\Sigma(L)$ are generated by the covering relation on $\mathcal{L}$. The factor lattice $\mathcal{L}/\ \Sigma(L)$ is called the skeleton of $\mathcal{L}$, and it will be denoted by $S(L)$.

Let $\mathcal{L}$ be a finite lattice and denote by $J_k(L)$ (resp. $M_k(L)$) the set of elements of $\mathcal{L}$ with exactly k lower (resp. upper) covers, i.e.

$$J_k(L) = \{a \in L;\ \ |\{b \in L;\ \ b \prec a\}| = k\},$$
$$M_k(L) = \{a \in L;\ \ |\{b \in L;\ \ a \prec b\}| = k\}.$$

It is clear that the zero of $\mathcal{L}$ is the only element of $J_0(L)$ and $J_1(L)$ is the set of all join-irreducible elements of $\mathcal{L}$ (except the zero).

Let $Cov(L)$ denote the set of all covering pairs in $\mathcal{L}$, i.e.

$$Cov(L) = \{(x, y) :\ \ x \prec y,\ x, y \in L\}.$$

In [7], using the Möbius function, Reuter proved a formula counting the numbers of elements in $J_k(L)$ ($M_k(L)$) for any finite lattice with a given glued tolerance relation. Let us recall that the Möbius function μ_P of a poset P can be given by the recursive formula (see e.g. [1]):

$$\begin{cases} \mu_P(x,x) = 1 & \text{for} \quad x \in P, \\ \mu_P(x,y) = -\Sigma_{x \le z < y} \mu_P(x,z) & \text{for} \quad x < z;\ x, z \in P. \end{cases}$$

Theorem 1. *([7]) Let Θ be a glued tolerance relation on a finite lattice $\mathcal{L}$ with the factor lattice P and blocks $\{L_p\}_{p \in P}$. Then for any $k \ge 0$*

$$|J_k(L)| = \sum_{r \le s} \mu_P(r,s) |J_k(L_r \cap L_s)|;$$
$$|M_k(L)| = \sum_{r \le s} \mu_P(r,s) |M_k(L_r \cap L_s)|.$$

Moreover,

$$|Cov(L)| = \sum_{r \leq s} \mu_P(r,s)|Cov(L_r \cap L_s)|.$$

As we see, to count elements of a lattice $\mathcal{L}$ we have to know not only the factor lattice P (the skeleton, for example) and blocks of the glued tolerance relation but intersections of all blocks, as well. All the information in the case of finite distributive lattices can be provided by the weighted double skeleton of the lattice, the notion of which we introduced in [5].

2. The main result

Let $\mathcal{D}$ be a finite distributive lattice with skeleton S. The blocks of the skeleton tolerance Θ are the maximal Boolean intervals of $\mathcal{D}$, we can denote them by $B_x = [0_x, 1_x]$ for $x \in S$. One can show that the subset $\{0_x\}_{x \in S}$ with the order inherited from $\mathcal{D}$ is a lattice isomorphic to the skeleton S (although the meet operations of these lattices may not agree). The same can be said about the subset $\{1_x\}_{x \in S}$ (now, the operations of join in $\mathcal{D}$ and the lattice of units can be different). Thus, these subsets need not form sublattices of $\mathcal{D}$. Let us consider the partially ordered subset $S^d = \{0_x\}_{x \in S} \cup \{1_x\}_{x \in S}$ of D. We shall call it the double skeleton of $\mathcal{D}$.

For simplicity we will write x instead of 0_x and x' instead of 1_x for $x \in S$. Thus, we can regard the double skeleton as a digraph, whose vertices are labeled by elements of some set S and its copy S' and whose arcs are determined just by the covering relation in the poset S^d. Let us observe that S and S' are not necessarily disjoint, hence some vertices can have two labels. It is also clear that if $a \prec b$ in the poset S^d, then $a < b$ in the lattice $\mathcal{D}$, and since all the maximal chains from a to b in a distributive lattice are of the same length, which will be denoted by $l[a,b]$, then in the digraph S^d we can introduce the weight function w assigning to every arc (a,b) the length of the interval $[a,b]$ in $\mathcal{D}$, i.e. $w(a,b) = l[a,b]$. The pair (S^d, w) is called the weighted double skeleton of $\mathcal{D}$.

Let $a \leq b$ in the poset S^d. Then there is a directed path from a to b in the weighted double skeleton and let $\bar{w}(a,b)$ denote the weight of the shortest path from a to b. In fact, in that case all the directed paths are of the same weight and $\bar{w}(a,b) = l[a,b]$.

Theorem 2. *If $\mathcal{D}$ is a finite distributive lattice with the weighted double skeleton (S^d, w), then for any $k \geq 0$*

$$|J_k(D)| = |M_k(D)| = \sum_{\substack{x \leq y \leq x' \\ x,y \in S}} \mu_S(x,y) \binom{\bar{w}(y,x')}{k}.$$

In particular,

$$|D| = \sum_{\substack{x \leq y \leq x' \\ x,y \in S}} \mu_S(x,y) 2^{\bar{w}(y,x')}.$$

Moreover,

$$|Cov(D)| = \sum_{\substack{x \leq y \leq x' \\ x,y \in S}} \mu_S(x,y) \bar{w}(y,x') 2^{\bar{w}(y,x')-1}.$$

Proof. Let $\mathcal{D}$ be a finite distributive lattice with the weighted double skeleton (S^d, w). Then maximal Boolean intervals of $\mathcal{D}$ can be written in the form $\mathcal{B}_x = [x, x']$ for $x \in S$. Let us observe that

$$\dim \mathcal{B}_x = l[x,x'] = w(x,x')$$

for any $x \in S$.

Moreover, if $x < y$ in S, then

$$B_x \cap B_y \neq \emptyset \text{ iff } y \leq x'.$$

In that case $\mathcal{B}_x \cap \mathcal{B}_y$ is also a Boolean interval of the dimension $l[y,x'] = w(y,x')$.

On the other hand, for any Boolean algebra $\mathcal{B}$ and any $0 \leq k \leq \dim B$ we have

$$|J_k(B)| = |M_k(B)| = \binom{\dim B}{k}.$$

Thus, using Theorem 1, we get

$$|J_k(D)| = |M_k(D)| = \sum_{\substack{x \leq y \\ x,y \in S}} \mu_S(x,y) |J_k(B_x \cap B_y)| = \sum_{\substack{x \leq y \leq x' \\ x,y \in S}} \mu_S(x,y) \binom{\bar{w}(y,x')}{k}.$$

In particular,

$$\begin{aligned} |D| &= \sum_{k \geq 0} |J_k(D)| = \sum_{k \geq 0} \sum_{\substack{x \leq y \leq x' \\ x,y \in S}} \mu_S(x,y) \binom{\bar{w}(y,x')}{k} \\ &= \sum_{\substack{x \leq y \leq x' \\ x,y \in S}} \mu_S(x,y) \sum_{k \geq 0} \binom{\bar{w}(y,x')}{k} = \sum_{\substack{x \leq y \leq x' \\ x,y \in S}} \mu_S(x,y) 2^{\bar{w}(y,x')}. \end{aligned}$$

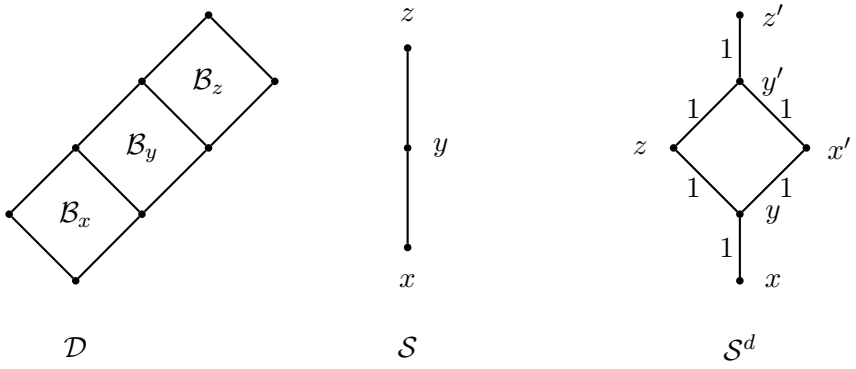


Figure 1:

Now, let us notice that for any m-dimensional Boolean algebra $\mathcal{B}$ we have

$$|Cov(B)| = m2^{m-1}.$$

Therefore, by Theorem 1,

$$|Cov(D)| = \sum_{\substack{x\leq y\\ x,y\in S}} \mu_S(x,y)|Cov(B_x \cap B_y)| = \sum_{\substack{x\leq y\leq x'\\ x,y\in S}} \mu_S(x,y)\bar{w}(y,x')2^{\bar{w}(y,x')-1}.$$

Example 1. Let us consider the distributive lattice $\mathcal{D}$ from Figure 1. Its skeleton S is the three-element chain.

For every poset P being a chain $x_1 \prec x_2 \prec \ldots \prec x_n$ we have

$$\mu_P(x_1,x_i) = \begin{cases} 1 & \text{if} \quad i=1,\\ -1 & \text{if} \quad i=2,\\ 0 & \text{otherwise.}\end{cases}$$

The weighted double skeleton S^d of $\mathcal{D}$ can be found in Figure 1. Thus, the number $|J_1(D)|$ of join-irreducible elements of $\mathcal{D}$ is counted by the formula:

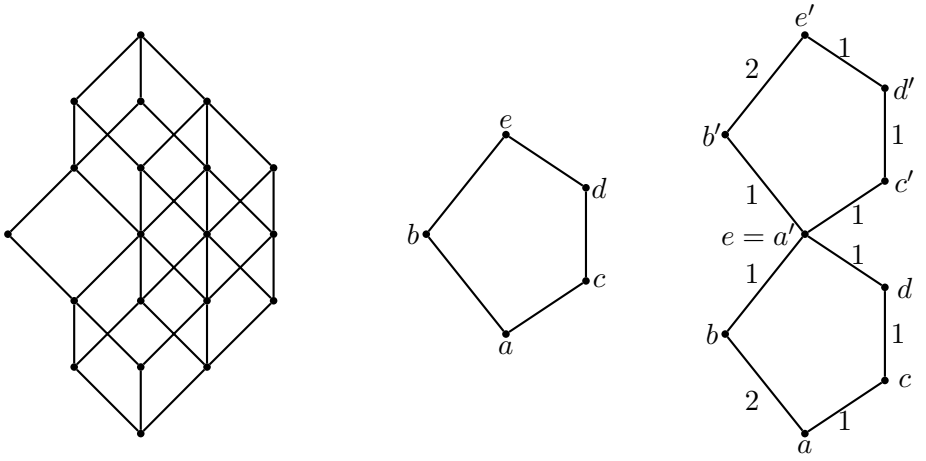
$$\begin{aligned}|J_1(D)| &= \bar{w}(x,x') + \bar{w}(y,y') + \bar{w}(z,z') - \bar{w}(y,x') - \bar{w}(z,y')\\ &= 2+2+2-1-1 = 4,\end{aligned}$$

and the total number of elements of $\mathcal{D}$ is given by

$$|D| = 2^2 + 2^2 + 2^2 - 2^1 - 2^1 = 8.$$

Moreover,

$$|Cov(D)| = 2\cdot 2^1 + 2\cdot 2^1 + 2\cdot 2^1 - 1\cdot 2^0 - 1\cdot 2^0 = 10.$$



$\mathcal{D}$ $\mathcal{S}$ $\mathcal{S}^d$

Figure 2:

Example 2. Let us consider the distributive lattice $\mathcal{D}$ from Figure 2, whose skeleton S is a pentagon. Since the skeleton of the pentagon is the trivial lattice, then $\mathcal{D}$ is an H-irreducible lattice (see [4]) and its double skeleton S^d consists of two copies of the skeleton having one element in common – the top element of the lattice of zeroes is at the same time the bottom element of the lattice of units of the maximal Boolean intervals of $\mathcal{D}$. The weighted double skeleton of $\mathcal{D}$ can be seen in Figure 2.

The Möbius function for the pentagon is given by the table below:

	a	b	c	d	e
a	1	-1	-1	0	1
b	x	1	x	x	-1
c	x	x	1	-1	0
d	x	x	x	1	-1
e	x	x	x	x	1

where x means that the value of μ for the given pair of elements does not exist.

Thus, the number of elements of $\mathcal{D}$ can be counted by the following formula:

$$\begin{aligned}|D| =& 2^{\bar{w}(a,a')} - 2^{\bar{w}(b,a')} - 2^{\bar{w}(c,a')} + 2^{\bar{w}(e,a')} + 2^{\bar{w}(b,b')} \\ & - 2^{\bar{w}(e,b')} + 2^{\bar{w}(c,c')} - 2^{\bar{w}(d,c')} + 2^{\bar{w}(d,d')} - 2^{\bar{w}(e,d')} + 2^{\bar{w}(e,e')} \\ =& 2^3 - 2^1 - 2^2 + 2^0 + 2^2 - 2^1 + 2^3 - 2^2 + 2^3 - 2^2 + 2^3 = 21.\end{aligned}$$

References

[1] M. Aigner. *Combinatorial Theory.* Springer, Berlin, 1979.

[2] H.J. Bandelt. Tolerance relations of lattices. *Bull. Austral. Math. Soc.*, **23**, 367-381, 1981.

[3] B. Ganter, R. Wille. *Formal Concept Analysis. Mathematical Foundations.* Springer, 1999.

[4] J. Grygiel. Some properties of H-irreducible lattices. *Bull. Sect. Logic*, **33** (2), 71–80, 2004.

[5] J. Grygiel. Weighted double skeletons. *Bull. Sect. Logic*, **35** (1), 37–48, 2006.

[6] Ch. Herrmann. S-verklebte Summen von Verbanden. *Math. Z.*, **130**, 255–274, 1973.

[7] K. Reuter. Counting formulas for glued lattices. *Order*, **1**, 265–276, 1985.

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XV, Częstochowa 2010

ON CONNECTED FUNCTIONS IN ORDERED SPACES

Jacek Jędrzejewski

Institute of Mathematics and Computer Science
Jan Długosz University in Częstochowa
al. Armii Krajowej 13/15, 42-200 Częstochowa, Poland
e-mail: jacek.m.jedrzejewski@gmail.com

Abstract. We consider some properties of functions defined in a topological space X with values in a topological space Y. The definitions 1, 2 and 3 define the same class of functions when X and Y are equal to $\mathbb{R}$ with natural topology. In this article we discuss some properties of those classes and give some sufficient conditions for the space X in which real functions defined in X form the same class.

1. Classes of connected functions

We shall consider some properties of functions defined in a topological space X with values in a topological space Y.

Definition 1. *[6] We shall say that a function $f : X \longrightarrow Y$ is connected if its graph is a connected set in $X \times Y$.*

The set of all functions which are connected will be denoted by $\mathscr{C}$.

Definition 2. *[6] We shall say that a function $f : X \longrightarrow Y$ is strongly connected if $f|E$ is a connected set for each connected subset E of X.*

The set of all functions which are strongly connected will be denoted by $\mathscr{C}_s$.

Definition 3. *[4] We shall say that a function $f : X \longrightarrow Y$ is locally strongly connected if for each x in X and its open neighbourhood U there exists open and connected neighbourhood E of x, $E \subset U$, such that $f|E$ is a connected set in the space $X \times Y$.*

The set of all functions which are locally strongly connected will be denoted by $\mathscr{C}_{ls}$.

The definitions 1, 2 and 3 define the same class of functions when X and Y are equal to $\mathbb{R}$ with natural topology.

In the article we shall discuss some properties of those classes and give some sufficient conditions for the space X in which real functions defined in X form the same class.

The terminology and properties concerned with ordered spaces are taken from the articles [1] and [2]. All other topological notions and properties are taken from [3] and [5].

The next properties follow immediately from the above definitions.

Property 1. *Each continuous function is strongly connected.*

Property 2. *Each continuous function defined in a connected space is connected.*

Property 3. *Each continuous function defined in a locally connected space is locally strongly connected.*

Theorem 1. *If a topological spaces X is connected and Y is an arbitrary topological space, then*

$$\mathscr{C}_s \subset \mathscr{C}.$$

This theorem can be completed to get a sufficient condition for a space X to be connected.

Theorem 2. *If a topological space Y has at least two elements and for topological space X*

$$\mathscr{C}_s \subset \mathscr{C},$$

then X is connected.

Theorem 3. *If a topological spaces X is locally connected and Y is an arbitrary topological space, then*

$$\mathscr{C}_s \subset \mathscr{C}_{ls}.$$

Theorem 4. *If a topological spaces X is connected and locally connected and Y is an arbitrary topological space, then*

$$\mathscr{C}_{ls} \subset \mathscr{C}.$$

Proof. For each point x from X there is an open and connected set U_x such that $f|U_x$ is a connected set in $X \times Y$. Of course,

$$\bigcup_{x \in X} U_x = X.$$

Let $\big(x_1, f(x_1)\big)$ and $\big(x_2, f(x_2)\big)$ be arbitrary points of the graph of the function f. The class of sets

$$\{U_x : x \in X\}$$

forms an open cover; then (see [3]) there exists a finite sequence of points $(t_1, \dots, t_n)$ of the set X such that

$$x_1 \in U_{t_1}, \quad x_2 \in U_{t_2} \quad \text{and} \quad U_{t_i} \cap U_{t_j} \neq \emptyset$$

if and only if $|i - j| \leq 1$.

The sets $f|U_{t_i}$ are connected, $f|U_{t_i}$ and $f|U_{t_{i+1}}$ are not disjoint. Hence the set $f|\bigcup_{i=1}^{n} U_{t_i}$ is connected and contains points $\big(x_1, f(x_1)\big)$ and $\big(x_2, f(x_2)\big)$.

We have proved that each two points of the graph of f can be joined by a connected set, therefore the graph of f is connected. □

Example 1. *Let us define a function* $f_1 \colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ *in the following way*

$$f_1(x, y) = \begin{cases} x & \text{if} \quad x > 0,\ y > 0, \\ 0 & \text{otherwise.} \end{cases}$$

The graph of this function is connected but has no other property.

Example 2. *Let* L_1 *be the union of segments* P_n *connecting points* $\left(\frac{1}{2^n}, 0\right)$, $\left(\frac{3}{2^{n+2}}, \frac{8}{10}\right)$ *and points* $\left(\frac{3}{2^{n+2}}, \frac{8}{10}\right)$, $\left(\frac{1}{2^{n+2}}, 0\right)$ *and a half-straight-line from* $(1, 0)$ *towards* $(2, 0)$.

Let L_2 *be the union of segments* Q_n *connecting points* $\left(\frac{1}{2^n}, \frac{1}{10}\right)$, $\left(\frac{3}{2^{n+2}}, \frac{9}{10}\right)$ *and points* $\left(\frac{3}{2^{n+2}}, \frac{9}{10}\right)$, $\left(\frac{1}{2^{n+2}}, \frac{1}{10}\right)$ *and a half-straight-line from* $\left(1, \frac{1}{10}\right)$ *towards* $\left(2, \frac{1}{10}\right)$.

Let L_3 *be the union of segments* S_n *connecting points* $\left(\frac{1}{2^n}, \frac{2}{10}\right)$, $\left(\frac{3}{2^{n+2}}, 1\right)$ *and points* $\left(\frac{3}{2^{n+2}}, 1\right)$, $\left(\frac{1}{2^{n+2}}, \frac{2}{10}\right)$ *and a half-straight-line from* $\left(1, \frac{2}{10}\right)$ *towards* $\left(2, \frac{2}{10}\right)$.

Let us define a function $f_2\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ *in the following way*

$$f_2(x,y) = \begin{cases} 0 & if & (x,y) \in L_1 \cup L_3, \\ 1 & if & (x,y) \in L_2, \\ 0 & if & x > 0,\ y > 1, \\ 0 & if & x > 0,\ y < 0, \\ 0 & if & x \leq 0, \\ \textit{continuous and linear} & \textit{in each vertical segment} & \textit{between lines } L_1,\ L_2, \\ \textit{continuous and linear} & \textit{in each vertical segment} & \textit{between lines } L_2,\ L_3, \\ 0 & \textit{otherwise.} & \end{cases}$$

The above-defined function f_2 is connected and locally strongly connected and, of course, has the local Darboux property, but it has no other, considered in the article, properties.

Example 3. *Let us define a function* $f_3\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ *in the following way*

$$f_3(x,y) = \begin{cases} f_2(x,y) & if \quad x > 0,\ y \in \mathbb{R}, \\ 1 & if \quad x \leq 0,\ y \in \left(\frac{1}{10}, \frac{9}{10}\right), \\ 10y & if \quad x \leq 0,\ y \in \left(0, \frac{1}{10}\right), \\ -10y + 10 & if \quad x \leq 0,\ y \in \left(\frac{9}{10}, 1\right), \\ 0 & if \quad x \leq 0,\ y \in (-\infty, 0) \cup (1, \infty). \end{cases}$$

The function f_3 is connected, strongly connected, locally strongly connected and has the local Darboux property, but it has no other properties.

The above examples complete all the relations among the considered properties.

2. Spaces in which classes of considered connected functions coincide

If we consider real functions defined in an interval of real numbers, i.e. functions $f\colon \mathbb{R} \longrightarrow \mathbb{R}$ or $f\colon (a,b) \longrightarrow \mathbb{R}$, then the four above-defined classes are equal. When we want to compare those classes, it is necessary to consider the space X to be connected (connected functions), locally connected (locally strongly connected functions). Nevertheless, those classes of functions are different if the domain of the functions is $\mathbb{R}^2$ as we have seen in the first part of the article.

If we assume that the space X has dimension 1, the situation is not better: the function f defined in the unit circle of the complex plane by the formula

$$f\left(e^{it}\right) = t \quad \text{if} \quad t \in [0, 2\pi]$$

is connected but it is not strongly connected.

It seems to be very useful the idea of cut points of a connected space, which means that a point x_0 is a cut point of a connected space X if the set $X \setminus \{x_0\}$ is not connected. A point x_0 is a strong cut point of a connected space X if the set $X \setminus \{x_0\}$ has (exactly) 2 connected components.

Similarly, if we assume that every point is a cut point of the space X, the situation is not sufficiently good. The function $f\colon X \longrightarrow Y$ defined by:

$$f(x,y) = \begin{cases} 1 + \sin\frac{1}{x} & \text{if} \quad x \in [-1,0),\ y = 0, \\ 2 + \sin\frac{1}{y} & \text{if} \quad x = 0,\ y \in (0,1], \\ 0 & \text{if} \quad x \in [0,1],\ y = 0, \end{cases}$$

where $X = [-1,1] \times \{0\} \cup \{0\} \times [0,1]$, is connected but it is not strongly connected.

In this way we come to the conclusion that comparing our classes of connected functions we should bound our considerations to functions which have connected or locally connected spaces for which each point is a strong cut point. However, such properties of topological spaces imply that they are linearly ordered spaces. That is the reason for assuming that the spaces X and Y are connected, locally connected and linearly ordered. In such a case it is possible to consider the fourth class of functions, i.e. functions which cut continuum.

If $\prec$ is an order relation in a topological space X, then this space is called ordered if the sets

$$\{x \in X : x \prec a\} \quad \text{and} \quad \{x \in X : a \prec x\}$$

form a subbase of the topology in X.

The sets

$$\{x \in X : a \prec x \wedge x \prec b\} \quad \text{and} \quad \{x \in X : a \prec x \wedge x \prec b\} \cup \{a, b\}$$

are called open and closed intervals in X. These sets are denoted by (a, b) and $[a, b]$, respectively.

The sets

$$\{x \in X : a \prec x\} \quad \text{and} \quad \{x \in X : x \prec b\}$$

are denoted by $(a, \rightarrow)$ and $(\leftarrow, b)$, respectively.

Of course, the class of all open intervals in a linearly ordered space form a base of this topology.

Lemma 1. *In a linearly ordered, connected and locally connected topological space each closed interval is compact.*

Proof. Let X be a linearly ordered and connected topological space, moreover let $[a, b]$ be an arbitrary closed interval in the space X. Suppose that $\{U_s : s \in S\}$ is an arbitrary open cover of the set $[a, b]$. Since X is a linearly ordered space, then each open set can be represented as a union of open intervals:

$$U_s = \bigcup_{t \in T_s} I_{s,t},$$

where $I_{s,t}$ are intervals in X and T_s are some sets of indexes. Then

$$[a, b] \subset \bigcup_{s \in S} \bigcup_{t \in T_s} I_{s,t}.$$

Since $[a, b]$ is a connected subset of the space X, then (see [3]) there exists a finite sequence $(s_1, t_1), \ldots, (s_n, t_n)$ of indexes such that

$$a \in I_{s_1,t_1}, \quad b \in I_{s_n,t_n}$$

and

$$I_{s_i,t_i} \cap I_{s_j,t_j} \neq \emptyset \iff |i - j| \leq 1. \tag{1}$$

Suppose now that some element x_0 from $[a, b]$ does not belong to the set $\bigcup_{i=1}^{n} I_{s_i,t_i}$. Let us assume for shortening of notation that $I_{s_i,t_i} = (a_i, b_i)$. Let

$$i_0 = \max \{i \in \{1, \ldots, n\} : a_i \prec x_0\}.$$

Hence (1) implies that

$$x_0 \in (a_{i_0}, b_{i_0}),$$

what contradicts to the assumption. Thus

$$[a,b] \subset \bigcup_{i=1}^{n} I_{s_i,t_i}$$

and consequently

$$[a,b] \subset \bigcup_{i=1}^{n} U_{s_i}.$$

It proves that the interval $[a,b]$ is compact. □

Theorem 5. *If topological spaces X and Y are linearly ordered, connected and locally connected topological spaces, then each connected function $f\colon X \longrightarrow Y$ is strongly connected.*

Proof. Suppose that there exists a connected function $f\colon X \longrightarrow Y$ which is not strongly connected. Then there exists a connected subset K of X such that $f|K$ is not a connected subset of the space Y. The set K can be one of the following sets:

$$[a,b], \quad (a,b), \quad [a,b), \quad (a,b], \quad (\leftarrow,b), \quad \text{or} \quad (a,\rightarrow).$$

Assume first that $K = [a,b]$. Since the set $f|K$ is not connected, then there are two nonempty separated sets A and B in $X \times Y$ such that

$$f|K = A \cup B.$$

Suppose that $(a, f(a)) \in A$. There are two possibilities:

1. $(b, f(b)) \in A$,
2. $(b, f(b)) \in B$.

In the first case, let

$$A_1 = A \cup f|(\leftarrow,a) \cup f|(b,\rightarrow), \quad B_1 = B.$$

Then

$$f = A_1 \cup B_1, \quad A_1 \neq \emptyset \neq B_1,$$

and the sets A_1 and B_1 are separated, which contradicts to the assumption.

If $(b, f(b)) \in B$, let

$$A_1 = A \cup f|(\leftarrow,a), \quad B_1 = B \cup f|(b,\rightarrow).$$

Then the sets A_1 and B_1 are nonempty and separated, which is impossible in view of connectivity of the function (set) f.

Let now $K = (a, b)$. Then there are nonempty and separated sets A and B such that $f|K = A \cup B$. There exist elements c and d in X such that $a \prec c \prec d \prec b$ and the sets A_1 and B_1 are nonempty and separated, where

$$A \cap ([c, d] \times Y), \quad B \cap ([c, d] \times Y).$$

It is impossible in view of connectivity of the function f.

Similar arguments can be applied in all remained cases for the set K. □

The next theorem is a simple corollary of theorem 4.

Theorem 6. *If topological spaces X and Y are linearly ordered, connected and locally connected, then each locally strongly connected function $f\colon X \longrightarrow Y$ is strongly connected.*

References

[1] R. Duda. On ordered topological spaces. *Fund. Math.*, **63**, 295–309, 1968.

[2] S. Eilenberg. Ordered topological spaces. *Amer. J. Math.*, **63**, 39–45, 1941.

[3] R. Engelking. *General Topology*, PWN, Warszawa, 1977.

[4] J. Jędrzejewski. On some properties of connected functions. *Scientific Issues, Jan Długosz University of Częstochowa, Mathematics*, **XIII**, 15–26, 2008.

[5] J.L. Kelley. *General Topology.* Springer, New York, 1955.

[6] P.E. Long. Connected mappings. *Duke Math. J.*, **35** (4), 677–682, 1968.

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XV, Częstochowa 2010

ENDOMORPHISM MONOID OF DIAMOND PRODUCT OF TWO COMMON COMPLETE BIPARTITE GRAPHS

Thiradet Jiarasuksakun, Tinnaluk Rutjanisarakul, Worapat Thongjua

Department of Mathematics, Faculty of Science
King Mongkut's University of Technology Thonburi (KMUTT)
126 Pracha-uthit Rd. Bangmod, Thungkru, Bangkok 10140 Thailand
e-mail: thiradet.jia@kmutt.ac.th

Abstract. An endomorphism of a graph $G = (V, E)$ is a mapping $f : V \longrightarrow V$ such that for all $x, y \in V$ if $\{x, y\} \in E$, then $\{f(x), f(y)\} \in E$. Let $End(G)$ be the class of all endomorphisms of graph G. The diamond product of graph $G = (V, E)$ (denoted by $G \diamond G$) is a graph defined by the vertex set $V(G \diamond G) = End(G)$ and the edge set $E(G \diamond G) = \{\{f, g\} \subset End(G) | \{f(x), g(x)\} \in E \text{ for all } x \in V\}$. Let $K_{m,n}$ be a complete bipartite graph on $m + n$ vertices. This research aims to study the algebraic property of $V(K_{m,n} \diamond K_{m,n}) = End(K_{m,n})$ after we have found that $K_{m,n} \diamond K_{m,n}$ is also a complete bipartite graph on $m^m n^n + n^m m^n$ vertices. The result shows that all of its vertices (endomorphisms) form a noncommutative monoid.

1. Introduction

In the graph theory [2, 5], a graph $G = (V, E)$ consists of a finite nonempty set V of objects called vertices, and a set E of 2-element subsets of V called edges. In this paper we use the following notation and classification of graphs.

- A path denoted P_n is a sequence of $n + 1$ vertices such that from each of its vertices there is an edge to the next vertex in the sequence.
- A cycle denoted C_n consists of n vertices connected in a closed chain.
- A complete graph denoted K_n is a graph on n vertices such that every two distinct vertices of G are adjacent.

- A graph G is called a bipartite graph if $V(G)$ can be partitioned into two subsets U and W, called partite sets, such that every edge of G joins a vertex of U and a vertex of W.
- A complete bipartite graph denoted $K_{m,n}$ is a graph on $m+n$ vertices such that one can partition V into two subsets U and W, where $|U| = m$ and $|W| = n$. Every edge of G joins a vertex of U and a vertex of W as well as every vertex of U is adjacent to every vertex of W.
- A $u - v$ walk in G is a sequence of vertices in G, beginning at u and ending at v such that consecutive vertices in the sequence are adjacent.
- A $u - v$ path in G is a $u - v$ walk in which no vertices are repeated.
- A graph G is called connected if G contains a $u - v$ path for every pair u, v of distinct vertices in G.
- A regular graph is a graph where each vertex has the same number of neighbors, i.e. every vertex has the same degree or valency. A regular graph with vertices of degree k is called a k-regular graph or regular graph of degree k.
- The distance between two vertices u and v in a graph (denoted by $d(u, v)$) is the number of edges in a shortest path connecting them. This is also known as the geodesic distance because it is the length of the graph geodesic between those two vertices. If there is no path connecting the two vertices, i.e. if they belong to different connected components, then conventionally the distance is defined as infinite.
- The diameter of a graph, denoted $\mathrm{diam}(G)$, is the maximum distance between any two vertices in the graph.

Definition 1. [1] *A homomorphism of a graph $G = (V, E)$ into a graph $H = (V', E')$ is a mapping $f : V \longrightarrow V'$ which preserves edges: for all $x, y \in V$, if $\{x, y\} \in E$, then $\{f(x), f(y)\} \in E'$. Let $Hom(G, H)$ be the class of all homomorphisms from a graph G into a graph H. In particular, an endomorphism of a graph $G = (V, E)$ is a mapping $f : V \longrightarrow V$ such that for all $x, y \in V$, if $\{x, y\} \in E$, then $\{f(x), f(y)\} \in E$. Let $End(G)$ be the class of all endomorphisms of graph G.*

From this definition, one can easily see that $Hom(G, H)$ may or may not exist. For example, $Hom(P_1, C_3)$ consists of 6 homomorphisms, while $Hom(C_3, P_1)$ is an empty set.

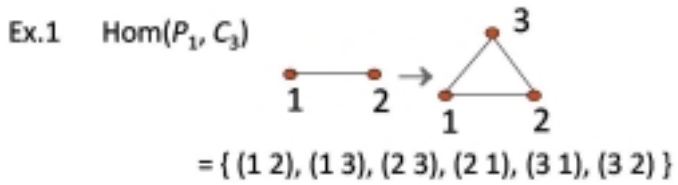


Figure 1: Hom(P_1, C_3).

Definition 2. [1] *The diamond product of a graph $G = (V, E)$ and a graph $H = (V', E')$ (denoted by $G \diamond H$) is a graph defined by the vertex set $V(G \diamond H) = Hom(G, H)$, where $Hom(G, H) \neq \emptyset$, and the edge set $E(G \diamond H) = \{\{f, g\} \subset Hom(G, H) | \{f(x), g(x)\} \in E' \text{ for all } x \in V\}$. In particular, the diamond product of a graph G with itself ($G \diamond G$) is defined by the vertex set $V(G \diamond G) = End(G)$ and the edge set $E(G \diamond G) = \{\{f, g\} \subset End(G) | \{f(x), g(x)\} \in E \text{ for all } x \in V\}$.*

An example of graph $P_1 \diamond C_3$ is shown below.

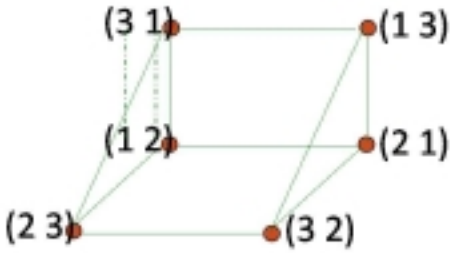


Figure 2: Graph $P_1 \diamond C_3$.

With this definition, there are some interesting results as follows:

Theorem 1. [3] *The graph $P_m \diamond P_n$ is connected for all positive integers m, n and $diam(P_m \diamond P_n) = n$.*

Theorem 2. [3] *Graphs $P_m \diamond C_n$ and $C_n \diamond P_m$ are connected for all positive integers m, n. $diam(P_m \diamond C_n) \leq m + n$ and $diam(C_n \diamond P_m) = n$.*

Theorem 3. [3] *If G is a connected graph, then the graph $P_m \diamond G$ is connected for all positive integers m, and $diam(P_m \diamond G) = diam(G) + 2m$.*

2. Some observations

In this paper, we study the diamond product of two complete bipartite graphs $K_{m,n}$.

- Denote $V(K_{m,n}) = \{1, 2, 3, ..., m, m+1, m+2, ..., m+n\}$, where $V_m = \{x \in V(K_{m,n}) \mid x \leq m\}$, and $V_n = \{x \in V(K_{m,n}) \mid m+1 \leq x \leq m+n\}$.
  Since $K_{m,n}$ is a complete bipartite graph, each vertex of V_m is adjacent to all vertices of V_n. Every edge joins a vertex of V_m and a vertex of V_n. We can define a function $h : V(K_{m,n}) \to \{0, 1\}$ such that

$$h(x) = \begin{cases} 0 & \text{if } x \in V_m, \\ 1 & \text{if } x \in V_n. \end{cases}$$

  By the definition of a complete bipartite graph, we obtain for all $x, y \in V(K_{m,n})$, $\{x, y\} \in E(K_{m,n})$ if and only if $\mid h(x) - h(y) \mid = 1$.

- Let $f : V(K_{m,n}) \to V(K_{m,n})$ be a homomorphism.
  Then $f \in V(K_{m,n} \diamond K_{m,n})$ if and only if

$$h(f(i)) = \begin{cases} 0 & \text{if } i \in V_m, \\ 1 & \text{if } i \in V_n \end{cases}$$

  or

$$h(f(i)) = \begin{cases} 1 & \text{if } i \in V_m, \\ 0 & \text{if } i \in V_n. \end{cases}$$

  For example, let us take a look at $K_{2,2} \diamond K_{2,2}$.

- We can define a norm as follows:

$$\|f - g\| = \max_{i \in V(K_{m,n})} |h(f(i)) - h(g(i))|.$$

3. Main results

Lemma 1. *For $f, g \in V(K_{m,n} \diamond K_{m,n})$, $\{f, g\} \in E(K_{m,n} \diamond K_{m,n})$ if and only if $\|f - g\| = 1$.*

Proof.
($\Rightarrow$) Let $\{f, g\} \in E(K_{m,n} \diamond K_{m,n})$. We have $\{f(i), g(i)\} \in E(K_{m,n})$ for all $i \in V(K_{m,n})$. Thus $|h(f(i)) - h(g(i))| = 1$ for all $i \in V(K_{m,n})$. This means that $\|f - g\| = 1$.

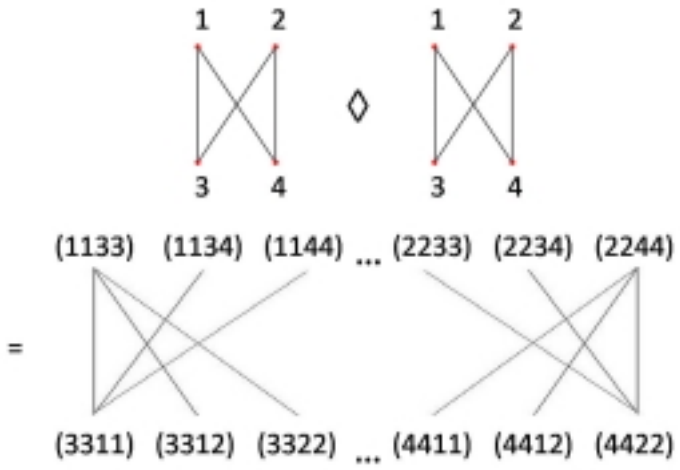


Figure 3: Graph $K_{2,2} \diamond K_{2,2}$

($\Leftarrow$) Let $\|f - g\| = 1$, where $f, g \in V(K_{m,n} \diamond K_{m,n})$. From the definition of norm, $\exists\ i_0 \in V(K_{m,n})$ such that $|h(f(i_0)) - h(g(i_0))| = 1$. Without loss of generality we may assume that $h(f(i_0)) = 0$ and $h(g(i_0)) = 1$.
If $i_0 \in V_m$, then we obtain

$$h(f(i)) = \begin{cases} 0 & \text{if } i \in V_m, \\ 1 & \text{if } i \in V_n \end{cases}$$

and

$$h(g(i)) = \begin{cases} 1 & \text{if } i \in V_m, \\ 0 & \text{if } i \in V_n. \end{cases}$$

So $|h(f(i)) - h(g(i))| = 1$, for all $i \in V(K_{m,n})$.
If $i_0 \in V_n$, then we obtain

$$h(f(i)) = \begin{cases} 1 & \text{if } i \in V_m, \\ 0 & \text{if } i \in V_n \end{cases}$$

and

$$h(g(i)) = \begin{cases} 0 & \text{if } i \in V_m, \\ 1 & \text{if } i \in V_n. \end{cases}$$

So $|h(f(i)) - h(g(i))| = 1$ for all $i \in V(K_{m,n})$. From both cases, we obtain $|h(f(i)) - h(g(i))| = 1$ for all $i \in V(K_{m,n})$. By the definitions of function h and diamond product, $\{f, g\} \in E(K_{m,n} \diamond K_{m,n})$. □

Theorem 4. *$K_{m,n} \diamond K_{m,n}$ is a complete bipartite graph on $m^m n^n + n^m m^n$ vertices.*

Proof.
First, let us define ${V_m}^\diamond = \left\{ f \in V(K_{m,n} \diamond K_{m,n}) \mid h(f(i)) = \left\{ \begin{array}{ll} 0 & \text{if } i \in V_m \\ 1 & \text{if } i \in V_n \end{array} \right\} \right.$
and ${V_n}^\diamond = \left\{ f \in V(K_{m,n} \diamond K_{m,n}) \mid h(f(i)) = \left\{ \begin{array}{ll} 1 & \text{if } i \in V_m \\ 0 & \text{if } i \in V_n \end{array} \right\} \right.$.
Obviously, $V(K_{m,n} \diamond K_{m,n}) = {V_m}^\diamond \cup {V_n}^\diamond$ and ${V_m}^\diamond \cap {V_n}^\diamond = \emptyset$.

To show that the graph of $K_{m,n} \diamond K_{m,n}$ is bipartite, we need to prove that $\{f, g\} \in E(K_{m,n} \diamond K_{m,n})$ if and only if f and g belong to different sets of vertices ${V_m}^\diamond$ and ${V_n}^\diamond$.

($\Rightarrow$) First, let f and g belong to the same set of vertices. Without loss of generality we can assume $f, g \in {V_m}^\diamond$. We have

$$\|f - g\| = \max_{i \in V(K_{m,n})} |h(f(i)) - h(g(i))|.$$

If $i \in V_m$, then $h(f(i)) = 0,\ h(g(i)) = 0$ and

$$\max_{i \in V_m} |h(f(i)) - h(g(i))| = \max |0 - 0| = 0.$$

If $i \in V_n$, then $h(f(i)) = 1,\ h(g(i)) = 1$ and

$$\max_{i \in V_n} |h(f(i)) - h(g(i))| = \max |1 - 1| = 0.$$

Therefore $\|f - g\| = 0$ implies that $\{f, g\} \notin E(K_{m,n} \diamond K_{m,n})$ by Lemma 1. Then we conclude that if f and g belong to the same sets of vertices, there is no edge $\{f, g\}$ in the graph $K_{m,n} \diamond K_{m,n}$.

($\Leftarrow$) Without loss of generality we can take $f \in {V_m}^\diamond$ and $g \in {V_n}^\diamond$. We have

$$\|f - g\| = \max_{i \in V(K_{m,n})} |h(f(i)) - h(g(i))|.$$

If $i \in V_m$, then $h(f(i)) = 0,\ h(g(i)) = 1$ and

$$\max_{i \in V_m} |h(f(i)) - h(g(i))| = \max |0 - 1| = 1.$$

If $i \in V_n$, then $h(f(i)) = 1,\ h(g(i)) = 0$ and

$$\max_{i \in V_n} |h(f(i)) - h(g(i))| = \max |1 - 0| = 1.$$

Then $|h(f(i)) - h(g(i))| = 1$ for all $i \in V(K_{m,n})$, and $\|f - g\| = 1$. Therefore $\{f, g\} \in E(K_{m,n} \diamond K_{m,n})$.

By definition, all the vertices $f \in {V_m}^\diamond$ have the same value of $h(f(i))$ for all $i \in V$, and all the vertices $g \in {V_n}^\diamond$ have the same value of $h(g(i))$ for all $i \in V$ such that $\|f - g\| = 1$. This means that each vertex of ${V_m}^\diamond$ is adjacent to all vertices of ${V_n}^\diamond$, making it a complete bipartite graph.

We know that $K_{m,n} \diamond K_{m,n}$ have two partite sets ${V_m}^\diamond$ and ${V_n}^\diamond$. From the definition of ${V_m}^\diamond$, an endomorphism maps each vertex of V_m into a vertex of V_m giving us m^m choices and maps each vertex of V_n into a vertex of V_n with n^n choices. Thus $|{V_m}^\diamond| = m^m n^n$. On the other hand, an endomorphism in ${V_n}^\diamond$ maps each vertex of V_m into a vertex of V_n giving us n^m choices and maps each vertex of V_n into a vertex of V_m with m^n choices. Thus $|{V_n}^\diamond| = n^m m^n$. Both cases combined, we obtain the number of vertices in the theorem. □

Corollary 1. *$K_{m,n} \diamond K_{m,n}$ is a regular graph if and only if $m = n$.*

Proof.
Since $K_{m,n} \diamond K_{m,n}$ is a complete bipartite graph, we may pick $f \in {V_m}^\diamond$ and $g \in {V_n}^\diamond$. From Theorem 4, we have the following:

- $\{f, k\} \in E(K_{m,n} \diamond K_{m,n})$ for all $k \in {V_n}^\diamond$.
  Thus $\deg(f) = |{V_n}^\diamond| = n^m \cdot m^n$ for all $f \in {V_m}^\diamond$.
- $\{g, h\} \in E(K_{m,n} \diamond K_{m,n})$ for all $h \in {V_m}^\diamond$.
  Thus $\deg(g) = |{V_m}^\diamond| = m^m \cdot n^n$ for all $g \in {V_n}^\diamond$.

Hence, $K_{m,n} \diamond K_{m,n}$ is a regular graph if and only if $\deg(f) = \deg(g)$, which implies $m = n$. □

Now let us consider the vertex set of $K_{m,n} \diamond K_{m,n}$ with operation of function composition.

Theorem 5. *The vertex (endomorphism) set of $K_{m,n} \diamond K_{m,n}$ with composition form a noncommutative monoid for all positive integers $m, n > 1$.*

Proof.
It is clear that $V(K_{m,n} \diamond K_{m,n})$ is a monoid. To show that in the case when $m, n > 1$, it is noncommutative we can take $f, g \in V(K_{m,n} \diamond K_{m,n})$ such that

$$f(i) = \begin{cases} i & \text{if } \ i \in V_m, \\ m+1 & \text{if } \ i \in V_n, \end{cases}$$

$$g(i) = \begin{cases} m+2 & \text{if } \ i \in V_m, \\ i-m & \text{if } \ i \in V_n. \end{cases}$$

Then we have $(f \circ g)(m) = f(g(m)) = f(m+2) = m+1$. But $(g \circ f)(m) = g(f(m)) = g(m) = m+2$. Thus $f \circ g \neq g \circ f$, making it a noncommutative monoid. □

Remark 1. *This noncommutative monoid is not a group since an endomorphism may not have an inverse. There exists a many-to-one endomorphism such as*

$$f(i) = \begin{cases} 1 & \text{if } i \in V_m, \\ m+1 & \text{if } i \in V_n. \end{cases}$$

Therefore, this endomorphism set forms only a noncommutative monoid, not a group.

Acknowledgements

The authors would like to thank Prof. Arworn and Ms. Damnernsawad who introduced us to the diamond product of graphs at the seminar at Chiangmai University in July 2008. Prof. Arworn also gave us some useful comments and suggestions.

References

[1] Sr. Arworn, P. Wojtylak. *Connectedness of Diamond Products*, preprint, 2008.

[2] G. Chartrand, P. Zhang. *Introduction to Graph Theory.* McGraw-Hill, 2005.

[3] J. Damnernsawad. *Diamond Product of Paths*, Master Degree Thesis, ChiangMai University, Thailand, 2007.

[4] T. Jiarasuksakun, T. Rutjanisarakul, W. Thongjua. *Diamond Product of Two Common Complete Bipartite Graphs*, Int. Conf. on Algebra and Geometry 2009 (ICAG 2009), Phuket, Thailand, 2009.

[5] D. West. *Introduction to Graph Theory*, 2nd edition. Prentice Hall, 2001.

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XV, Częstochowa 2010

$\Psi_{\mathcal{I}}$-DENSITY TOPOLOGY

Ewa Łazarow[a], Agnieszka Vizváry[b]

[a] *Institute of Mathematics*
Academia Pomeraniensis in Słupsk
ul. Arciszewskiego 22b, 62-200 Słupsk, Poland
e-mail: elazarow@toya.net.pl

[b] *Faculty of Mathematics and Computer Science*
University of Łódź
ul. Banacha 22, 90-238 Łódź, Poland
e-mail: vizvary@gazeta.pl

Abstract. The purpose of this paper is to study the notion of a $\Psi_{\mathcal{I}}$-density point and $\Psi_{\mathcal{I}}$-density topology, generated by it analogously to the classical $\mathcal{I}$-density topology on the real line. The idea arises from the note by Taylor [3] and Terepeta and Wagner-Bojakowska [2].

We introduce the following notation:

$\mathbb{N}$ the set of positive integers,
$\mathbb{R}$ the set of real numbers,
$\mathbb{R}_+$ the set of positive real numbers,
$\mathcal{S}$ σ-algebra of subsets of $\mathbb{R}$ having the Baire property,
$\mathcal{I}$ σ-ideal of subsets of $\mathbb{R}$ of the first category,
$\mathcal{C}$ the family of all nondecreasing continuous functions $\psi : \mathbb{R}_+ \to (0, 1]$ such that $\lim\limits_{x \to 0^+} \psi(x) = 0$.

We say that two sets A and B are equivalent ($A \sim B$) if $A \triangle B \in \mathcal{I}$, where $A \triangle B$ is the symmetric difference of A and B. Additionally, if $A \subset \mathbb{R}$, $\alpha \in \mathbb{R}$ and $x_0 \in \mathbb{R}$, then $-A = \{x \in \mathbb{R} : \ -x \in A\}$, $\alpha \cdot A = \{\alpha \cdot x \in \mathbb{R} : \ x \in A\}$, $A' = \mathbb{R} \setminus A$ and $A - x_0 = \{x \in \mathbb{R} : \ x + x_0 \in A\}$. For each $x \in \mathbb{R}^+$, let $[x] = \max\{n \in \mathbb{N} \cup \{0\} : \quad n \leq x \ \}$.

Definition 1. [1] *We say that 0 is a point of $\mathcal{I}$-density of a set $A \in \mathcal{S}$ if for each increasing sequence of positive integers $\{n_m\}_{m\in\mathbb{N}}$ there exists a subsequence $\{n_{m_p}\}_{p\in\mathbb{N}}$ such that*

$$\{x:\ \chi_{n_{m_p}\cdot A\cap[-1,1]}(x) \not\longrightarrow 1\} \in \mathcal{I}.$$

A point x_0 is a point of $\mathcal{I}$-density of a set $A \in \mathcal{S}$ if 0 is a point of $\mathcal{I}$-density of the set $A - x_0$. A point x_0 is a point of $\mathcal{I}$-dispersion of a set $A \in \mathcal{I}$ if x_0 is a point of $\mathcal{I}$-density of the set $\mathbb{R} \setminus A$.

Let

$$\Phi(A) = \{x \in \Re:\ \ x \text{ is } \mathcal{I}\text{-density point of } A\}$$

for $A \in \mathcal{S}$, and $\mathcal{T}_\mathcal{I} = \{A \in \mathcal{S}:\ A \subset \Phi(A)\}$. We recall the following theorems.

Theorem 1. [1] *0 is a point of $\mathcal{I}$-density of a set $A \in \mathcal{S}$ if and only if for each sequence $\{t_n\}_{n\in\mathbb{N}} \subset \mathbb{R}_+$ such that $\lim_{n\to\infty} t_n = +\infty$ there exists a subsequence $\{t_{n_k}\}_{k\in\mathbb{N}}$ such that*

$$\{x \in [-1,1]:\ \chi_{t_{n_k}\cdot A\cap[-1,1]}(x) \not\longrightarrow 1\} \in \mathcal{I}.$$

Theorem 2. [1] *For any $A \in \mathcal{S}$ and $B \in \mathcal{S}$,*

i) *if $A \subset B$, then $\Phi(A) \subset \Phi(B)$,*

ii) $\Phi(\emptyset) = \emptyset$, $\Phi(\mathbb{R}) = \mathbb{R}$,

iii) *if $A \sim B$, then $\Phi(A) = \Phi(B)$,*

iv) $\Phi(A \cap B) = \Phi(A) \cap \Phi(B)$,

v) $A \sim \Phi(A)$.

Theorem 3. [1] *$\mathcal{T}_\mathcal{I}$ is a topology on the real line stronger than the Euclidean topology.*

Definition 2. *Let $\psi \in \mathcal{C}$.*

I. *We say that 0 is a point of right-hand $\psi_\mathcal{I}$-dispersion of a set A from $\mathcal{S}$ if for each sequence $\{(h_n, m_n)\}_{n\in\mathbb{N}}$ with the following properties*

- $\{(h_n, m_n)\}_{n\in\mathbb{N}} \subset \mathbb{R}_+ \times (\mathbb{N} \cup \{0\})$,
- *the sequence $\{h_n\}_{n\in\mathbb{N}}$ is decreasing,*
- $\lim\limits_{n\to\infty} h_n = 0$,
- *for each $n \in \mathbb{N}$, $m_n \in \{0, \dots, \left[\frac{1}{\psi(h_n)}\right] - 1\}$*

there exists a subsequence $\{(h_{n_k}, m_{n_k})\}_{k\in\mathbb{N}}$ *such that*

$$\{x \in [0,1];\ \chi_{A_k}(x) \not\longrightarrow 0\} \in \mathcal{I},$$

where

$$A_k = \left(\frac{1}{h_{n_k}\psi(h_{n_k})} \cdot A - m_{n_k}\right) \cap [0,1].$$

II. *We say that 0 is a point of left-hand $\psi_{\mathcal{I}}$-dispersion of a set $A \in \mathcal{S}$ if 0 is a right-hand point of $\psi_{\mathcal{I}}$-dispersion of the set $-A$.*

III. *We say that 0 is a point of $\psi_{\mathcal{I}}$-dispersion of a set $A \in \mathcal{S}$ if 0 is a point of right-hand and left-hand $\psi_{\mathcal{I}}$-dispersion of the set A.*

IV. *We say that $x_0 \in \mathbb{R}$ is a point of $\psi_{\mathcal{I}}$-dispersion of a set $A \in S$ if 0 is a point of $\psi_{\mathcal{I}}$-dispersion of the set $A - x_0$.*

V. *We say that $x_0 \in \mathbb{R}$ is a point of $\psi_{\mathcal{I}}$-density of a set $A \in S$ if x_0 is a point of $\psi_{\mathcal{I}}$-dispersion of the set $\mathbb{R} \setminus A$.*

Lemma 1. *Let $\psi \in \mathcal{C}$ and $\{(a_n, b_n)\}_{n\in\mathcal{N}}$ be a sequence of open intervals such that $\lim\limits_{n\to\infty} b_n = 0$ and, for each $n \in \mathbb{N}$,*

i) $0 < a_{n+1} < b_{n+1} < a_n$,

ii) $b_{n+1} \leq b_n\psi(b_n)$,

iii) $b_n - a_n \leq b_n\psi(b_n)$.

Let $G = \bigcup\limits_{n=1}^{\infty}(a_n, b_n)$. Then, for each sequence of positive real numbers $\{h_n\}_{n\in\mathbb{N}}$ such that $\lim\limits_{n\to\infty} h_n = 0$ there exists a subsequence $\{h_{n_k}\}_{k\in\mathbb{N}}$ satisfying the condition

$$\left\{ x \in [0,1] : \chi_{\frac{1}{h_{n_k}}\cdot G\cap[0,1]}(x) \not\longrightarrow 0 \right\} \in \mathcal{I}.$$

Proof. Let $\{h_n\}_{n\in\mathbb{N}}$ be an arbitrary sequence of positive real numbers such that $\lim\limits_{n\to\infty} h_n = 0$. We can assume that, for each $n \in \mathbb{N}$, there exists $p_n \in \mathbb{N}$ such that

$$b_{p_n+1} < h_n \leq b_{p_n}.$$

We shall consider two cases.

a) There exists positive integer n_0 such that, for each $n \geq n_0$,

$$b_{p_n+1} \leq h_n \leq a_{p_n}.$$

Assume that $n_0 = 1$. We consider a sequence $\left\{\frac{1}{h_n} \cdot b_{p_n+1}\right\}_{n\in\mathbb{N}}$. Then there exist $\alpha \in [0,1]$ and an increasing sequence of positive integers $\{n_k\}_{k\in\mathcal{N}}$ such that

$$\lim_{k\to\infty} \frac{1}{h_{n_k}} \cdot b_{p_{n_k}+1} = \alpha.$$

Hence

$$0 \leq \lim_{k\to\infty} \frac{1}{h_{n_k}} \cdot \left(b_{p_{n_k}+1} - a_{p_{n_k}+1}\right) \leq \lim_{k\to\infty} \frac{1}{h_{n_k}} \cdot b_{p_{n_k}+1} \cdot \psi(b_{p_{n_k}+1}) = 0$$

and

$$\lim_{k\to\infty} \frac{1}{h_{n_k}} \cdot a_{p_{n_k}+1} = \alpha.$$

By the above and

$$0 \leq \lim_{k\to\infty} \frac{1}{h_{n_k}} \cdot b_{p_{n_k}+2} \leq \lim_{k\to\infty} \frac{1}{h_{n_k}} \cdot b_{p_{n_k}+1} \cdot \psi(b_{p_{n_k}+1}) = 0,$$

we infer that

$$\left\{ x \in [0,1] : \chi_{\frac{1}{h_{n_k}} \cdot G\cap[0,1]}(x) \not\longrightarrow 0 \right\} \subset \{0, \alpha, 1\}.$$

b) Now we assume that, for each $n \in \mathbb{N}$, there exists $k_n \in \mathbb{N}, k_n \geq n$ such that

$$a_{p_{n_k}} < h_{k_n} < b_{p_{n_k}}.$$

Then

$$1 \leq \lim_{k\to\infty} \frac{1}{h_{k_n}} \cdot b_{p_{n_k}} \leq \lim_{k\to\infty} \frac{1}{a_{p_{n_k}}} \cdot b_{p_{n_k}} \leq \lim_{k\to\infty} \frac{1}{b_{p_{n_k}}(1 - \psi(b_{p_{n_k}}))} \cdot b_{p_{n_k}} = 1$$

and

$$\lim_{k\to\infty} \frac{1}{h_{k_n}} \cdot \left(b_{p_{n_k}} - a_{p_{n_k}}\right) \leq \lim_{k\to\infty} \frac{1}{h_{k_n}} b_{p_{n_k}} \psi(b_{p_{n_k}}) = 0.$$

Hence

$$\lim_{k\to\infty} \frac{1}{h_{k_n}} \cdot a_{p_{n_k}} = 1.$$

Additionaly

$$\lim_{k\to\infty}\frac{1}{h_{k_n}}\cdot b_{p_{n_k}+1}\leq\lim_{k\to\infty}\frac{1}{h_{k_n}}\cdot b_{p_{n_k}}\psi(b_{p_{n_k}})=0,$$

therefore

$$\left\{x\in[0,1]:\chi_{\frac{1}{h_{n_k}}\cdot G\cap[0,1]}(x)\not\longrightarrow 0\right\}\subset\{0,1\}.\qquad\square$$

Theorem 4. *Let $\psi\in\mathcal{C}$. If 0 is a point of right-hand $\psi_{\mathcal{I}}$–dispersion of a set $A\in S$, then it is a point of a right-hand $\mathcal{I}$-dispersion of the set A.*

Proof. Let $\{t_n\}_{n\in\mathbb{N}}$ be a decreasing sequence of positive real numbers such that $\lim\limits_{n\to\infty}t_n=0$. We may assume that, for each $n\in\mathbb{N}$, there exists a positive h_n such that

$$t_n=h_n\psi(h_n).$$

Then $\lim\limits_{n\to\infty}h_n=0$. Let, for each $n\in\mathbb{N}$, $m_n=0$.

The sequence $\{(h_n,m_n)\}_{n\in\mathbb{N}}$ satisfies the conditions of Definition 2, therefore there exists a sequence $\{(h_{n_k},m_{n_k})\}_{k\in\mathbb{N}}$ such that

$$\limsup_{k\to\infty}\left(\frac{1}{h_{n_k}\psi(h_{n_k})}\cdot A-m_{n_k}\right)\cap[0,1]\in\mathcal{I}.$$

By

$$\limsup_{k\to\infty}\left(\frac{1}{h_{n_k}\psi(h_{n_k})}\cdot A-m_{n_k}\right)\cap[0,1]=\limsup_{k\to\infty}\left(\frac{1}{t_{n_k}}\cdot A\right)\cap[0,1],$$

the proof is complete. $\square$

Theorem 5. *Let $\psi\in\mathcal{C}$. There exists an open set G such that 0 is a point of right-hand $\mathcal{I}$-dispersion of the set G and 0 is not a point of right-hand $\psi_{\mathcal{I}}$-dispersion of the set G.*

Proof. We shall define a sequence of open intervals $\{(a_n,b_n)\}_{n\in\mathbb{N}}$ such that

i) $0<a_{n+1}<b_{n+1}<a_n$,

ii) $b_{n+1}<\min\{\frac{1}{n},b_n\psi(b_n)\}$,

iii) $b_n-a_n=b_n\psi(b_n)$,

iv) $\frac{1}{\psi(b_n)}\in\mathbb{N}$

for each $n\in\mathbb{N}$.

Let b_1 be a positive real number such that $\psi(b_1) \in \left\{\frac{1}{2}, \frac{1}{3}, ...\right\}$. Let $n \in \mathbb{N}$. Assume that we have defined positive real numbers $b_1, ..., b_n$. Now we shall define a positive b_{n+1} fulfilling the following properties:

$$\psi(b_{n+1}) \in \left\{\frac{1}{2}, \frac{1}{3}, ...\right\} \quad \text{and} \quad b_{n+1} < \min\left\{\frac{1}{n}, b_n\psi(b_n)\right\}.$$

For each $n \in \mathbb{N}$, we put $a_n = b_n - b_n\psi(b_n)$. Then, for each $n \in \mathbb{N}$,

$$a_{n-1} = b_{n-1}(1 - \psi(b_{n-1})) \geq b_{n-1} \cdot \frac{1}{2} \geq b_{n-1} \cdot \psi(b_{n-1}) > b_n.$$

Set $G = \bigcup_{n=1}^{\infty}(a_n, b_n)$.

By Lemma 1, 0 is a point of right-hand $\mathcal{I}$-dispersion of the set G. Now we prove that 0 is not a point of right-hand $\psi_{\mathcal{I}}$-dispersion of the set G.

Let $\{(h_n, m_n)\}_{n\in\mathbb{N}}$ be a sequence such that $h_n = b_n$, $m_n = \left[\frac{1}{\psi(h_n)}\right] - 1$ for each $n \in \mathbb{N}$, and let $\{(h_{n_k}, m_{n_k})\}_{k\in\mathbb{N}}$ be an arbitrary subsequence of $\{(h_n, m_n)\}_{n\in\mathbb{N}}$. We shall show that

$$(0,1) \subset \limsup_{k\to\infty}\left(\frac{1}{h_{n_k}\psi(h_{n_k})} \cdot G - m_{n_k}\right).$$

Let $k \in \mathbb{N}$. Then

$$\frac{1}{h_{n_k}\psi(h_{n_k})} \cdot G - m_{n_k} \supset \frac{1}{h_{n_k}\psi(h_{n_k})} \cdot (a_{n_k}, b_{n_k}) - m_{n_k} =$$

$$\left(\frac{1}{b_{n_k}\psi(b_{n_k})}(b_{n_k} - b_{n_k}\psi(b_{n_k})) - \left[\frac{1}{\psi(b_{n_k})}\right] + 1, \frac{1}{b_{n_k}\psi(b_{n_k})} \cdot b_{n_k} - \left[\frac{1}{\psi(b_{n_k})}\right] + 1\right) =$$

$$= \left(\frac{1}{\psi(b_{n_k})}(1 - \psi(b_{n_k})) - \frac{1}{\psi(b_{n_k})} + 1, \frac{1}{\psi(b_{n_k})} - \frac{1}{\psi(b_{n_k})} + 1\right) = (0,1).$$

By the above, 0 is not a point of right-hand $\psi_{\mathcal{I}}$-dispersion of the set G. □

Theorem 6. *Let $\psi \in \mathcal{C}$. There exists an open set G such that 0 is an accumulation point of the set G and 0 is a point of right-hand $\psi_{\mathcal{I}}$-dispersion of the set G.*

Proof. We define sequences of real positive numbers $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ such that

1) $b_{n+1} \leq \frac{1}{n} a_n \psi(a_n)$,

2) $0 < b_n - a_n \leq \frac{1}{n} a_n \psi(a_n)$,

for each $n \in \mathbb{N}$, and

3) $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n = 0$.

Let b_1 be an arbitrary real positive number. Let $n \in \mathbb{N}$. Assume that we have defined numbers $b_1, ..., b_{n-1}$ and $a_1, ..., a_{n-1}$. Let b_n be a real positive number such that $b_n \leq \frac{1}{n-1} a_{n-1}\psi(a_{n-1})$. By the continuity of a function $g(x) = x + \frac{1}{n} x\psi(x)$ and by $b_n < b_n + \frac{1}{n} b_n \psi(b_n)$, there exists a_n such that $a_n < b_n$ and $a_n + \frac{1}{n} a_n \psi(a_n) = b_n$.

Set $G = \bigcup\limits_{n=1}^{\infty} (a_n, b_n)$. We shall show that 0 is a point of right-hand $\psi_{\mathcal{I}}$-dispersion of G. Let $\{(h_n, m_n)\}_{n\in\mathbb{N}}$ be an arbitrary sequence satisfying the conditions of Definition 2. We consider the following possibilities:

a) Assume that there exists a subsequence $\{(h_{n_k}, m_{n_k})\}_{k\in\mathbb{N}}$ such that for each $k \in \mathbb{N}$, $m_{n_k} = 0$. Then, in view of Lemma 1, 0 is a point of a right-hand $\mathcal{I}$–dispersion of G. Since $\lim\limits_{k\to\infty} h_{n_k}\psi(h_{n_k}) = 0$, we may choose a subsequence $\left\{h_{n_{k_p}}\right\}_{p\in\mathbb{N}}$ such that

$$\limsup_{p\to\infty} \left(\frac{1}{h_{n_{k_p}} \psi(h_{n_{k_p}})} G - m_{n_{k_p}} \right) \cap [0,1] = \limsup_{p\to\infty} \frac{1}{h_{n_{k_p}} \psi(h_{n_{k_p}})} \cdot G \cap [0,1] \in \mathcal{I}.$$

b) Assume that there exists a subsequence $\{(h_{n_k}, m_{n_k})\}_{k\in\mathbb{N}}$ such that

$$[m_{n_k} h_{n_k} \psi(h_{n_k}), (m_{n_k} + 1) h_{n_k} \psi(h_{n_k})] \cap G = \emptyset$$

for each $k \in \mathbb{N}$.

Then, for each $k \in \mathbb{N}$, $\left(\frac{1}{h_{n_k}\psi(h_{n_k})} \cdot G - m_{n_k} \right) \cap [0,1] = \emptyset$. Hence

$$\limsup_{k\to\infty} \left(\frac{1}{h_{n_k} \psi(h_{n_k})} \cdot G - m_{n_k} \right) \cap [0,1] = \emptyset.$$

c) If none of the cases a) and b) is true, then there exists $n_0 \in \mathbb{N}$ such that for each $n \geq n_0$, $m_n \geq 1$ and

$$[m_n h_n \psi(h_n), (m_n + 1) h_n \psi(h_n)] \cap G \neq \emptyset.$$

We can assume that for each $n \in \mathbb{N}$ there exists $r_n \in \mathbb{N}, r_n > 1$ such that

$$[m_n h_n \psi(h_n), (m_n + 1) h_n \psi(h_n)] \cap (a_{r_n}, b_{r_n}) \neq \emptyset.$$

Therefore

$$a_{r_n} \leq (m_n+1)h_n\psi(h_n) \leq \left[\frac{1}{\psi(h_n)}\right] h_n\psi(h_n) \leq h_n$$

and, by

$$b_{r_n+1} \leq \frac{1}{r_n} a_{r_n}\psi(a_{r_n}) \leq 1 \cdot h_n\psi(h_n) \leq m_n h_n\psi(h_n),$$

we have

$$[m_n h_n\psi(h_n), (m_n+1)h_n\psi(h_n)] \cap \bigcup_{j=r_n+1}^{\infty} (a_j, b_j) = \emptyset.$$

Additionaly, by $a_{r_n-1} > h_n$,

$$[m_n h_n\psi(h_n), (m_n+1)h_n\psi(h_n)] \cap \bigcup_{j=1}^{r_n-1} (a_j, b_j) = \emptyset.$$

Let $n \in \mathbb{N}$ and

$$x_n \in [m_n h_n\psi(h_n), (m_n+1)h_n\psi(h_n)] \cap (a_{r_n}, b_{r_n}).$$

Then $\frac{1}{h_n\psi(h_n)} \cdot x_n - m_n \in [0,1]$, for all $n \in \mathbb{N}$. Thus, there exists $\alpha \in [0,1]$ and a subsequence $\left\{\frac{1}{h_{n_k}\psi(h_{n_k})} x_{n_k} - m_{n_k}\right\}_{k\in\mathbb{N}}$ such that

$$\lim_{k\to\infty} \left(\frac{1}{h_{n_k}\psi\left(h_{n_k}\right)} x_{n_k} - m_{n_k}\right) = \alpha.$$

By

$$\begin{aligned} 0 &\leq \lim_{k\to\infty} \frac{1}{h_{n_k}\psi(h_{n_k})} \cdot (b_{r_{n_k}} - a_{r_{n_k}}) \\ &\leq \lim_{k\to\infty} \frac{1}{h_{n_k}\psi(h_{n_k})} \cdot \frac{1}{r_{n_k}} \cdot a_{r_{n_k}}\psi(a_{r_{n_k}}) \\ &\leq \lim_{k\to\infty} \frac{1}{h_{n_k}\psi(h_{n_k})} \cdot \frac{1}{r_{n_k}} \cdot h_{n_k}\psi(h_{n_k}) \\ &= \lim_{k\to\infty} \frac{1}{r_{n_k}} = 0, \end{aligned}$$

we infer that

$$\lim_{k\to\infty} \left(\frac{1}{h_{n_k}\psi(h_{n_k})} b_{r_{n_k}} - m_{n_k}\right) = \alpha$$

and

$$\lim_{k\to\infty}\left(\frac{1}{h_{n_k}\psi(h_{n_k})}a_{r_{n_k}}-m_{n_k}\right)=\alpha.$$

Thus

$$\limsup_{k\to\infty}\left(\frac{1}{h_{n_k}\psi(h_{n_k})}\cdot G-m_{n_k}\right)\cap[0,1]\subset\{\alpha\}. \qquad \square$$

Theorem 7. *Let $\psi_1\in\mathcal{C}$. There exist a function $\psi_2\in\mathcal{C}$ and an open set G such that 0 is a point of right-hand $\psi_{1,\mathcal{I}}$-dispersion of the set G, but 0 is not a point of right-hand $\psi_{2,\mathcal{I}}$-dispersion of the set G.*

Proof. We define a sequence of open intevals $\{(a_n,b_n)\}_{n\in\mathbb{N}}$ such that

1) $0<a_{n+1}<b_{n+1}<a_n$,

2) $b_{n+1}\leq\frac{1}{n}a_n\psi_1(a_n)$,

3) $b_n-a_n\leq\frac{1}{n}a_n\psi_1(a_n)$,

4) $\frac{b_n-a_n}{b_n}<\frac{b_{n-1}-a_{n-1}}{b_{n-1}}$,

5) $\frac{b_n}{b_n-a_n}\in\mathbb{N}$,

for each $n\in\mathbb{N}$, and

6) $\lim\limits_{n\to\infty}b_n=0$.

Let $b_1\in(0,1)$ and $k\in\mathbb{N}\setminus\{1\}$. Assume that we have defined numbers $a_1,...,a_{k-1}$ and $b_1,...,b_{k-1}$. Let b_k be an arbitrary positive number such that $b_k\leq\frac{1}{k-1}a_{k-1}\psi_1(a_{k-1})$.

We consider two functions: $g(x)=x+\frac{1}{k}x\psi_1(x)$ and $h(x)=1-\frac{x}{b_k}$. Since $g(b_k)=b_k+\frac{1}{k}b_k\psi_1(b_k)>b_k$, therefore, by continuity of a function g, we have $\alpha\in(0,b_k)$ such that $g(\alpha)=b_k$ and, for each $x\in(\alpha,b_k)$, $g(x)>b_k$. Let p be a positive integer such that

$$\frac{1}{p}<\min\left\{\frac{b_{k-1}-a_{k-1}}{b_{k-1}},h(\alpha)\right\}.$$

Then

$$0=h(b_k)<\frac{1}{p}<h(\alpha)$$

and, by continuity of h, we can choose $a_k\in(\alpha,b_k)$ such that $h(a_k)=\frac{1}{p}$.

Set $G = \bigcup_{n=1}^{\infty}(a_n, b_n)$. Let $\psi_2 \in \mathcal{C}$ be a function such that, for each $n \in \mathbb{N}$, $\psi_2(b_n) = \frac{b_n - a_n}{b_n}$. In a similar way as in Theorem 6, one can prove that 0 is a point of right-hand $\psi_{1,\mathcal{I}}$-dispersion of the set G.

We shall show that 0 is not a point of right-hand $\psi_{2,\mathcal{I}}$-dispersion of the set G. Let $h_n = b_n$ for each $n \in \mathbb{N}$ and $m_n = [\frac{1}{\psi_2(b_n)}] - 1$. The sequence $\{(h_n, m_n)\}_{n\in\mathbb{N}}$ satisfies the conditions of Definition 2. Let $\{(h_{n_k}, m_{n_k})\}_{n\in\mathbb{N}}$ be an arbitrary subsequence of $\{(h_n, m_n)\}_{n\in\mathbb{N}}$. Then, for each $k \in \mathbb{N}$,

$$\frac{1}{h_{n_k}\psi_2(h_{n_k})} \cdot G - m_{n_k} \supset \frac{1}{h_{n_k}\psi_2(h_{n_k})} \cdot (a_{n_k}, b_{n_k}) - m_{n_k} = (0, 1).$$

Thus

$$(0,1) \subset \limsup_{k\to\infty} \left(\frac{1}{h_{n_k}\psi_2(h_{n_k})} \cdot G - m_{n_k} \right) \cap [0,1].$$

Definition 3. *Let $\psi \in \mathcal{C}$. For a set $A \in \mathcal{S}$, we define $\Phi_\psi(A)$ to be the set of all points of $\psi_\mathcal{I}$-density of the set A.*

Theorem 8. *Let $\psi \in \mathcal{C}$. Then, for any $A, B \in \mathcal{S}$,*

1) $\Phi_\psi(\emptyset) = \emptyset, \Phi_\psi(\mathbb{R}) = \mathbb{R}$,

2) *If $A \subset B$, then $\Phi_\psi(A) \subset \Phi_\psi(B)$,*

3) *If $A \sim B$, then $\Phi_\psi(A) = \Phi_\psi(B)$,*

4) $\Phi_\psi(A \cap B) = \Phi_\psi(A) \cap \Phi_\psi(B)$,

5) $A \sim \Phi_\psi(A)$.

Proof. The conditions 1) and 2) are obvious. Assume that $A \sim B$ and $x \in \Phi_\psi(A)$. Without loss of generality, one can assume that $x = 0$. We only show that if 0 is a point of right-hand $\psi_\mathcal{I}$-dispersion of the set A', then 0 is a point of right-hand $\psi_\mathcal{I}$-dispersion of the set B'.

Let $\{(h_n, m_n)\}_{n\in\mathbb{N}}$ be an arbitrary sequence which satisfies conditions of Definition 2. We observe that

$$B' = (B' \cap A') \cup (B' \setminus A'),$$

where $B' \setminus A' = A \setminus B \in \mathcal{I}$, and $B' \cap A' \subset A'$. 0 is a point of right-hand $\psi_\mathcal{I}$-dispersion of the set A', thus it is a point of right-hand $\psi_\mathcal{I}$-dispersion of

the set $A' \cap B'$. Therefore, there exists a subsequence $\{(h_{n_k}, m_{n_k})\}_{k\in\mathbb{N}}$ such that

$$\limsup_{k\to\infty}\left(\frac{1}{h_{n_k}\psi(h_{n_k})}\cdot(A'\cap B') - m_{n_k}\right)\cap[0,1]\in\mathcal{I}.$$

We define the sets P, P_1, P_2 in the following way:

$$\begin{aligned}
P = &\quad \limsup_{k\to\infty}\left(\frac{1}{h_{n_k}\psi(h_{n_k})}\cdot B' - m_{n_k}\right)\cap[0,1],\\
P_1 = &\quad \limsup_{k\to\infty}\left(\frac{1}{h_{n_k}\psi(h_{n_k})}\cdot(A'\cap B') - m_{n_k}\right)\cap[0,1],\\
P_2 = &\quad \limsup_{k\to\infty}\left(\frac{1}{h_{n_k}\psi(h_{n_k})}\cdot(B'\setminus A') - m_{n_k}\right)\cap[0,1].
\end{aligned}$$

Then $P\subset P_1\cup P_2$. The set P_1 is of the first category, and

$$P_2 = \bigcap_{r=1}^{\infty}\bigcup_{k=r}^{\infty}\left(\frac{1}{h_{n_k}\psi(h_{n_k})}\cdot(B'\setminus A') - m_{n_k}\right)\cap[0,1]\in\mathcal{I}.$$

Thus $P\in\mathcal{I}$.

We have proved that $\Phi_\psi(A)\subset\Phi_\psi(B)$. In a similar way, we can prove that $\Phi_\psi(B)\subset\Phi_\psi(A)$.

Now we shall show condition 4). Since $A\cap B\subset A$ and $A\cap B\subset B$, therefore, by condition 2), we have $\Phi_\psi(A\cap B)\subset\Phi_\psi(A)\cap\Phi_\psi(B)$.

Let $x\in\Phi_\psi(A)\cap\Phi_\psi(B)$. We can assume that $x=0$. Let $\{(h_n,m_n)\}_{n\in\mathbb{N}}$ be an arbitrary sequence which satisfies the conditions of Definition 2. Since 0 is a point of right-hand $\psi_{\mathcal{I}}$-dispersion of the set A', therefore there exists a subsequence $\{(h_{n_k},m_{n_k})\}_{k\in\mathbb{N}}$ such that

$$\limsup_{k\to\infty}\left(\frac{1}{h_{n_k}\psi(h_{n_k})}\cdot A' - m_{n_k}\right)\cap[0,1]\in\mathcal{I}.$$

Additionaly, 0 is a point of right-hand $\psi_{\mathcal{I}}$-dispersion of the set B', thus there exists a subsequence $\{(h_{n_{k_p}},m_{n_{k_p}})\}_{k\in\mathbb{N}}$, such that

$$\limsup_{k\to\infty}\left(\frac{1}{h_{n_{k_p}}\psi(h_{n_{k_p}})}\cdot B' - m_{n_{k_p}}\right)\cap[0,1]\in\mathcal{I}.$$

Then

$$\limsup_{k\to\infty}\left(\frac{1}{h_{n_{k_p}}\psi(h_{n_{k_p}})}\cdot(A\cap B)' - m_{n_{k_p}}\right)\cap[0,1]\subset H,$$

where

$$\begin{aligned} H &= \limsup_{k\to\infty}\left(\frac{1}{h_{n_{k_p}}\psi(h_{n_{k_p}})}\cdot A' - m_{n_{k_p}}\right)\cap[0,1]\cup \\ &\cup \limsup_{k\to\infty}\left(\frac{1}{h_{n_{k_p}}\psi(h_{n_{k_p}})}\cdot B' - m_{n_{k_p}}\right)\cap[0,1]\in\mathcal{I}. \end{aligned}$$

Hence, 0 is a point of right-hand $\psi_{\mathcal{I}}$-dispersion of the set $(A\cap B)'$. In a similar way, we can show that 0 is a point of left-hand $\psi_{\mathcal{I}}$-dispersion of the set $(A\cap B)'$.

Now we shall show condition 5). Let $A\in\mathcal{S}$. Then $A=(G\setminus P_1)\cup P_2$, where G is an open set, P_1 i P_2 are sets of the first category and $P_1\subset G$, $P_2\cap G=\emptyset$. By 3), we have $\Phi_\psi(A)=\Phi_\psi(G)$ and $G\subset\Phi_\psi(G)$. Thus

$$A\setminus\Phi_\psi(A)=A\setminus\Phi_\psi(G)\subset A\setminus G\in\mathcal{I}.$$

By Theorem 4, $\Phi_\psi(A)\subset\Phi(A)$ and by Theorem 2, $A\sim\Phi(A)$, therefore $\Phi_\psi(A)\setminus A\subset\Phi(A)\setminus A\in\mathcal{I}$. □

Definition 4. *Let, for $\psi\in\mathcal{C}$,*

$$\mathcal{T}_\psi=\{A\in\mathcal{S}:A\subset\Phi_\psi(A)\}.$$

By theorems 3, 4, 5 and 8 we have the following

Theorem 9. *Let $\psi\in\mathcal{C}$. $\mathcal{T}_\psi$ is a topology on the real line, stronger than the Euclidean topology and weaker than the $\mathcal{I}$–topology.*

Lemma 2. *Assume that we have a sequences of real numbers $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty}a_n=\lim_{n\to\infty}b_n=0$ and, for each $n\in\mathbb{N}$, $0<b_{n+1}<a_n<b_n$. Then there exists a function $\psi\in\mathcal{C}$ such that 0 is not a point of $\psi_{\mathcal{I}}$-dispersion of the set $G=\bigcup_{n=1}^{\infty}(a_n,b_n)$.*

Proof. First we define values of the finction ψ at points of the sequence $\{b_n\}_{n\in\mathbb{N}}$. Set $\psi(b_1)=\frac{1}{\left[\frac{b_1}{b_1-a_1}\right]+1}$, $a_2'=\max\{a_2,b_2(1-\psi(b_1))\}$ and $\psi(b_2)=\frac{1}{\left[\frac{b_2}{b_2-a_2'}\right]+1}$. Assume that for $n\in\mathbb{N}$ we have defined the points $a_1',\ldots,a_n'$ and the real numbers $\psi(b_1),\ldots,\psi(b_n)$ in the following way:

- $a_{i+1}'=\max\left\{a_{i+1},b_{i+1}(1-\frac{1}{i}\psi(b_i))\right\}$ if $i\in\{1,\ldots,n-1\}$,

- $\psi(b_{i+1}) = \frac{1}{\left[\frac{b_{i+1}}{b_{i+1}-a'_{i+1}}\right]+1}$ if $i \in \{1, \dots, n-1\}$.

Put $a'_{n+1} = \max\left\{a_{n+1}, b_{n+1}(1-\frac{1}{n}\psi(b_n))\right\}$ and $\psi(b_{n+1}) = \frac{1}{\left[\frac{b_{n+1}}{b_{n+1}-a'_{n+1}}\right]+1}$.

We observe that $\psi(b_{n+1}) < \frac{1}{n}\psi(b_n)$. Indeed

$$\frac{1}{n}\psi(b_n) \geq 1 - \frac{a'_{n+1}}{b_{n+1}} = \frac{1}{\frac{b_{n+1}}{b_{n+1}-a'_{n+1}}} > \frac{1}{\left[\frac{b_{n+1}}{b_{n+1}-a'_{n+1}}\right]+1} = \psi(b_{n+1}).$$

Let $\psi \in \mathcal{C}$ be a function such that, for any $n \in \mathbb{N}$ and $x \in [a_n, b_n]$, $\psi(x) = \psi(b_n)$. In a similar way as in Theorem 4, we can show that 0 is not a point of right-hand $\psi_{\mathcal{I}}$-dispersion of the set G. □

Definition 5. *We denote by $\mathcal{H}$ the Hashimoto topology, where*

$$\mathcal{H} = \{U \setminus P : U - \text{ an open set }, P \in \mathcal{I}\}.$$

Theorem 10. $\bigcap_{\psi\in\mathcal{C}} \mathcal{T}_\psi = \mathcal{H}$.

Proof. It is obvious that $\mathcal{H} \subset \bigcap_{\psi\in\mathcal{C}} \mathcal{T}_\psi$. Let $A \in \mathcal{S}$ and $A \notin \mathcal{H}$. Then $A = (G \setminus P_1) \cup P_2$, where G is an open set, $P_1, P_2 \in \mathcal{I}$, $P_1 \subset G$ and $P_2 \cap G = \emptyset$.

Set $H = \mathrm{Int}(\mathrm{Cl}(G))$ and $R = H \setminus (G \cup P_2)$. By $A \notin \mathcal{H}$, we know that P_2 is not a subset of H. It is easy to see that $\mathrm{Int}\,(\mathbb{R} \setminus H) \neq \emptyset$ and the set $\mathbb{R} \setminus H$ has no isolated points.

Let $x_0 \in P_2 \cap (\mathbb{R} \setminus H)$ and $\{(c_n, d_n)\}_{n\in\mathbb{N}}$ be a sequence of all components of the set $\mathrm{Int}\,(\mathbb{R} \setminus H)$. We consider the following cases:

a) $x_0 \in \mathrm{Int}\,(\mathbb{R} \setminus H)$. Then, for an arbitrary function $\psi \in \mathcal{C}$, x_0 is a point of right-hand $\psi_{\mathcal{I}}$-dispersion of the set H. Thus, x_0 is a point of right-hand $\psi_{\mathcal{I}}$-dispersion of the set $G \subset H$. Since $\Phi_\psi(A) = \Phi_\psi(G)$, we have $x_0 \notin \Phi_\psi(A)$. Therefore, $A \not\subset \Phi_\psi(A)$, and $A \notin \mathcal{T}_\psi$.

b) There exists $n_0 \in \mathbb{N}$ such that $x_0 = c_{n_0}$ or $x_0 = d_{n_0}$. Then x_0 is a point of right-hand or left-hand $\psi_{\mathcal{I}}$–density of the set $\mathbb{R} \setminus H$ for arbitrary function $\psi \in \mathcal{C}$, respectively, and, as above, $x_0 \in A \setminus \Phi_\psi(A)$.

c) There exists a sequence $\{c_{n_k}\}_{k\in\mathbb{N}}$ which converges to x_0 from the right or there exists a sequence $\{d_{n_k}\}_{k\in\mathbb{N}}$ which converges to x_0 from the left. Then, by Lemma 2, there exists a function $\psi \in \mathcal{C}$ such that x_0 is not a point of $\psi_{\mathcal{I}}$-dispersion of the set $\bigcup_{k=1}^{\infty}(c_{n_k}, d_{n_k})$. Thus, $x_0 \notin \Phi_\psi(A)$ and $A \notin \mathcal{T}_\psi$. Therefore, $A \notin \bigcap_{\psi\in\mathcal{C}} \mathcal{T}_\psi$. □

References

[1] W. Poreda, E. Wagner-Bojakowska, W. Wilczyński. A category analogue of the density topology. *Fund. Math.*, CXXV, 167–173, 1985.

[2] M. Terepeta, E. Wagner-Bojakowska. ψ-density topologies. *Rend. Circ. Mat. Palermo*, Ser. II, XLVIII, 451–476, 1999.

[3] S.J. Taylor. On strenghtening the Lebesgue density theorem. *Fund. Math.*, XLVI, 305–315, 1959.

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XV, Częstochowa 2010

ON SOME GENERALIZATIONS OF GOŁĄB–SCHINZEL FUNCTIONAL EQUATION

Janusz Matkowski[a,b], **Jolanta Okrzesik**[c]

[a] *Faculty of Mathematics, Computer Science and Econometrics*
University of Zielona Góra, Podgórna 50
65246 Zielona Góra, Poland
e-mail: J.Matkowski@wmie.uz.zgora.pl

[b] *Institute of Mathematics, Silesian University*
Bankowa 14, 0007 Katowice, Poland

[c] *Department of Mathematics, ATH Bielsko-Biała, Poland*
e-mail: jokrzesik@ath.bielsko.pl

Abstract. Composite functional equations in several variables generalizing the Gołąb–Schinzel equation are considerd and some simple methods allowing us to determine their one-to-one solutions, bijective solutions or the solutions having exactly one zero are presented. For an arbitrarily fixed real p, the functional equation

$$\phi\left([p\,\phi(y)+(1-p)]x+[(1-p)\phi(x)+p\,]y\right)=\phi(x)\phi(y),\quad x,y\in\mathbb{R},$$

being a special generalization of the Gołąb–Schinzel equation, is considered.

1. Introduction

Composite functional equations in several variables, i.e. equations involving the superpositions of unknown functions, represent an important class of equations. The translation equation (cf. Aczél [1], p. 245),

$$\phi(\phi(x,s),t)=\phi(x,s+t),$$

the Gołąb–Schinzel equation ([2], see also [1], pp. 311–312)

$$\phi(x+y\phi(x))=\phi(x)\phi(y), \tag{1}$$

or the equation [3]

$$\phi(x+y\phi(x))+\phi(x-y\phi(x))=2\phi(x)\phi(y), \tag{2}$$

are the examples. In section 1, we consider more general functional equations than (1) and (2) and give some conditions allowing us to determine their one-to-one solutions, bijective solutions or the solutions having exactly one zero. In section 2, for an arbitrarily fixed real p, we deal with the functional equation

$$\phi\left([p\,\phi(y)+(1-p)]x+[(1-p)\phi(x)+p\,]y\right)=\phi(x)\phi(y), \quad x,y\in\mathbb{R},$$

being a special generalization of equation (1).

2. Main result

Let X be a set. For a function $\phi: X\to X$ and a positive integer number k, by the symbol ϕ^k we denote the kth iteration of the function ϕ.

The following result reduces the problem of determining the solutions of a functional equation of a composite type to an application of the implicit function theorem.

Theorem 1. *Let $m, n\in\mathbb{N}$ be fixed. Let I, $I_1\subseteq\mathbb{R}$ be intervals such that $0\in I_1$ and $I_1\subset I$. Let $G:(I\times I_1)^2\mapsto I$ and $H:(I\times I_1^n)\times(I\times I_1^m)\longmapsto I_1$. Suppose that for all $x,y\in I$; $x_1,\dots,x_n,y_2,\dots,y_m\in I_1$,*

$$H(x,x_1,x_2,\dots,x_n,y,0,y_2,\dots,y_m)=0. \tag{3}$$

If a function $\phi: I\longmapsto I_1$ satisfies the functional equation

$$\phi(G(x,\phi(x),y,\phi(y)))=H(x,\phi(x),\phi^2(x),\dots,\phi^n(x),y,\phi(y),\phi^2(y)\dots,\phi^m(y)) \tag{4}$$

for all x, $y\in I$ and there exists exactly one $z_0\in I$ such that $\phi(z_0)=0$, then

$$G(x,\phi(x),z_0,0)=z_0, \qquad x\in I.$$

Proof. Taking $y=z_0$ in equation (4) and applying condition (3), we get

$$\phi(G(x,\phi(x),z_0,0))=0, \qquad x\in I.$$

Since ϕ has exactly one zero, we obtain $G(x,\phi(x),z_0,0)=z_0$ for all $x\in I$. This completes the proof. □

Remark 1. *Equation (4) generalizes the Gołąb–Schinzel equation (1).*

In what follows, for $p \in \mathbb{R}$ and $\phi : X \to (0,\infty)$ the symbol $X \ni x \to [\phi(x)]^p$ stands for the superposition of the power function $(0,\infty) \ni u \to u^p$ and ϕ.

Now we present some applications of Theorem 1.

Corollary 1. *Let k, $l \in \mathbb{N}$ be fixed and let $\phi : \mathbb{R} \longmapsto \mathbb{R}$ be a function with exactly one zero point. Then ϕ satisfies the functional equation*

$$\phi\left(x + y[\phi(x)]^{\frac{2k-1}{2l-1}}\right) = \phi(x)\phi(y), \qquad x, y \in \mathbb{R}, \tag{5}$$

if and only if for some $c \in \mathbb{R}$, $c \neq 0$,

$$\phi(x) = (cx+1)^{\frac{2l-1}{2k-1}}, \qquad x \in \mathbb{R}\,. \tag{6}$$

Proof. In Theorem 1 take $I = I_1 = \mathbb{R}$, $n = m = 1$ and define $G : \mathbb{R}^4 \longmapsto \mathbb{R}$ by

$$G(x, x_1, y, y_1) := x + y(x_1)^{\frac{2k-1}{2l-1}}, \qquad x,\, x_1,\, y\,, y_1 \in \mathbb{R},$$

and $H : \mathbb{R}^4 \longmapsto \mathbb{R}$ by

$$H(x, x_1, y, y_1) := x_1 y_1, \qquad x,\, x_1,\, y,\, y_1 \in \mathbb{R}.$$

Suppose that $\phi : \mathbb{R} \to \mathbb{R}$ satisfies equation (5) and has exactly one zero $z_0 \in \mathbb{R}$. Since $H(x, x_1, z_0, 0) = 0$ for all x, $x_1 \in \mathbb{R}$, the assumptions of Theorem 1 are fulfilled. From (5), applying Theorem 1, we get

$$G(x, \phi(x), z_0, 0) = z_0, \qquad x \in \mathbb{R},$$

that is

$$x + z_0[\phi(x)]^{\frac{2k-1}{2l-1}} = z_0, \qquad x \in \mathbb{R},$$

whence $z_0 \neq 0$ and

$$\phi(x) = \left(1 - \frac{x}{z_0}\right)^{\frac{2l-1}{2k-1}}, \qquad x \in \mathbb{R}.$$

Putting here $c := -\frac{1}{z_0}$, we obtain (6). Since ϕ given by (6) satisfies equation (5), the proof is completed. □

Remark 2. *It is known that (cf. [1], pp. 132-133) if $\phi : \mathbb{R} \longmapsto \mathbb{R}$ is a continuous solution of the Gołąb–Schinzel equation*

$$\phi(x + y\phi(x)) = \phi(x)\phi(y), \qquad x, y \in \mathbb{R},$$

then there exists $c \in \mathbb{R} \setminus \{0\}$ *such that either*

$$\phi(x) = \sup\{cx + 1, 0\}, \qquad x \in \mathbb{R},$$

or there exists $c \in \mathbb{R}$ *such that*

$$\phi(x) = cx + 1, \qquad x \in \mathbb{R}, \tag{7}$$

or

$$\phi(x) = 0, \qquad x \in \mathbb{R}.$$

The second solution can be obtained from Corollary 1 in a different way. Taking $k = l$ in the equation (5) and applying Corollary 1, we obtain (7) as a only solution having only zero in $\mathbb{R}$.

Corollary 2. *Let* $a < 0$ *and* $p \in \mathbb{R}$, $p > 0$, *be fixed. Suppose that* $\phi : [a, \infty) \longmapsto [0, \infty)$ *has exactly one zero in* $[a, \infty)$. *A function* ϕ *satisfies the functional equation*

$$\phi\left(x + y[\phi(x)]^p\right) = \phi(x)\phi(y), \qquad x \geq a,\ y \geq 0, \tag{8}$$

if and only if

$$\phi(x) = \left(1 - \frac{x}{a}\right)^{\frac{1}{p}}, \qquad x \geq a. \tag{9}$$

Proof. Suppose that $\phi : [a, \infty) \longmapsto [0, \infty)$ satisfies equation (8) and $z_0 \geq a$ is the only zero of ϕ. In Theorem 1 take $n = m = 1$, $I := [a, \infty)$, $I_1 := [0, \infty)$, the function $G : (I \times I_1)^2 \longmapsto I$ defined by

$$G(x, x_1, y, y_1) := x + y(x_1)^p, \qquad x, y \in I,\ \ x_1, y_1 \in I_1,$$

and the function $H : (I \times I_1)^2 \longmapsto I_1$ defined by

$$H(x, x_1, y, y_1) := x_1 y_1, \qquad x, y \in I,\ \ x_1, y_1 \in I_1.$$

Since $H(x, x_1, y, 0) = 0$, for all $x,\ y \in I,\ x_1 \in I_1$, the assumptions of Theorem 1 are satisfied. Therefore

$$G(x, \phi(x), z_0, 0) = z_0, \qquad x \in I,$$

so $x + z_0[\phi(x)]^p = z_0$ for all $x \geq a$. It follows that $z_0 \neq 0$ and, consequently,

$$\phi(x) = (1 - \frac{x}{z_0})^{\frac{1}{p}}, \qquad x \geq a.$$

Since ϕ is non–negative, we have $1 - \frac{x}{z_0} \geq 0$ for all $x \in [a, \infty)$. Thus $z_0 = a$. Since the converse implication is easy to verify, the proof is completed. $\square$

Remark 3. *Note that for $p = 0$ equation (8) in Corollary 2 becomes the Cauchy functional equation.*

Theorem 2. *Let $n \in \mathbb{N}$ be fixed. Let I, I_1 be intervals such that $I_1 \subset I \subseteq \mathbb{R}$. Let $G : (I \times I_1)^2 \longmapsto I$ and $H : (I \times I_1^n)^2 \longmapsto I_1$ be given functions. Suppose that H is symmetric, that is*

$$H(x, x_1, x_2, \dots, x_n, y, y_1, y_2, \dots, y_n) = H(y, y_1, y_2, \dots, y_n, x, x_1, x_2, \dots, x_n) \tag{10}$$

for all $x, y \in I$, $x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n \in I_1$.

If $\phi : I \longmapsto I_1$ is a solution of the functional equation

$$\phi(G(x, \phi(x), y, \phi(y))) = H(x, \phi(x), \phi^2(x), \dots, \phi^n(x), y, \phi(y), \phi^2(y) \dots, \phi^n(y)) \tag{11}$$

for all $x, y \in I$, then

$$\phi(G(x, \phi(x), y, \phi(y))) = \phi(G(y, \phi(y), x, \phi(x))), \qquad x, y \in I.$$

If, moreover ϕ is one-to-one function, then

$$G(x, \phi(x), y, \phi(y)) = G(y, \phi(y), x, \phi(x)), \qquad x, y \in I. \tag{12}$$

Proof. Suppose that $\phi : I \longmapsto I_1$ satisfies Eq. (11) and $H : (I \times I_1^n)^2 \longmapsto I_1$ satisfies condition (10). Then for all x, $y \in I$ we have

$$\begin{aligned}\phi(G(x, \phi(x), y, \phi(y))) &= H(x, \phi(x), \phi^2(x), \dots, \phi^n(x), y, \phi(y), \phi^2(y) \dots, \phi^n(y)) \\ &= H(y, \phi(y), \phi^2(y), \dots, \phi^n(y), x, \phi(x), \phi^2(x) \dots, \phi^n(x)) \\ &= \phi(G(y, \phi(y), x, \phi(x))),\end{aligned}$$

so,

$$\phi(G(x, \phi(x), y, \phi(y))) = \phi(G(y, \phi(y), x, \phi(x))), \qquad x, y \in I.$$

If ϕ is one-to-one, then obviously equality (12) holds true.

Remark 4. *If the function G in Theorem 2 is not symmetric, then in general equality (12) allows us to obtain the one-to-one solutions of (11).*

Applying Theorem 2 we obtain

Corollary 3. *Let $a,\ p \in \mathbb{R}$ be fixed and such that $a < 0,\ p \neq 0$. A one-to-one function $\phi : (a,\infty) \longmapsto (0,\infty)$ satisfies the functional equation*

$$\phi\left(x + y[\phi(x)]^p\right) = \phi(x)\phi(y)\,, \qquad x > a,\ y \geq 0, \tag{13}$$

if, and only if,

$$\phi(x) = \left(1 - \frac{x}{a}\right)^{\frac{1}{p}}, \qquad x > a. \tag{14}$$

Proof. In Theorem 2 take $n = 1,\ I = (a,\infty),\ I_1 = (0,\infty)$, the function $G : (I \times I_1)^2 \longmapsto I$ defined by

$$G(x, x_1, y, y_1) := x + y(x_1)^p, \qquad x,\ y \in I,\ \ x_1,\ y_1 \in I_1,$$

and $H : (I \times I_1)^2 \longmapsto I_1$ defined by

$$H(x, x_1, y, y_1) := x_1 y_1 \qquad x,\ y \in I,\ \ x_1,\ y_1 \in I_1,$$

Since

$$H(x, x_1, y, y_1) = H(y, y_1, x, x_1), \qquad x,\ y \in I,\ \ x_1,\ y_1 \in I_1,$$

the assumptions of Theorem 2 are satisfied. Applying Theorem 2, we have from (12):

$$x + y[\phi(x)]^p = y + x[\phi(y)]^p, \qquad x,\ y \in I,$$

whence

$$\frac{[\phi(x)]^p - 1}{x} = \frac{[\phi(y)]^p - 1}{y}, \qquad x,\ y \in I, \qquad x,\ y \neq 0.$$

So, there exists a constant $c \in \mathbb{R} \setminus \{0\}$ such that

$$x^{-1}([\phi(x)]^p - 1) = c$$

for all $x \in I,\ x \neq 0$. Hence

$$\phi(x) = (cx + 1)^{\frac{1}{p}}, \qquad x > a,\ \ x \neq 0.$$

Equation (13) implies that

$$cx + 1 > 0, \qquad x > a,$$

and, consequently, $ca+1 \geq 0$. On the other hand, if ϕ satisfies equation (13), then obviously the following inequality

$$x + y[(cx+1)^{\frac{1}{p}}]^p > a, \qquad x,\, y > a,$$

is true, which means that

$$x + y[(cx+1)] > a, \qquad x,\, y > a.$$

It follows that $a + ca^2 + a \geq a$, so $ca + 1 \leq 0$. Both inequalities imply that $ca + 1 = 0$, whence $c = -\frac{1}{a}$, and ϕ has to be of the form (14).

To show that the function ϕ given by (14) satisfies equation (13), let us note that

$$x + y[\phi(x)]^p > a, \qquad x, y > a.$$

In fact, this inequality is equivalent to $(x-a)(y-a) > 0$. Now, it is easy to verify that (14) satisfies equation (13). This completes the proof. □

Remark 5. *Taking $a, p \in \mathbb{R}$, $a < 0$, and $p > 0$, we can show in the same way that the one–to–one function $\phi : [a, +\infty) \longmapsto [0, +\infty)$ satisfies the functional equation*

$$\phi\left(x + y[\phi(x)]^p\right) = \phi(x)\phi(y), \qquad x,\, y \geq a,$$

if and only if

$$\phi(x) = \left(1 - \frac{x}{a}\right)^{\frac{1}{p}}, \qquad x \geq a.$$

Remark 6. *Let I, $I_1 \subseteq \mathbb{R}$ be intervals. Let $G : (I \times I_1)^2 \longmapsto I$ and $H : I_1 \times I_1 \longmapsto I_1$ be the given functions. Assume that $\phi : I \longmapsto I_1$, $\phi(I) = I_1$ is a bijective solution of the functional equation*

$$\phi(G(x, \phi(x), y, \phi(y))) = H(\phi(x), \phi(y)), \qquad x, y \in I. \tag{15}$$

Then the function $\phi^{-1} : I_1 \longmapsto I$ satisfies the (non-composite) functional equation

$$G(\phi^{-1}(x), x, \phi^{-1}(y), y) = \phi^{-1}(H(x, y)), \qquad x, y \in I_1. \tag{16}$$

In fact, putting $\phi^{-1}(x)$ in place of x and $\phi^{-1}(y)$ in place of y in equation (15), we obtain (16).

Sometimes the above remark allows us to determine effectively the bijective solutions for functional equations of form (15). We have the following

Corollary 4. *Let $k,\ l \in \mathbb{N}$ be fixed and let $I = I_1 = \mathbb{R}$. The bijection function $\phi : \mathbb{R} \longmapsto \mathbb{R}$ satisfies functional equation*

$$\phi\left(x + y[\phi(x)]^{\frac{2k-1}{2l-1}}\right) = \phi(x)\phi(y), \qquad x, y \in \mathbb{R}, \tag{17}$$

if and only if

$$\phi(x) = (cx+1)^{\frac{2l-1}{2k-1}}, \qquad x \in \mathbb{R}, \tag{18}$$

for some $c \in \mathbb{R},\ c \neq 0$.

Proof. According to Remark 6, a bijection $\phi : \mathbb{R} \mapsto \mathbb{R}$ satisfies equation (17) if and only if $\phi^{-1} : \mathbb{R} \mapsto \mathbb{R}$ satisfies the equation

$$\phi^{-1}(x) + \phi^{-1}(y)x^{\frac{2k-1}{2l-1}} = \phi^{-1}(xy), \qquad x, y \in \mathbb{R}.$$

Putting here $y = 0$, we obtain

$$\phi^{-1}(x) = \phi^{-1}(0)\left(1 - x^{\frac{2k-1}{2l-1}}\right), \qquad x \in \mathbb{R},$$

which implies (18). □

3. A special generalization of Gołąb–Schinzel functional equation

In this section we examine the functional equation

$$\phi\left([p\,\phi(y) + (1-p)]x + [(1-p)\phi(x) + p\,]y\right) = \phi(x)\phi(y), \quad x, y \in \mathbb{R}, \tag{19}$$

where $p \in \mathbb{R}$ is an arbitrarily fixed parameter. For $p = 0$ or $p = 1$ it reduces to the classical Gołąb–Schinzel equation.

Theorem 3. *Let $p \in \mathbb{R}$ be fixed.*

1. *If $p \neq \frac{1}{2}$, then the one-to-one function $\phi : \mathbb{R} \longmapsto \mathbb{R}$ satisfies (19) if and only if*

$$\phi(x) = cx + 1, \qquad x \in \mathbb{R},$$

for some $c \in \mathbb{R} \setminus \{0\}$.

2. *If $p = \frac{1}{2}$, then bijection $\phi : \mathbb{R} \longmapsto \mathbb{R}$ satisfies (19) if and only if*

$$\phi(x) = cx + 1, \qquad x \in \mathbb{R},$$

for some $c \in \mathbb{R} \setminus \{0\}$.

Proof. Take $n = 1$, $I = \mathbb{R}$ and define $G : (\mathbb{R} \times \mathbb{R})^2 \longmapsto \mathbb{R}$ by

$$G(x, x_1, y, y_1) := [py_1 + (1-p)]x + [(1-p)x_1 + p]y, \qquad x,\, y,\, x_1,\, y_1 \in \mathbb{R},$$

and $H : (\mathbb{R} \times \mathbb{R})^2 \longmapsto \mathbb{R}$ by

$$H(x, x_1, y, y_1) := x_1 y_1, \qquad x,\, y,\, x_1,\, y_1 \in \mathbb{R}.$$

Note, that

$$H(x, x_1, y, y_1) = H(y, y_1, x, x_1), \qquad x,\, y,\, x_1,\, y_1 \in \mathbb{R}.$$

Applying Theorem 2, we obtain

$$[p\phi(y) + (1-p)]x + [(1-p)\phi(x) + p]y$$
$$= [p\phi(x) + (1-p)]y + [(1-p)\phi(y) + p]x$$

for all $x,\, y \in \mathbb{R}$, whence

$$(2p-1)[x(\phi(y)-1)] = (2p-1)[y(\phi(x)-1)], \qquad x,\, y \in \mathbb{R}.$$

If $p \neq \frac{1}{2}$, hence we get

$$\frac{\phi(x)-1}{x} = \frac{\phi(y)-1}{y}, \qquad x,\, y \in \mathbb{R} \setminus \{0\}.$$

Therefore, there exists a constant $c \in \mathbb{R} \setminus \{0\}$ such that $\phi(x) = cx + 1$ for all $x \in \mathbb{R} \setminus \{0\}$. Putting $x = y = 0$ in equation (19), we get $[\phi(0)] = [\phi(0)]^2$, consequently we obtain either $\phi(0) = 0$ or $\phi(0) = 1$. Since ϕ is one–to–one and $\phi(-\frac{1}{c}) = 0$, the case $\phi(0) = 0$ cannot occur. Thus $\phi(x) = cx + 1$ for all $x \in \mathbb{R}$.

For $p = \frac{1}{2}$ equation (19) has the form:

$$\phi\left(\frac{1}{2}[x(\phi(y)+1) + y(\phi(x)+1)]\right) = \phi(x)\phi(y), \qquad x,\, y \in \mathbb{R}. \tag{20}$$

If a bijection $\phi : \mathbb{R} \longmapsto \mathbb{R}$ satisfies (20), then according to the Remark 2 the function $\phi^{-1} : \mathbb{R} \longmapsto \mathbb{R}$ satisfies the equation

$$2\phi^{-1}(xy) = (y+1)\phi^{-1}(x) + (x+1)\phi^{-1}(y), \qquad x,\, y \in \mathbb{R}.$$

Putting here $y = 0$, we get

$$2\phi^{-1}(0) = \phi^{-1}(x) + (x+1)\phi^{-1}(0), \qquad x \in \mathbb{R}.$$

Hence, as $\phi^{-1}(0) \neq 0$,

$$\phi^{-1}(x) = \phi^{-1}(0)(1-x), \qquad x \in \mathbb{R},$$

whence

$$\phi(x) = 1 - \frac{1}{\phi^{-1}(0)}\, x, \qquad x \in \mathbb{R}. \qquad \square$$

Theorem 4. *Let $p \in \mathbb{R}$ be fixed. A function $\phi : \mathbb{R} \longmapsto \mathbb{R}$ satisfies equation (19) and has exactly one zero if and only if there exists a constant $c \in \mathbb{R} \setminus \{0\}$ such that*

$$\phi(x) = cx + 1, \qquad x \in \mathbb{R}.$$

Proof. Note that substitution of $(1-p)$ for p in equation (19) gives the same equation. Thus, without any loss of generality, we can assume that $p \neq 1$. Take $n = m = 1$, $I = I_1 = \mathbb{R}$, and define $H : (\mathbb{R} \times \mathbb{R})^2 \mapsto \mathbb{R}$ by

$$H(x, x_1, y, y_1) := x_1 y_1, \qquad x,\, x_1,\, y,\, y_1 \in \mathbb{R}, \tag{21}$$

and $G : (\mathbb{R} \times \mathbb{R})^2 \mapsto \mathbb{R}$ by

$$G(x, x_1, y, y_1) := [py_1 + (1-p)]x + [(1-p)x_1 + p]y, \qquad x, x_1, y, y_1 \in \mathbb{R}.$$

Suppose that $\phi : \mathbb{R} \to \mathbb{R}$ satisfies equation (19) and $z_0 \neq 0$ is a unique zero of ϕ. Note that if $y = z_0$, then $H(x, x_1, z_0, 0) = 0$ for all $x,\, x_1 \in \mathbb{R}$, so the function (21) satisfies the condition (3) of Theorem 1. Therefore, if ϕ satisfies equation (19), then

$$(1-p)x + [(1-p)\phi(x) + p]z_0 = z_0, \qquad x \in \mathbb{R}.$$

Hence we obtain $\phi(x) = 1 - \frac{x}{z_0}$ for all $x \in \mathbb{R}$. $\square$

References

[1] J. Aczél. *Lectures on Functional Equations and Their Applications*, Academic Press, New York, 1966.

[2] S. Gołąb, A. Schinzel. Sur l'équation fonctionnelle $f[x + yf(x)] = f(x)f(y)$. *Publ. Math. Debrecen*, **6**, 113–125, 1959.

[3] P. Kahlig, J. Matkowski. On some extension of Gołąb–Schinzel functional equation. *Ann. Math. Siles.*, **8**, 13–31, 1994.

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XV, Częstochowa 2010

THE BOUNDED LOCAL OPERATORS IN THE BANACH SPACE OF HÖLDER FUNCTIONS

Janusz Matkowski[a,b], **Małgorzata Wróbel**[c]

[a] *Faculty of Mathematics, Computer Science and Econometrics*
University of Zielona Góra, Podgórna 50
65246 Zielona Góra, Poland
e-mail: J.Matkowski@wmie.uz.zgora.pl

[b] *Institute of Mathematics, Silesian University*
Bankowa 14, 0007 Katowice, Poland

[c] *Institute of Mathematics and Computer Science*
Jan Długosz University in Częstochowa
Armii Krajowej 13/15, 42-200 Częstochowa, Poland
e-mail: m.wrobel@ajd.czest.pl

Abstract. It is known that every locally defined operator acting between two Hölder spaces is a Nemytskii superposition operator. We show that if such an operator is bounded in the sense of the norm, then its generator is continuous.

1. Introduction

Let $I \subset \mathbb{R}$ be an arbitrary interval and by $\mathbb{R}^I$ we denote the set of all functions $\varphi : I \to \mathbb{R}$. For a given two-place function $h : I \times \mathbb{R} \to \mathbb{R}$, the mapping $K : \mathbb{R}^I \to \mathbb{R}^I$ defined by

$$K(\varphi)(x) := h(x, \varphi(x)), \quad \varphi \in \mathbb{R}^I,\ x \in I,$$

is called a Nemytskii superposition operator of the generator h.

It is known that every locally defined operator mapping the set of continuous functions $C(I, \mathbb{R})$ into itself must be a superposition operator [2]. Moreover, K maps $C(I, \mathbb{R})$ into itself if and only if its generator h is continuous. At this background it is surprising enough that there are discontinuous

functions $h : I \times \mathbb{R} \to \mathbb{R}$ generating the superpositions operators K which map the space of continuously differentiable functions $C^1(I, \mathbb{R})$ into itself (cf. [1, p. 209]). In [3] it has been proved that if a locally defined operator maps the Banach space $H_\phi(I, \mathbb{R})$ of all Hölder functions $\varphi : I \to \mathbb{R}$ into $H_\psi(I, \mathbb{R})$, then it is a Nemytskii superposition operator. The purpose of this paper is to show that if, additionally, K is bounded with respect to $H_\phi(I, \mathbb{R})$–norm, then its generator must be continuous.

2. Main result

Let $\phi : (0, \infty) \to (0, \infty)$ satisfy the following condition:

(*i*) ϕ is strictly increasing, $\phi(0+) := \lim_{t \to 0+} \phi(t) = 0$ and the function

$$(0, \infty) \ni t \to \frac{\phi(t)}{t}$$

is decreasing.

Let us note the following (easy to verify)

Remark 1. If $\phi : (0, \infty) \to (0, \infty)$ satisfies condition (*i*), then ϕ is subadditive and continuous.

Let $I \subset \mathbb{R}$ be an interval and let $x_0 \in I$ be arbitrarily fixed. For a given $\phi : (0, \infty) \to (0, \infty)$, having the above properties, by $H_\phi(I, \mathbb{R})$ we denote the Banach space of all Hölder functions $\varphi : I \to \mathbb{R}$ equipped with the norm

$$\|\varphi\|_\phi := |\varphi(x_0)| + \sup_{x,y \in I, x \neq y} \frac{|\varphi(x) - \varphi(y)|}{\phi(|x - y|)}.$$

Clearly, $\varphi \in H_\phi(I, \mathbb{R})$ if and only if there exists a constant $c > 0$ such that

$$|\varphi(x) - \varphi(y)| \leq c\phi(|x - y|), \quad x, y \in I.$$

Let us notice that if $\phi(t) = t^\alpha$ for some $\alpha \in (0, 1]$, then $H_\alpha(I, \mathbb{R}) := H_\phi(I, \mathbb{R})$ is the classical Hölder functions space and $H_1(I, \mathbb{R})$ becomes the Banach space of Lipschitz functions.

Definition. *Let $\phi, \psi : (0, \infty) \to (0, \infty)$ satisfy condition (i). An operator $K : H_\phi(I, \mathbb{R}) \to H_\psi(I, \mathbb{R})$ is said to be locally defined if for any open interval $J \subset \mathbb{R}$ and for any functions $\varphi, \psi \in H_\phi(I, \mathbb{R})$,*

$$\varphi\big|_{J \cap I} = \psi\big|_{J \cap I} \Rightarrow K(\varphi)\big|_{J \cap I} = K(\psi)\big|_{J \cap I},$$

where $\phi\big|_{J \cap I}$ denotes the restriction of φ to $J \cap I$.

In [3] the following result was proved:

Theorem 1. ([3], Corollary 2). *Let $I \subset \mathbb{R}$ be an interval. If a locally defined operator K maps $H_\phi(I,\mathbb{R})$ into $H_\psi(I,\mathbb{R})$, then there exists a unique function $h : I \times \mathbb{R} \to \mathbb{R}$ such that*

$$K(\varphi)(x) = h(x, \varphi(x)), \quad (x \in I),$$

for all $\varphi \in H_\phi(I,\mathbb{R})$, that is K is a Nemytskii operator of the generator h.

We say that an operator $K : H_\phi(I,\mathbb{R}) \to H_\psi(I,\mathbb{R})$ is bounded if it maps the convergent sequences of $H_\phi(I,\mathbb{R})$ into bounded sequences in $H_\psi(I,\mathbb{R})$.

The main result reads as follows:

Theorem 2. *Let $I \subset \mathbb{R}$ be an interval. If a locally defined operator $K : H_\phi(I,\mathbb{R}) \to H_\psi(I,\mathbb{R})$ is bounded, then there exists a continuous function $h : I \times \mathbb{R} \to \mathbb{R}$ such that*

$$K(\varphi)(x) = h(x, \varphi(x)); \quad \varphi \in H_\phi(I,\mathbb{R}), \quad (x \in I).$$

Proof. By Theorem 1, there exists a function $h : I \times \mathbb{R} \to \mathbb{R}$ such that the formula of our result holds true. We shall show that h is continuous.

Without any loss of generality we can assume that $I = [a, b)$, where $0 < b \leq +\infty$, and that

$$\|\varphi\|_\phi := |\varphi(a)| + \sup_{x,y \in I, x \neq y} \frac{|\varphi(x) - \varphi(y)|}{\phi(|x - y|)}.$$

First we show that h is continuous with respect to the second variable. To this end let us fix $(x_0, y_0) \in I$ and choose arbitrarily a real sequence $(y_n)_{n\in\mathbb{N}}$ such that

$$y_n \neq y_0, \quad n \in \mathbb{N}, \quad \lim_{n\to\infty} y_n = y_0. \tag{1}$$

Let $(x_n)_{n\in\mathbb{N}}$ be a sequence such that $x_n \in I,\ n \in \mathbb{N}$, and

$$|x_n - x_0| = \phi^{-1}\left(\sqrt{|y_n - y_0|}\right), \quad n \in \mathbb{N}.$$

Hence we obtain

$$\frac{|y_n - y_0|}{\phi(|x_n - x_0|)} = \frac{|y_n - y_0|}{\phi\left(\phi^{-1}\left(\sqrt{|y_n - y_0|}\right)\right)} = \sqrt{|y_n - y_0|}, \quad n \in \mathbb{N}. \tag{2}$$

Define the functions $P_{y_n} : I \to \mathbb{R}$, $\varphi_n : I \to \mathbb{R}$, $n \in \mathbb{N}$, by the following formulas:

$$P_{y_n}(x) := y_n, \quad n \in \mathbb{N}, \tag{3}$$

$$\varphi_n(x) = \begin{cases} y_0, & \text{for} \quad x \in [a, x_0], \\ \dfrac{y_n - y_0}{x_n - x_0}(x - x_0) + y_0 & \text{for} \quad x \in (x_0, x_n), \ n \in \mathbb{N}, \\ y_n, & \text{for} \quad x \in [x_n, b). \end{cases} \tag{4}$$

and put

$$\varphi_0(x) = y_0, \quad x \in I.$$

Of course,

$$P_{y_n}, \varphi_n \in H_\phi(I, \mathbb{R}), \quad n \in \mathbb{N}.$$

Since

$$\|P_{y_n} - \varphi_0\|_\phi = |y_n - y_0|, \quad n \in \mathbb{N},$$

applying (1) and (2), we get

$$\lim_{n\to\infty} \|P_{y_n} - \varphi_0\|_\phi = 0, \qquad \lim_{n\to\infty} \|\varphi_n - \varphi_0\|_\phi = 0. \tag{5}$$

Making use of (3), (4), the triangle inequality and by the definition of the norm, we have

$$\begin{aligned} |h(x_0, y_n) - h(x_0, y_0)| &\le |h(x_n, y_n) - h(x_0, y_n)| + |h(x_n, y_n) - h(x_0, y_0)| \\ &= |h(x_n, P_{y_n}(x_n) - h(x_0, P_{y_n}(x_0)| \\ &+ |h(x_n, \varphi_n(x_n)) - h(x_0, \varphi_n(x_0))| \\ &= |K(P_{y_n})(x_n) - K(P_{y_n})(x_0)| \\ &+ |K(\varphi_n)(x_n) - K(\varphi_n)(x_0)| \\ &= \frac{|K(P_{y_n})(x_n) - K(P_{y_n})(x_0)|}{\psi(|x_n - x_0|)} \psi(|x_n - x_0|) + \\ &+ \frac{|K(\varphi_n)(x_n) - K(\varphi_n)(x_0)}{\psi(|x_n - x_0|)} \psi(|x_n - x_0|) \\ &\le \|K(P_{y_n})\|_\psi \psi(|x_n - x_0|) + \|K(\varphi_n)\|_\psi \cdot \psi(|x_n - x_0|). \end{aligned}$$

Taking into account (5), the equality $\psi(0+) = 0$, the boundedness of the operator K and letting n tend to the infinity, we get the continuity of h with respect to the second variable.

To show that h is continuous fix $(x_0, y_0) \in I \times \mathbb{R}$, take two arbitrary sequences $x_n \in I$, $y_n \in \mathbb{R}$, $n \in \mathbb{N}$, convergent to x_0 and y_0, respectively, and define $P_{y_n} : I \to \mathbb{R}$, $n \in \mathbb{N} \cup \{0\}$, by

$$P_{y_n}(x) = y_n, \quad n \in \mathbb{N} \cup \{0\}.$$

Hence, by the triangle inequality and by the definition of the norm, we have

$$\begin{aligned}
|h(x_n,y_n)-h(x_0,y_0)| &\le |h(x_n,y_n)-h(x_0,y_n)|+|h(x_0,y_n)-h(x_0,y_0)| \\
&= |h(x_n,P_{y_n}(x_n))-h(x_0,P_{y_n}(x_0)| \\
&+|h(x_0,y_n)-h(x_0,y_0)| \\
&= |(K(P_{y_n})(x_n)-K(P_{y_n})(x_0)| \\
&+|h(x_0,y_n)-h(x_0,y_0)| \\
&= \frac{|K(P_{y_n})(x_n)-K(P_{y_n})(x_0)|}{\psi(|x_n-x_0|)}\cdot\psi(|x_n-x_0|) \\
&+|h(x_0,y_n)-h(x_0,y_0)| \\
&\le \|K(P_{y_n})\|_\psi\psi(|x_n,x_0|)+|h(x_0,y_n)-h(x_0,y_0)|.
\end{aligned}$$

Since, by the definition of P_{y_n}, $n\in\mathbb{N}\cup\{0\}$,

$$\lim_{n\to\infty}\|P_{y_n}-P_{y_0}\|_\phi=0,$$

applying the boundedness of the operator K, the equality $\psi(0+)=0$ and the first part of the proof, i.e. the continuity of h with respect to the second variable, letting n tend to the infinity, we get the required claim. □

Remark 2. Taking in the above theorem a compact interval $I\subset\mathbb{R}$, one gets Theorem 7.3 from [1].

To construct an example showing that the assumption of the boundedness of K is essential, we need the following

Lemma. *Let $(X,d),(Y,\rho)$ be metric spaces. Suppose $A,B\subset X$ are closed, $intA\cap intB=\emptyset$ and adjacent in the following sense: for any $x\in A$, $y\in B$ there exists a point $z\in\delta A\cap\delta B$ such that*

$$d(x,y)=d(x,z)+d(z,y). \tag{6}$$

If the functions $f:A\to Y$ and $g:B\to Y$ are Lipschitz continuous and

$$f(z)=g(z)\quad for\ all\quad z\in\delta A\cap\delta B,$$

then the function $h:(A\cup B)\to Y$ defined by

$$h(x):=\begin{cases} f(x) & for\quad x\in A,\\ g(x) & for\quad x\in B\end{cases}$$

is Lipschitz continuous. (Here δA stands for the boundary of A.)

Proof. Since f and g are Lipschitz continuous, there is $c \in \mathbb{R}_+$ such that

$$\rho(f(x), f(y)) \le cd(x, y) \quad \text{for } x, y \in A; \quad \rho(g(x), g(y)) \le cd(x, y) \ \text{ for } x, y \in B.$$

Take $x, y \in A \cup B$ and assume that $x \in A$ and $y \in B$. By assumption, there is $z \in \delta A \cap \delta B$ such that (6) holds. Hence, by the triangle inequality,

$$\rho(h(x), h(y)) \le \rho(h(x), h(z)) + \rho(h(z), h(y)) = \rho(f(x), f(z)) + \rho(g(z), g(y))$$

$$\le cd(x, z) + cd(z, y) = cd(x, y).$$

As the remaining two cases are obvious, the proof is complete. □

Example. Define a two-place function $h : [0, 1] \times \mathbb{R} \to \mathbb{R}$ by the formula

$$h(x, y) := \begin{cases} 0 & \text{if} \quad y \le 0, \\ \dfrac{y}{\sqrt{x}} & \text{if} \quad 0 < y \le \sqrt{x}, \\ 1 & \text{if} \quad y > \sqrt{x}. \end{cases} \tag{7}$$

Observe that h is continuous in $[0, 1] \times \mathbb{R}\backslash\{(0, 0)\}$ and discontinuous at the point $(0, 0)$. In fact we have more, namely outside of any neighbourhood of $(0, 0)$, by Lemma, the function h is Lipschitzian.

Denote by $\mathcal{F}[0, 1]$ the set of all functions $\varphi : [0, 1] \to \mathbb{R}$. Let $K : \mathcal{F}[0, 1] \to \mathcal{F}[0, 1]$ be the Nemytskii composition (so locally defined) operator generated by h, i.e.

$$K(\varphi)(x) := h(x, \varphi(x)), \quad x \in [0, 1].$$

We shall show that K maps the space $H_1([0, 1], \mathbb{R})$ of all Lipschitz continuous functions $\varphi : [0, 1] \to \mathbb{R}$ into itself.

Take $\varphi \in H_1([0, 1], \mathbb{R})$. If $\varphi(0) \neq 0$, then as h is Lipschitz continuous outside any neighbourhood of $(0, 0)$, the function $K(\varphi)$, as composition of Lipschitz continuous functions, is Lipschitz continuous in $[0, 1]$, so $K(\varphi) \in H_1([0, 1], \mathbb{R})$. If $\varphi(0) = 0$, then $K(\varphi)\big|_{[\varepsilon,1]}$ is Lipschitz continuous for any $\varepsilon \in (0, 1]$. In view of Lemma, it is enough to show that $K(\varphi)\big|_{[0,\varepsilon]}$ is Lipschitz continuous. To this end assume that φ satisfies the Lipschitz condition with a constant c, that is

$$|\varphi(x) - \varphi(\overline{x})| \le c|x - \overline{x}|, \quad x, \overline{x} \in [0, 1].$$

Setting $\overline{x} = 0$, we hence get

$$|\varphi(x)| \le cx, \quad x \in [0, 1],$$

so the graph of the function φ is contained in the triangle set

$$D := \{(x, y) : x \in [0, 1],\ |y| \leq cx\}.$$

If φ is nonpositive on any subinterval of $I \subset [0, 1]$, then, by the definition of h, we have $K(\varphi)\big|_I = 0$ and, obiously, $K(\varphi)$ is Lipschitz continiuous on I with zero Lipschitz constant. Therefore, it is enough to confine our considerations to the case when the graph of $\varphi\big|_{[0,\varepsilon]}$ is contained in the set

$$D_\varepsilon := \{(x, y) : x \in [0, \varepsilon],\ 0 \leq y \leq cx\}.$$

Let us choose $\varepsilon > 0$ such that $c < \frac{1}{\sqrt{\varepsilon}}$. Then, clearly $cx < \sqrt{x}$ for all $x \in (0, \varepsilon]$. Since for all $(x, y) \in D_\varepsilon$ we have

$$\left|\frac{\partial}{\partial x}h(x, y)\right| = \left|-\frac{y^2}{2x\sqrt{x}}\right| \leq \frac{(cx)^2}{2x\sqrt{x}} \leq \frac{c^2\sqrt{\varepsilon}}{2}$$

and

$$\left|\frac{\partial}{\partial y}h(x, y)\right| = \frac{2y}{\sqrt{x}} \leq \frac{2cx}{\sqrt{x}} \leq 2c\sqrt{\varepsilon},$$

we infer that $h\big|_{D\varepsilon}$ is Lipschitz continuous. It follows that $K(\varphi)\big|_{[0,\varepsilon]}$, as a composition of Lipchitz functions, is Lipschitz continuous.

We claim that K is unbounded. To see this take a sequence of constant functions convergent to zero, $\varphi_k : [0, 1] \to \mathbb{R},\ k \in \mathbb{N}$, defined by $\varphi_k(x) = \frac{1}{\sqrt{k}}$. According to (7), we get

$$K(\varphi_k)(x) = \begin{cases} 1 & \text{for} \quad 0 \leq x < \dfrac{1}{k} \\ \dfrac{1}{\sqrt{kx}} & \text{for} \quad \dfrac{1}{k} \leq x \leq 1 \end{cases} \qquad k \in \mathbb{N}.$$

Since

$$\|K(\varphi_k)\|_\psi \geq \left|\frac{\varphi_k(x) - \varphi_k(\overline{x})}{x - \overline{x}}\right|, \quad x, \overline{x} \in [0, 1], \quad x \neq y,$$

setting $x = \frac{4}{k},\ \overline{x} = 0$, for all $k \geq 4$, we get

$$\|K(\varphi_k)\|_\psi \geq \frac{k}{8}, \quad k \geq 4,$$

which shows that K is not bounded. □

References

[1] J. Appell, P.P. Zabrejko. *Nonlinear Superposition Operators.* Cambridge University Press, Cambridge, 1990.

[2] K. Lichawski, J. Matkowski, J. Miś. Locally defined operators in the space of differentiable functions. *Bull. Polish Acad. Sci. Math.*, **37**, 315–125, 1989.

[3] M. Wróbel. Locally defined operators in Hölder's spaces. *Nonlinear Analysis*, 2010. doi: 10.1016/j.na.2010.08.046.

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XV, Częstochowa 2010

A NOTE ON SI-SPACES AND MI-SPACES

Stanislav P. Ponomarev

Institute of Mathematics
Pomeranian University
ul Arciszewskiego 22a, 76-200 Słupsk, Poland
e-mail: p35st9@poczta.onet.pl

Abstract. We show that if there exists a second κ-category (or κ-Baire) SI-space, then there exists a second κ-category (resp. κ-Baire) MI-space. Next we discuss some properties of real functions on such spaces.

1. Preliminaries and basic definitions

The topic of our research stems from the ω-problem formulated below (see also [1]), which initially and formally had nothing in common with the spaces under discussion. The connections appeares in the way of analyzing the problem for non-metrizable spaces.

Although we retain all the definitions and notation from [1], we recall some of them for convenience of the reader.

Let $X = (X, \tau)$ be a topological space. To each function $F : X \to \mathbb{R}$ we associate the upper and lower Baire functions

$$M(F, \cdot) : X \to \overline{\mathbb{R}}, \;\; m(F, \cdot) : X \to \overline{\mathbb{R}}$$

defined in a usual way (see [1]). It is well known that $M(F, \cdot)$ is upper semicontinuous (USC), while $m(F, \cdot)$ is lower semicontinuous (LSC) on X.

The value

$$\omega(F, x) = M(F, x) - m(F, x) \in [0, \infty]$$

is called the oscillation of F at a point x.

We can also give an equivalent definition:

$$\omega(F, x) = \inf_{U} \sup_{x', x'' \in U} (F(x') - F(x'')),$$

where the infimum is taken over all elements U of a neighborhood base τ_x of τ at x.

Let $X = (X, \tau)$ be a topological space and a USC function $f : X \to [0, \infty]$ be given. If there exists a function $F : X \to \mathbb{R}$ such that

$$\forall x \in X : \ \omega(F, x) = f(x),$$

then we call F an ω-primitive for f.

By the "ω-problem" on a topological space X we mean the problem of the existence of an ω-primitive for a given USC function $f : X \to [0, \infty]$.[1]

In what follows, we consider only dense-in-themselves topological spaces and finite USC functions f.

In [2] it was shown that the ω-problem is solvable for each metric space. For a non-metrizable space the ω-problem need not be solvable what was shown in the case of an irresolvable space (see, e.g. [1], Theorem 4).

The notion of a resolvable (irresolvable) space was introduced in [3], where the basic properties of such spaces were given. Further, we will discuss the following two special classes of irresolvable spaces introduced in [3].

A dense-in-itself topological space $X = (X, \tau)$ is called an MI-space (or simply, MI) if every dense subset of (X, τ) is open.

A dense-in-itself topological space $X = (X, \tau)$ is called an SI-space (or simply, SI) if X has no resolvable subsets. Each MI-space is an SI-space [3].

We often write X instead of (X, τ). Closure of E is denoted by $\overline{E}$. The phrase "$E \subset X$ is τ-open (or τ-closed, τ-dense, etc.)" means that E is so with respect to the topology τ on X. Similarly, by $\mathrm{Int}_\tau E$ we denote the interior of E with respect to the topology τ. The symbol τ is omitted when no confusion could arise.

2. On second category MI-spaces and Baire MI-spaces

The notions of a first category (second category) set and of a Baire space will be considered in some generalized sense. Namely, we adopt the following definitions (see [4], [5]). Let κ be a cardinal, $\kappa > \aleph_0$.

Definition 1. *A set $E \subset X = (X, \tau)$ is of the first κ-category if it can be written in the form*

$$E = \bigcup_{\alpha \in A} E_\alpha,$$

where $\mathrm{card}\, A < \kappa$ and each E_α is nowhere dense in X.

A set $E \subset X = (X, \tau)$ is of the second κ-category if it is not of the first κ-category.

[1] Problems of this type in various settings and different terminology have been studied by many authors. Some results can be found in References which, however, are far from being complete.

Definition 2. *A topological space $X = (X, \tau)$ is called κ-Baire if the intersection of fewer than κ dense open subsets of X is dense in X.*

Recall that the definitions of a "usual" first (second) category set and of a Baire space correspond to $\kappa = \aleph_1$ and that each second κ-category set (κ-Baire space) is at the same time a "usual" second category set (resp. Baire space).

Definition 3. ([5]). *A space $X = (X, \tau)$ is called κ-SIB if it is a κ-Baire SI-space. We also say that X is a κ-SIB-space.*

In a similar way, we give

Definition 4. *A space $X = (X, \tau)$ is called κ-MIB (or κ-MIB-space) if X is a κ-Baire MI-space.*

Although initiated by the ω-problem, the propositions we are going to prove in this section were motivated by [5] and [6].

In [5] the authors obtained consistency and existence results concerning κ-SIB-spaces. Their methods used the theory of ideals on cardinals.

Our goal is far more simple. Namely, we are going only to show that if there exists a κ-SIB-space (or a second κ-category SI-space), then there exists a corresponding MI-space, i.e. a κ-MIB-space (or, respectively, a second κ-category MI-space). Some properties of functions and the ω-problem for such spaces will be discussed in Section 3.

Let $X = (X, \tau)$ be a topological space. Following [6], let $D(X, \tau)$ denote the family of all dense subsets of (X, τ).

By $\mathfrak{F}(X, \tau)$ we denote the family of filters $\mathscr{F}$ on (X, τ) consisting of dense subsets of (X, τ). It is clear that $\mathfrak{F}(X, \tau)$ is partially ordered by the usual inclusion relation.

Lemma 1. ([6], *Lemma* 3.3). Let $X = (X, \tau)$ be a topological space. Then there exists an ultrafilter $\mathscr{F}_m \in \mathfrak{F}(X, \tau)$.

Given a topological space (X, τ) and a filter $\mathscr{F} \in \mathfrak{F}(X, \tau)$, one may produce a finer topology $\hat{\tau}$ on X generated by the family $\tau \cup \mathscr{F}$. By definition, the basis for $\hat{\tau}$ consists of all intersections $U \cap E$, where $\emptyset \neq U \in \tau$ and $E \in \mathscr{F}$ (see [6]).

It is convenient to state the next two theorems of this section in the form of the following Proposition from [6]. Only category and baireness will be new items and this is exactly the object of our consideration.

Lemma 2. ([6], *Proposition* 3.4). Let $X = (X, \tau)$ be a dense-in-itself T_1 (or Hausdorff) space. Let $\mathscr{F}_m \in \mathfrak{F}(X, \tau)$ be an ultrafilter. Define $\hat{\tau}$ to be the topology generated by $\tau \cup \mathscr{F}_m$. Then

(i) $D(X, \hat{\tau}) = \mathscr{F}_m$;

(ii) $(X, \hat{\tau})$ is an MI–space which is T_1 (respectively, Hausdorff);

(iii) if (X, τ) is connected, then so is $(X, \hat{\tau})$.

Lemma 3. *(*[3]*, Theorem* 29*).* Every dense subset of an SI–space has dense interior.

Now we will prove the first main result of this section.

Theorem 1. *Assume that there exists a second κ-category T_1 (or Hausdorff) space (X, τ) which is SI. Let $\mathscr{F}_m \in \mathfrak{F}(X, \tau)$ be an ultrafilter and let $\hat{\tau}$ be a topology on X generated by $\tau \cup \mathscr{F}_m$. Then*

(i) $D(X, \hat{\tau}) = \mathscr{F}_m$;

(ii) $(X, \hat{\tau})$ is a T_1 (respectively, Hausdorff) MI-space;

(iii) $(X, \hat{\tau})$ is of second κ-category; thus $(X, \hat{\tau})$ is a second κ-category MI-space;

(iv) if (X, τ) is connected, then so is $(X, \hat{\tau})$.

Proof. Assertions (i), (ii), (iv) follow straightforward from Lemma 2. We only need to prove (iii). Assume that (iii) does not hold. Then there exists a set A, $\operatorname{card} A < \kappa$, such that

$$X = \bigcup_{\alpha \in A} E_\alpha,$$

where each E_α is $\hat{\tau}$-nowhere dense in X (i.e. nowhere dense in $(X, \hat{\tau})$). Therefore $X \setminus X_\alpha$ is $\hat{\tau}$-dense, hence τ-dense in X because $\tau \subset \hat{\tau}$. Since (X, τ) is SI, we have by Lemma 3 that $\operatorname{Int}_\tau(X \setminus E_\alpha)$ is τ-dense in X. It follows that $X \setminus \operatorname{Int}_\tau(X \setminus E_\alpha)$ is τ-closed and τ-nowhere dense in X. Since $E_\alpha \subset X \setminus \operatorname{Int}_\tau(X \setminus E_\alpha)$, we conclude that every E_α is τ-nowhere dense in X; a contradiction because (X, τ) is of the second κ-category. □

Lemma 4. *(*[3]*, Theorem* 33*).* If X is an MI–space and $E \subset X$, then $\operatorname{Int} E = \emptyset$ if and only if E is closed and discrete (the empty set is considered as discrete).

Next we will prove our second main result replacing second κ-category spaces by κ-Baire spaces.

Theorem 2. *Assume that there exists a dense-in-itself T_1 (or Hausdorff) κ-SIB-space (X,τ). Let $\mathscr{F}_m \in \mathfrak{F}(X,\tau)$ be an ultrafilter and let $\hat{\tau}$ be a topology on X generated by $\tau \cup \mathscr{F}_m$. Then*

(i) $D(X,\hat{\tau}) = \mathscr{F}_m$;

(ii) $(X,\hat{\tau})$ is an MI-space which is T_1 (respectively, Hausdorff);

(iii) $(X,\hat{\tau})$ is a κ-Baire space;

Thus $(X,\hat{\tau})$ is a κ-MIB-space which is T_1 (respectively, Hausdorff);

(iv) moreover, if (X,τ) is connected, then so is $(X,\hat{\tau})$.

Proof. As in Theorem 1, claims (i), (ii), (iv) follow immediately from Lemma 2. It only remains to prove (iii). Assume that (iii) does not hold. Then there exists a nonempty set $G \in \hat{\tau}$ which is of the first κ-category in $(X,\hat{\tau})$.

Let us prove that in this case the set $X \setminus G$ should be dense in $(X,\hat{\tau})$.

Since the family $\{W \cap E :\ W \in \tau \setminus \{\emptyset\},\ E \in \mathscr{F}_m\}$ is a basis of the topology $\hat{\tau}$, it suffices to show that

$$\forall\, E \in \mathscr{F}_m\ \forall\, W \in \tau \setminus \{\emptyset\} :\ E \cap W \cap (X \setminus G) \neq \emptyset. \tag{3}$$

Assume that this does not hold. Then there exist $E_0 \in \mathscr{F}_m$ and $W_0 \in \tau \setminus \{\emptyset\}$ such that $E_0 \cap W_0 \cap (X \setminus G) = \emptyset$. It follows that $E_0 \subset (X \setminus W_0) \cup G$, and therefore $(X \setminus W_0) \cup G \in \mathscr{F}_m$, because $\mathscr{F}_m$ is a filter.

Then we have

$$\forall\, E \in \mathscr{F}_m\ : E \cap ((X \setminus W_0) \cup G) \in \mathscr{F}_m,$$

hence $E \cap ((X \setminus W_0) \cup G) = (E \setminus W_0) \cup (E \cap G)$ is dense in (X,τ) for each $E \in \mathscr{F}_m$. Since $\emptyset \neq W_0$ is τ-open, this yields that $E \cap G$ is τ-dense in W_0 for each $E \in \mathscr{F}_m$. In other words,

$$\forall\, E \in \mathscr{F}_m\ \forall\, V \in \tau \setminus \{\emptyset\},\ V \subset W_0,\ : V \cap (E \cap G) = (V \cap E) \cap G \neq \emptyset. \tag{4}$$

Since $V \cap E \in \hat{\tau} \setminus \{\emptyset\}$, Eq. (4) implies that a $\hat{\tau}$-open set $G \cap W_0$ is $\hat{\tau}$-dense in a τ-open, hence $\hat{\tau}$-open, set W_0. It follows that $W_0 \setminus G$ is $\hat{\tau}$-nowhere dense in a $\hat{\tau}$-open set W_0.

This implies, recalling that G is, by assumption, first κ-category in $(X,\hat{\tau})$, that W_0 is also first κ-category in $(X,\hat{\tau})$ what follows immediately in view of the equality

$$W_0 = (W_0 \setminus G) \cup (W_0 \cap G).$$

Therefore, there exists a set A, $\operatorname{card} A < \kappa$, such that

$$W_0 = \bigcup_{\alpha \in A} T_\alpha, \tag{5}$$

where each T_α is nowhere dense in $(X, \hat{\tau})$. Since $\mathrm{Int}_{\hat{\tau}} T_\alpha = \emptyset$ and $(X, \hat{\tau})$ is MI, we have that every T_α is $\hat{\tau}$-closed and $\hat{\tau}$-discrete (Lemma 4). As $(X, \hat{\tau})$ is dense-in-itself, each $X \setminus T_\alpha$ is dense in $(X, \hat{\tau})$. Since (X, τ) is κ-Baire, a τ-open set W_0 is of the second κ-category in (X, τ), therefore it follows by (5) that there exist $\beta \in A$ and $\Omega \subset W_0$, $\Omega \in \tau \setminus \{\emptyset\}$, such that T_β is τ-dense in Ω.

Since the set $X \setminus T_\beta$ is $\hat{\tau}$-dense in X, it is also τ-dense in X. In particular, $X \setminus T_\beta$ is τ-dense in Ω.

We have

$$\Omega = (\Omega \cap T_\beta) \cup (\Omega \cap (X \setminus T_\beta)),$$

where each of the two terms is τ-dense in Ω.

But this means that a τ-open set Ω is resolvable, which is impossible, because (X, τ) is an SI–space.

Consequently, we have shown that if (3) does not hold, then we get a contradiction. Thus $X \setminus G$ is $\hat{\tau}$-dense in X. But this is again a contradiction because G is nonempty and $\hat{\tau}$-open.

We finally conclude that $(X, \hat{\tau})$ has no nonempty first κ-category open subsets, i.e. $(X, \hat{\tau})$ is κ-Baire, as claimed. □

To complete this section, let us make the following

Remark 1. *In* [9] *it was shown that there is a model of the theory* **ZF** *in which all the subsets of the real line are Lebesgue measurable. Let* $\mathbb{R}_s$ *denote the real line in that model and* τ_d *denote the usual density topology on* $\mathbb{R}_s$.

QUESTION*: is* **ZF** *consistent with the conjunction of the following two statements:*

(a) each subset of $\mathbb{R}$ *is Lebesgue measurable,*

(b) almost each point of any set $E \subset \mathbb{R}$ *is its point of density?*

If the answer is in affirmative, then $(\mathbb{R}_s, \tau_d)$ *is a Baire space which is MI. Indeed, the complement of each* τ_d*-dense set* $E \subset \mathbb{R}_s$ *would be of measure zero, whence* E *is* τ_d*-open in* $\mathbb{R}_s$.

3. Some properties of real functions on Baire SI- and MI-spaces

Recall that if X is a topological space and $\varphi : X \to \mathbb{R}$ a USC (or LSC) function, then the set of points at which φ is discontinuous is of the first category (and F_σ) in X (see, e.g. [8], Theorem 1), and if X is a Baire space, then the complement of that set is dense in X. We also recall that by $\omega(F, x)$ we denote the oscillation of F at $x \in X$ (cf. ()). Since $\omega(F, \cdot)$ may take the value $\infty (:= +\infty)$, we consider $[0, \infty]$ with its standard topology of a one-point compactification of $[0, \infty)$.

Given a mapping $\varphi : X \to Y$ between topological spaces, we denote by $\mathscr{C}(\varphi)$ and $\mathscr{D}(\varphi)$ the sets of continuity and discontinuity points of φ, respectively.

Definition 5. ([1]). *A topological space X is said to be resolvable at a point $x_0 \in X$ if each open neighborhood of x_0 contains a nonempty open subset which is resolvable.*

We will use the following proposition which is the main result of [1].

Lemma 5. ([1], *Theorem* 3). *Let $X = (X, \tau)$ be a topological space. In order that X be resolvable at a point x_0, it is necessary and sufficient that the following condition be satisfied. There exist an open neighborhood G of x_0 and a function $F : G \to \mathbb{R}$ such that $0 < \omega(F, x_0) < \infty$ and $\omega(F, \cdot)$ is quasicontinuous at x_0.*

Theorem 3. *Let $X = (X, \tau)$ be a Baire SI-space. Then for each function $F : X \to \mathbb{R}$ we have*

(a) $\mathscr{C}(F) = \mathscr{C}(\omega(F, \cdot))$.

(b) The F_σ-set $\mathscr{D}(F)$ is nowhere dense.

Proof. The set $E_\infty = \{x \in X : \ \omega(F, x) = \infty\}$ is obviously closed. First we will show that E_∞ is nowhere dense. Indeed, assume that this is not the case. Then there exists an open set U such that $\omega(F, x) = \infty$ for each $x \in U$. It follows that $E_n = \{x \in U : \ F(x) > n\}$ is dense in U for each $n \in \mathbb{N}$. Since U is an SI-subspace of X, we have by Lemma 3 that $IntE_n$ is dense in U. The subspace U is a Baire subspace, this yields $\bigcap_{n=1}^{\infty} E_n \neq \emptyset$. But then it follows that $F(x) = \infty$ at each $x \in \bigcap_{n=1}^{\infty} E_n$, which is clearly impossible. Thus, E_∞ is nowhere dense in X.

To prove (a), first observe that the inclusion $\mathscr{C}(F) \subset \mathscr{C}(\omega(F, \cdot))$ is obvious. The reverse inclusion may be proved as follows. Let $x_0 \in \mathscr{C}(\omega(F, \cdot))$. The case $\omega(F, x_0) = \infty$ is impossible what follows immediately from the fact that E_∞ is nowhere dense. So we have $\omega(F, x_0) < \infty$. We claim that $\omega(F, x_0) = 0$. Indeed, if not, we would get, by Lemma 5, that X is resolvable at x_0, a contradiction because X is SI. Thus $\omega(F, x_0) = 0$, i.e. $x_0 \in \mathscr{C}(F)$. This shows that $\mathscr{C}(\omega(F, \cdot)) \subset \mathscr{C}(F)$ which completes the proof of Claim (a).

Put $E_0 = X \setminus E_\infty$. Since $\omega(F, \cdot)$ is USC and finite on a dense open set E_0 (which is a Baire subspace of X), the set $E_0 \cap \mathscr{C}(\omega(F, \cdot)) = E_0 \cap \mathscr{C}(F)$ is dense in E_0, hence by Lemma 3, has a dense interior, because E_0 is SI. Therefore, $\mathscr{D}(F) = E_\infty \cup (E_0 \setminus \mathscr{C}(F)) = X \setminus \mathscr{C}(F)$ is a nowhere dense subset of X which proves Claim (b). □

Similar proposition holds for MI-spaces. Namely, we have

Theorem 4. *Let $X = (X, \tau)$ be a Baire MI-space. Then for each function $F : X \to \mathbb{R}$ we have*

(a) $\mathscr{C}(F) = \mathscr{C}(\omega(F, \cdot))$.*

(b) $\mathscr{D}(F)$ is a discrete closed set.*

Proof. Since each MI-space is an SI-space, Claim (a*) follows from Claim (a) of Theorem 3. By Claim (b) of Theorem 3, we have Int $\mathscr{D}(F) = \emptyset$, whence by Lemma 4, Claim (b*) follows. □

As a consequence, we obtain the following simple criteria for the existence of ω-primitives on Baire SI- and MI-spaces.

Theorem 5. *(A) Let $X = (X, \tau)$ be a Baire SI-space. Then a USC function $f : X \to [0, \infty)$ has an ω-primitive $F : X \to \mathbb{R}$ if and only if f vanishes on a dense subset of X.*

(B) Let $X = (X, \tau)$ be a Baire MI-space. Then a USC function $f : X \to [0, \infty)$ has an ω-primitive $F : X \to \mathbb{R}$ if and only if f vanishes outside of a closed and discrete subset of X.

In either of the cases (A),(B) one may take $F = f$.

Proof of (A). Assume that F is an ω-primitive for f. Then applying Claim (a) of Theorem 3, we get $\mathscr{C}(F) = \mathscr{C}(\omega(F, \cdot)) = \mathscr{C}(f)$. This implies, in view of Claim (b) of Theorem 3, that $f(x) = \omega(F, x) = 0$ at each point x of the dense set $X \setminus \mathscr{D}(F)$.

Conversely, if a USC function $f : X \to [0, \infty)$ vanishes on a dense set E, then it is easy to see that $\forall x \in X : \omega(f, x) = f(x)$.

Proof of (B). Assume that a USC function $f : X \to [0, \infty)$ has an ω–primitive $F : X \to \mathbb{R}$. By Theorem 4, the set $\mathscr{D}(F)$ of points at which F is discontinuous is closed and discrete. Therefore, $f(x) = \omega(F, x) = 0$ at each $x \in X \setminus \mathscr{D}(F)$.

Conversely, assume that there is a closed and discrete set $E \subset X$ such that a USC function $f : X \to [0, \infty)$ vanishes outside E. Since X is dense in itself and $f \geq 0$ is USC, we easily deduce that the equality $\omega(f, x) = f(x)$ holds for each $x \in X$. In other words, f is an ω-primitive for itself. □

References

[1] S.P. Ponomarev. A criterion for the local resolvability of a space and the ω-problem. *J. Appl. Anal.*, **13**, No. 1, 83–96, 2007.

[2] J. Ewert, S.P. Ponomarev. On the existence of ω–primitives on arbitrary metric spaces. *Math. Slovaca*, **53** (1), 51–57, 2003.

[3] E. Hewitt. A problem of set-theoretic topology. *Duke Math. J.*, **10**, 309–333, 1943.

[4] R.C. Haworth, R.A. McCoy. Baire spaces. *Diss. Math.*, CXLI, Warszawa, PWN, 1977.

[5] K. Kunen, A. Szymanski, F. Tall. Baire irresolvable spaces and ideal theory. *Ann. Math. Sil.*, **14**, 98–107, 1986.

[6] G. Bezhanishvili, R. Mines, P.J. Morandi. Scattered, Hausdorf-reducible, and hereditarily irresolvable spaces. *Topology and Its Applications*, **132**, 291–306, 2003.

[7] R.M. Solovay. A model of set-theory in which every set of reals is Lebesgue measurable. *Ann. Math.*, (2), **92**, 1–56, 1970.

[8] M.K. Fort, Jr. Category theorems. *Fund. Math.*, **42**, 276–288, 1955.

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XV, Częstochowa 2010

AXISYMMETRIC SOLUTIONS TO THE CAUCHY PROBLEM FOR TIME-FRACTIONAL DIFFUSION EQUATION IN A CIRCLE

Yuriy Povstenko[a,b]

[a] *Institute of Mathematics and Computer Science*
Jan Długosz University in Częstochowa
al. Armii Krajowej 13/15, 42-200 Częstochowa, Poland
e-mail: j.povstenko@ajd.czest.pl

[b] *European University of Informatics and Economics in Warsaw (EWSIE)*
ul. Białostocka 22/11, 03-741 Warsaw, Poland

Abstract. The Cauchy problems for time-fractional diffusion equation with delta pulse initial value of a sought-for function is studied in a circle domain in the axisymmetric case under zero Dirichlet and Neumann boundary conditions, respectively. The Caputo fractional derivative is used. The Laplace and finite Hankel integral transforms are employed. The results are illustrated graphically.

1. Introduction

The time-fractional diffusion equation

$$\frac{\partial^\alpha u}{\partial t^\alpha} = a\Delta u, \qquad 0 < \alpha \leq 2, \tag{1}$$

is a mathematical model of a wide range of important physical phenomena in amorphous and porous materials, fractals, disordered media, dielectrics and semiconductors, geophysical and geological processes, medicine and biological systems [1–8].

In Eq. (1), we use the Caputo fractional derivative [9]

$$\frac{\mathrm{d}^\alpha u}{\mathrm{d}t^\alpha} = \frac{1}{\Gamma(n-\alpha)} \int_0^t (t-\tau)^{n-\alpha-1} \frac{\mathrm{d}^n u(\tau)}{\mathrm{d}\tau^n}\, \mathrm{d}\tau, \qquad n-1 < \alpha < n, \tag{2}$$

where $\Gamma(x)$ is the gamma function. The Laplace transform rule for the Caputo derivative has the following form:

$$\mathcal{L}\left\{\frac{\mathrm{d}^\alpha u(t)}{\mathrm{d}t^\alpha}\right\} = s^\alpha \mathcal{L}\left\{u(t)\right\} - \sum_{k=0}^{n-1} u^{(k)}(0^+) s^{\alpha-1-k}, \qquad n-1 < \alpha < n, \quad (3)$$

with s being the transform variable.

Several problems for time-fractional diffusion equation in a cylinder were considered in [10–14]. In this paper we investigate the Cauchy problems with delta function initial value of a sought-for function in a circle domain under zero Dirichlet and Neumann boundary conditions, respectively, and compare the obtained results with the corresponding solution in an infinite domain.

2. The Cauchy problem in an infinite domain

In order to gain a better insight of the considered problem in a circle, we recall the corresponding result for the infinite domain [15]. Let us study the Cauchy problem for time-fractional diffusion equation under delta-function initial condition for a sought-for function:

$$\frac{\partial^\alpha u}{\partial t^\alpha} = a\left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r}\right), \qquad 0 < t < \infty, \quad 0 \le r < \infty, \quad (4)$$

$$t = 0: \quad u = \frac{p}{2\pi r}\,\delta_+(r), \quad 0 < \alpha \le 2, \quad (5)$$

$$t = 0: \quad \frac{\partial u}{\partial t} = 0, \qquad 1 < \alpha \le 2. \quad (6)$$

As usually, we impose the zero condition at infinity:

$$\lim_{r\to\infty} u(r,t) = 0. \quad (7)$$

Using the Laplace transform with respect to time t and the Hankel transform with respect to the spatial coordinate r, we obtain

$$u^* = \frac{p}{2\pi}\,\frac{s^{\alpha-1}}{s^\alpha + a\xi^2}, \quad (8)$$

where the asterisk denotes the transforms.

Inversion of the Laplace transform is carried out in terms of the Mittag-Leffler functions

$$E_\alpha(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n + 1)}, \qquad \alpha > 0, \;\; z \in C, \quad (9)$$

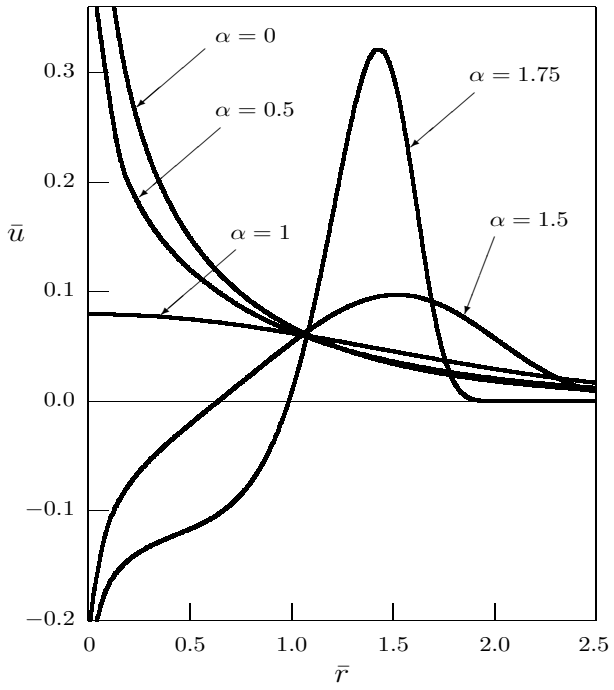


Fig. 1. Dependence of solution on the similarity variable
(the Cauchy problem with the delta pulse initial condition)

due to the following formula [9]

$$\mathcal{L}^{-1}\left\{\frac{s^{\alpha-1}}{s^{\alpha}+a\xi^2}\right\} = E_{\alpha}(-a\xi^2 t^{\alpha}). \tag{10}$$

Thus, we get

$$u = \frac{p}{2\pi}\int_0^{\infty} E_{\alpha}(-a\xi^2 t^{\alpha})\, J_0(r\xi)\, \xi\, d\xi. \tag{11}$$

The similarity variable $\bar{r}$, new integration variable η and nondimensional solution $\bar{u}$ are defined as

$$\bar{r} = \frac{r}{\sqrt{a}t^{\alpha/2}}, \qquad \eta = \sqrt{a}t^{\alpha/2}\xi, \qquad \bar{u} = \frac{at^{\alpha}}{p}\, u. \tag{12}$$

Hence,

$$\bar{u} = \frac{1}{2\pi}\int_0^{\infty} E_{\alpha}(-\eta^2)\, J_0(\bar{r}\eta)\, \eta\, d\eta. \tag{13}$$

The behavior of the solution at the origin was analyzed in [15], where it was shown that only the fundamental solution to the classical diffusion equation

($\alpha = 1$) has no singularity at the origin. For $0 \leq \alpha < 1$ and $1 < \alpha < 2$ the solution has the logarithmic singularity at the origin:

$$\bar{u} \sim -\frac{1}{2\pi\Gamma(1-\alpha)} \ln \bar{r}. \tag{14}$$

Dependence of nondimensional solution $\bar{u}$ on nondimensional distance $\bar{r}$ is shown in Fig. 1.

3. The Cauchy problem in a circle with zero Dirichlet boundary condition

Consider the following initial-boundary value problem for time-fractional diffusion equation:

$$\frac{\partial^\alpha u}{\partial t^\alpha} = a\left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r}\right), \quad 0 < t < \infty, \quad 0 \leq r < R, \tag{15}$$

$$t = 0: \quad u = \frac{p}{2\pi r}\,\delta_+(r), \quad 0 < \alpha \leq 2, \tag{16}$$

$$t = 0: \quad \frac{\partial u}{\partial t} = 0, \qquad 1 < \alpha \leq 2. \tag{17}$$

$$r = R: \quad u = 0. \tag{18}$$

The finite Hankel transforms are used in cylindrical coordinates in the domain $0 \leq r \leq R$. The form of the finite Hankel transform depends on the type of boundary conditions at $r = R$. We restrict ourselves to the finite Hankel transform of the zeroth order. For Dirichlet boundary conditions with the given boundary value of a function at $r = R$ we have [16]

$$\mathcal{H}^{(\mathrm{D})}\{f(r)\} = f^*(\xi_n) = \int_0^R f(r)\, J_0(\xi_n r)\, r\, \mathrm{d}r \tag{19}$$

with the inverse transform

$$\mathcal{H}^{-1(\mathrm{D})}\{f^*(\xi_n)\} = f(r) = \frac{2}{R^2}\sum_{n=1}^{\infty} f^*(\xi_n)\frac{J_0(\xi_n r)}{J_1^2(\xi_n R)}, \tag{20}$$

where ξ_n are positive zeros of the transcendental equation

$$J_0(R\xi_n) = 0. \tag{21}$$

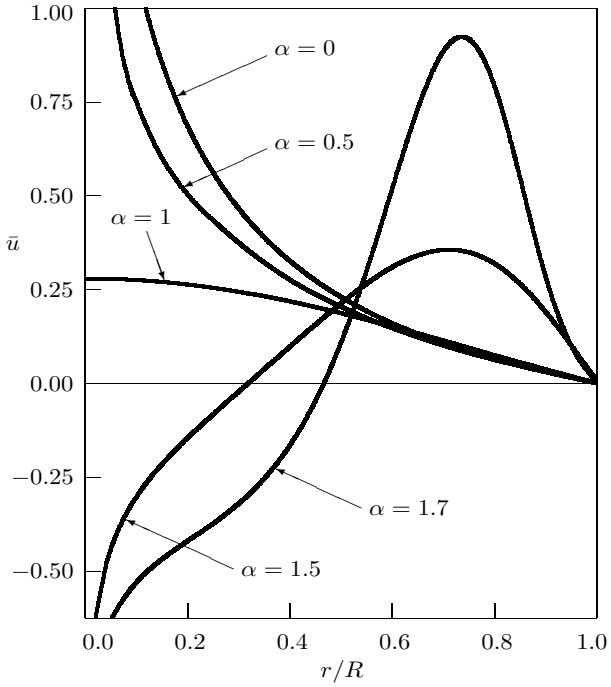


Fig. 2. Dependence of solution on distance
(the zero Dirichlet boundary condition; $\kappa = 0.5$).

The following formula plays important role in applications of the finite Hankel transform:

$$\mathcal{H}^{(\mathrm{D})}\left\{\frac{\mathrm{d}^2 f(r)}{\mathrm{d}r^2} + \frac{1}{r}\frac{\mathrm{d}f(r)}{\mathrm{d}r}\right\} = -\xi_n^2 f^*(\xi_n) + R\xi_n J_1(\xi_n R) f(R). \tag{22}$$

The integral transform technique allows us to get the solution in the transform domain:

$$u^* = \frac{p}{2\pi}\,\frac{s^{\alpha-1}}{s^\alpha + a\xi_n^2}, \tag{23}$$

and after inversion we arrive at the series representation of the solution:

$$u = \frac{p}{\pi R^2}\sum_{n=1}^{\infty} E_\alpha(-a\xi^2 t^\alpha)\,\frac{J_0(r\xi_n)}{J_1^2(R\xi_n)}. \tag{24}$$

Introdusing nondimensional quantities

$$\eta_n = R\xi_n, \quad \kappa = \frac{\sqrt{a}t^{\alpha/2}}{R}, \quad \bar{r} = \frac{r}{R}, \quad \bar{u} = \frac{R^2}{p}\,u, \tag{25}$$

we have

$$\bar{u} = \frac{1}{\pi} \sum_{n=1}^{\infty} E_\alpha(-\kappa^2 \eta_n^2) \, \frac{J_0(\bar{r}\eta_n)}{J_1^2(\eta_n)}. \tag{26}$$

Figure 2 shows the dependence of the solution (26) on distance for $\kappa = 0.5$.

4. The Cauchy problem in a circle with zero Neumann boundary condition

Now we study the time-fractional diffusion equation in a circle under delta pulse initial condition and zero Neumann boundary condition:

$$\frac{\partial^\alpha u}{\partial t^\alpha} = a \left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} \right), \quad 0 < t < \infty, \quad 0 \le r < R, \tag{27}$$

$$t = 0: \quad u = \frac{p}{2\pi r} \, \delta_+(r), \quad 0 < \alpha \le 2, \tag{28}$$

$$t = 0: \quad \frac{\partial u}{\partial t} = 0, \quad 1 < \alpha \le 2, \tag{29}$$

$$r = R: \quad \frac{\partial u}{\partial r} = 0. \tag{30}$$

For the Neumann boundary condition with the given value of normal derivative of a function, the corresponding finite Hankel transform is defined as [16]:

$$\mathcal{H}^{(\mathrm{N})}\{f(r)\} = f^*(\xi_n) = \int_0^R r f(r) J_0(r\xi_n) \, \mathrm{d}r, \tag{31}$$

having the inverse

$$\mathcal{H}^{-1(\mathrm{N})}\{f^*(\xi_n)\} = f(r) = \frac{2}{R^2} \sum_{n=0}^{\infty} f^*(\xi_n) \frac{J_0(r\xi_n)}{[J_0(R\xi_n)]^2}, \tag{32}$$

where ξ_n are nonnegative roots of the transcendental equation

$$J_1(R\xi_n) = 0. \tag{33}$$

To obtain the correct results, it should be emphasized that Eq. (33) also has the root $\xi_0 = 0$ which should be taken into consideration.

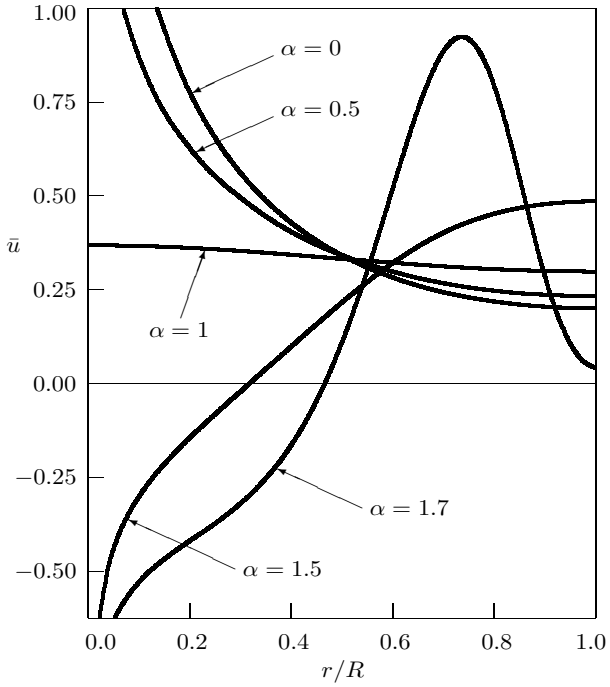


Fig. 3. Dependence of solution on distance
(the zero Neumann boundary condition; $\kappa = 0.5$).

The following formula explains importance of the finite Hankel transform of such a type for Neumann boundary value problems:

$$\mathcal{H}^{(\mathrm{N})}\left\{\frac{\mathrm{d}^2 f}{\mathrm{d}r^2} + \frac{1}{r}\frac{\mathrm{d}f}{\mathrm{d}r}\right\} = -\xi_n^2 f^*(\xi_n) + RJ_0(R\xi_n)\left(\frac{\mathrm{d}f}{\mathrm{d}r}\right)_{r=R}. \tag{34}$$

Thus, we obtain

$$u^* = \frac{p}{2\pi}\,\frac{s^{\alpha-1}}{s^\alpha + a\xi_n^2} \tag{35}$$

and

$$u = \frac{p}{\pi R^2}\sum_{n=0}^{\infty} E_\alpha(-a\xi^2 t^\alpha)\,\frac{J_0(r\xi_n)}{J_0^2(R\xi_n)} \tag{36}$$

or in terms of nondimensional quantities (25)

$$\bar{u} = \frac{1}{\pi}\sum_{n=0}^{\infty} E_\alpha(-\kappa^2\eta_n^2)\,\frac{J_0(\bar{r}\eta_n)}{J_0^2(\eta_n)}. \tag{37}$$

Dependence of the solution (37) on distance for $\kappa = 0.5$ is depicted in Fig. 3.

5. Concluding remarks

The results given by Eqs. (26) and (37) and displayed in Figures 2 and 3 are the primary results of this paper. The parameter κ describes nondimensional time and in the case of the wave equation ($\alpha = 2$) the values $0 < \kappa < 1$ and $\kappa = 1$ correspond to two characteristic cases: the wave front does not yet arrive at the boundary, and the wave front arrives at the boundary. For $0 \leq \alpha < 1$ and $1 < \alpha < 2$ in the case $\kappa = 0.5$ the solution does not "feel" the type of the boundary condition: the curves in Figs. 2 and 3 are very similar and do not differ essentially from the corresponding curves obtained for unbounded domain (see Fig. 1), including the logarithmic singularity at the origin. But for $\kappa = 1$ the situation changes substantially.

References

[1] A. Pękalski, K. Sznajd-Weron (Eds.) *Anomalous Diffusion: From Basics to Applications.* Springer, Berlin, 1999.

[2] R. Hilfer (Ed.) *Applications of Fractional Calculus in Physics.* World Scientific, Singapore, 2000.

[3] R. Metzler, J. Klafter. The random walk's guide to anomalous diffusion: A fractional dynamics approach. *Phys. Rep.*, **339**, 1–77, 2000.

[4] G.M. Zaslavsky. Chaos, fractional kinetics, and anomalous transport. *Phys. Rep.*, **371**, 461–580, 2002.

[5] B.J. West, M. Bologna, P. Grigolini. *Physics of Fractal Operators.* Springer, New York, 2003.

[6] R. Metzler, J. Klafter. The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics. *J. Phys. A: Math. Gen.*, **37**, R161–R208, 2004.

[7] R.L. Magin. *Fractional Calculus in Bioengineering.* Begell House Publishers, Connecticut, 2006.

[8] V.V. Uchaikin. *Method of Fractional Derivatives.* Artishock, Ulyanovsk, 2008. (In Russian).

[9] A.A. Kilbas, H.M. Srivastava, J.J. Trujillo. *Theory and Applications of Fractional Differential Equations.* Elsevier, Amsterdam, 2006.

[10] Y.Z. Povstenko. Fractional radial diffusion in a cylinder. *J. Mol. Liq.*, **137**, 46–50, 2008.

[11] N. Özdemir, D.Karadeniz. Fractional diffusion-wave problem in cylindrical coordinates. *Phys. Lett. A*, **372**, 5968–5972, 2008.

[12] N. Özdemir, D. Karadeniz, B.B. Iskender. Fractional optimal control problem of a distributed system in cylindrical coordinates. *Phys. Lett. A*, **373**, 221–226, 2009.

[13] E.K. Lenzi, L.R. da Silva, A.T. Silva, L.R. Evangelista, M.K. Lenzi. Some results for a fractional diffusion equation with radial symmetry in a confined region. *Physica A*, **388**, 806–810, 2009.

[14] H. Qi, J. Liu. Time-fractional radial diffusion in hollow geometries. *Meccanica*, **45**, 577–583, 2010.

[15] Y. Povstenko. Analysis of fundamental solutions to fractional diffusion-wave equation in polar coordinates. *Scientific Issues, Jan Długosz University of Częstochowa, Mathematics*, **XIV**, 97–104, 2009.

[16] I.N. Sneddon. *The Use of Integral Transforms.* McGraw-Hill, New York, 1972.

PART II
COMPUTER SCIENCE

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XV, Częstochowa 2010

ON SOME SPECIFICATION LANGUAGES OF CRYPTOGRAPHIC PROTOCOLS

Paweł Dudek, Mirosław Kurkowski

Institute of Computer and Information Sciences
Częstochowa University of Technology
ul. Dąbrowskiego 73, 42-200 Częstochowa, Poland
e-mail: pdudek@icis.pcz.pl, mkurkowski@icis.pcz.pl

Abstract. A key element of the security systems in computer networks are cryptographic protocols (CP). These protocols are concurrent algorithms used to provide relevant system security goals. Their main purpose is, for example, a mutual authentication (identification) of communicating parties (users, servers), distribution of new keys and session encryption. Literature indicates numerous errors in protocol constructions. Thus, there is a need to create methods for CP specification and verification.

In this paper, we investigate a problem of CP specification. The paper discusses the so-called Common Language – the simplest language of CP specification and HLPSL – a specification language used in the European verification project Avispa. Finally, we introduce PTL – the new language developed for CP specification which allows fully automatic verification.

1. Introduction

It is well known that today each IT system and computer network must meet certain security properties [6]. CP are now commonly used in various applications (banking, emails, encrypted Web pages, instant messaging networks, etc.) for achieving security goals. They are also widely used as essential components of larger systems such as communication protocols for wider application. Good examples are the systems of Kerberos, SSL and Zfone. A pioneering role in the area of CP has the paper published in 1978 by Needham and Schroeder [2]. In their work the authors presented main ideas of applying cryptographic techniques in order to solve problems related to the

authentication of communicating parties in communication networks. Suggested layouts of authentication protocols can use symmetric and asymmetric cryptography.

CP are concurrent algorithms, designed to attain certain specific objectives during the transfer, including the transactions carried out electronically. In general, these protocols are algorithms whose implementations are performed in a concurrent way and may be used for cooperating computers, computer networks or simply across multiple CPUs. This is a significant difference between them and ordinary sequential algorithms. CP can also specify concurrent processes as communicating sequential processes with each other from time to time through the exchange of data (the parameters) or the use of common resources. Cryptographic protocols are those concurrent processes, which work using cryptographic algorithms.

A specification of any cryptographic protocol has to contain:

- the number of parties involved in the protocol,
- the nature of the participation of the parties,
- the goal of the protocol,
- actions comprised in the implementation of the protocol.

Basic security goals which CP need to ensure are the following:

- mutual authentication (confirmation of identity) of communicating parties,
- confidentiality of transmitted information,
- integrity of transmitted data,
- distribution of session key.

Actions performed during the execution of the protocol can be divided into internal and external ones. External actions are those which rely on the mutual exchange of transmitted information. Description of those actions will specify sources of any sent message (senders), recipients of sent messages and their contents. It must also indicate, respectively, which part of the sent letter has to be encrypted and how. Internal actions are all the other actions that each party must perform on its own during the execution of the protocol. As examples, one can give generating new, confidential information, encrypting and deciphering cryptograms, comparing data or performing mathematical operations on locally held data.

Applying cryptographic protocols in order to ensure adequate security purposes in computer systems requires special attention with regard to the correctness of their executions. Incorrect work of protocols can lead to different

sorts of losses of users resources [4]. Cryptographic protocols are usually short and not too complicated in their structure, so often entirely informal arguments are used to justify that they operate properly and to all system users that the protocol actually does what it is expected to do [2].

In most cases, however, it is difficult to imagine all possible executions of these complex systems. This becomes especially difficult when dealing with programs that are executed concurrently on many computers, where the partial results of these performances may affect the implementation of the next instruction. For these reasons, the method of verifying the correctness of software systems is constantly an extensively developed area of computer science.

Basically, we can distinguish two main groups of verification methods:

1. Testing of real or virtual systems (simulations).
2. Formal modeling and verification.

In the first case, the verification process simply consists in testing the systems already implemented or simulating their performances by computers (eg. virtual machines). After carrying out several such tests or simulations, unfortunately, we can only say that so far the implementation of all programs works properly.

The second direction of research, namely formal modeling and verification, involves creating special mathematical structures which model processes taking place during protocol executions. It is therefore, in some sense, the creation of a new, specific types of simulation. However, as numerous examples show, this type of modeling can sometimes prove formally that certain undesirable behavior of the system will never occur.

Creating mathematical structures simulating the implementation of cryptographic protocols is not an easy process. This work, however, requires to use a specially constructed languages for protocols specification. In [2] a simple language for the specification of protocols has been applied, known simply as Common Language (CL) [1, 6]. As an example, we show below protocol specification using CL and some information about it.

2. Common Language – CL

Common Language has never been formalized. However, the grammar of the basic version is not too complicated. The protocol is described as a sequence of steps, specifying the sender of the message, recipient and content of the sent letter [1].

Each step is specified as follows:

$$A \rightarrow B \ : \ M,$$

where A is, of course, the sender of the message, B is the recipient and M is the message.

The grammar of messages is the following:

$$M \ : \ A \mid T \mid K \mid N \mid L \mid M, M \mid < M >_K,$$

where A belongs to a set of users, K to a set of cryptographic keys, T to the set of timestamps, L is a life time of T. The keys used in the specification, of course, may be symmetric or asymmetric. In the first case, we denote by K_{AB} the key, where A and B are their owners; in the second case, the symbol K_A denotes the public key of A and K_A^{-1} its private key. M, M is simply a concatenation of messages, and by writing $< M >_K$ we understand the ciphertext M encrypted with the key K.

Here, as an example, we show a specification for some version of Kerberos Protocol using the Common Language. The basic version of this protocol is as follows: we have two parties A and B, which share the server S with different secret keys. The main goal of this protocol is to generate by A a session key in order to conduct communication with B.

Protocol specification:

1. $A \rightarrow S \ : \ A, B,$
2. $S \rightarrow A \ : < T, L, K, B >_{K_{AS}}, < T, L, K, A >_{K_{BS}},$
3. $A \rightarrow B \ : < A, T >_K, < T, L, K, A >_{K_{BS}},$
4. $B \rightarrow A \ : < T >_K.$

In the first step of the protocol, the user A sends to the server S a message consisting of its identifier and the name of B. In this way, S possesses information with whom A wants to communicate. In the second step, the server generates two messages with a timestamp T, the ticket duration L and a newly generated, random session key K. S encrypts all of them using a secret key shared with B. Then it gets a timestamp, the duration L and the identifier B, and encrypts everything using secret key shared with A. Next, S sends two encrypted messages to A. In the third step, A generates a message containing its identifier and timestamp, encrypts them using the session key K newly obtained from S and sends it to B. A also sends to B a message encrypted by the server using a common key for B and S. Then B possesses the key K and creates a message consisting of the timestamp T, encrypts it using K and sends it to A.

Executing a protocol assumes the existence of an ideal clock allocating time in compliance with clocks of all users of the server. This is achieved by synchronizing every few minutes to a secure server clock time. The key server S needs to remember all the keys that it shares with users. However, the session key is created for the purpose of communication between A and B, then the server forgets about the result.

Obviously, as one can see from the above example, CL is very simple and it is probably difficult to imagine a simpler protocol specification language. However, it is important to note that it requires additional information about, for example, the description of internal actions during the protocol execution, including generating new elements such as keys, nonces (pseudo-randoms numbers generated for a single session) or timestamps. There is also no information about how users compose sent messages. That is the reason why CL cannot be used in fully automatic verification.

3. HLPSL Language

Currently, the world's most recognizable system of formal verification of cryptographic protocols is the AVISPA system (Automated Validation of Internet Security Protocols and Applications) [7, 10]. This system was created through cooperation of several institutions: Universities of Genova, Zurich, Nancy and subsidiaries of Siemens in Munich. For this project a special role-based, high-level language HLPSL for CP specification (High Level Protocol Specification Language) was created.

In HLPSL each participant has a defined primary role (basic role), which is described by various parameters related to the behavior of participants during the execution protocol. These roles define how users can transfer their information during the executions of the protocol. Data included in those roles determine the information, which a participant can use initially, and the initial state of the knowledge. Additionally, roles describe how the users knowledge might change during the execution of the protocol. The specification given in basic roles can be used later by one or more users who can play a particular role in the protocol execution. Then, to create the composed roles we describe how the individual members communicate among themselves by means of repeated basic roles. In this way, we obtain a specification schema for data exchange during the whole protocol execution. Roles are independent processes, which have a specific name, replaced by the value of initialization parameters, also contain local declarations. Actions of simple roles are specified in order to describe transitions in the form of a change in the role depending on the events occurred, while the complex roles determine the way in which pre-defined roles are combined.

The HLPSL specification also defines additional parameters of verification. Additionally, in a batch file there will be determined security properties which are to be examined and the size of the protocol performances in the needed searching space. The declaration and definition of the objectives which we want to achieve during the verification takes place in an another special section of specification.

HLPSL allows testing of the following security properties:

- maintaining the confidentiality of the information,
- strong user authentication on the basis of the message,
- weak user authentication based on a certain message.

Fixed data or variable used in specification must have assigned a unique type. The list of examples of variables is the following:

- *agent* – for users identifiers, for the intruder the letter *i* is reserved,
- *public_key* – for public keys of agents. Given a public key *pk* (resp. private), its reversed private (resp. public) key is obtained through the structures *inv(pk)*,
- *symmetric_key* – for keys used in symmetric encryption,
- *nat* – for the scope of variables of this type the natural numbers are used. *Nat* type is usually used to describe states,
- *protocol_id* – for identifiers used in the studied properties,
- *message* – for representing any message,
- *text* – for nonces.

Correct messages are defined as the submission of the concatenation operation '.' and/or encryption '_' (message_key) of basic data types. There is no difference between the descriptions of symmetric and asymmetric encryptions. Assuming that we have a type of agent, the agent A, the nonce N_a and the symmetric key K, the following messages are correct:

1. N_a – nonce N_a is a message,
2. $A.N_a$ – the message containing the identifier of agent A with a value N_a,
3. $\{A.N_a\}_K$ – the proper message encrypted with the key K.

A channel is a variable that connects communicating parties and exchanges messages between them. HLPSL's channels contain intruder acting in that channel. The model available in HLPSL is the well known Dolev-Yao model [3] (denoted by *dy*) in which the attacker is a network of canals.

The four predefined goal predicates listed above contain the following information:

- *secret(E,id,S)*: declares the information E as a secret shared by the agents from a set S; this secret will be identified by the constant *id* in the goal section;
- *witness(A,B,id,E)*: for a (weak) authentication property of A by B on E, declares that an agent A is witness for the information E; this goal will be identified by the constant *id* in the goal section;
- *request(B,A,id,E)*: for a strong authentication property of A by B on E, declares that an agent B requests checking the value E; this goal will be identified by the constant *id* in the goal section;
- *wrequest(B,A,id,E)*: similar to *request* property, but in this case for a weak authentication property.

Summing up, the language HLPSL is a very complex language which allows the full specification of cryptographic protocols. It is clear, however, that it has been specially designed deliberately to be used by a specific tool, namely the verification system Avispa. That is why one can reflect on its versatility. It is obvious that if we would like to apply the specification of the protocol in HLPSL in another tool in the study and application, we need appropriate special translators.

4. VerICS system and PTL language

VerICS [5, 11] is an original tool for automatic or semi-automatic verification of concurrent systems. The system allows verification of various properties of systems containing the temporal aspects. One module of VerICS is solely devoted for the CP verification. The results obtained by the VerICS team so far are competitive to the other results obtained in Europe and worldwide [7, 8, 9, 10, 12]. In the case of CP verification, a special mathematical model of CP executions has been developed. This model allows testing various executions of CP. The formalism has been designed so as to be able to identify accurately the correct sequences of steps protocol performances that make up executions performances.

For this project, a simple specification language called ProTocol Language (PTL) has been proposed. In this approach, the protocol is defined as a sequence of steps, and each of them is defined as an ordered pair of the form: (α_1, α_2). The component α_1 defines external actions of the protocol (messaging), while the component α_2 defines internal ones.

Both components contain basic and complete information about the specified protocol. More precisely:

$$\alpha_1 = (P, Q, M), \alpha_2 = (t, X, G, \tau),$$

where P is the step initiator, Q is the owner and M is the sent message. So far, there are no differences between this specification and the specification in the CL language. In this approach, we have yet more information: t is a variable indicating time when the step's execution starts, X is a set of information needed to compose a message, G is a set of sensitive information generated for a given step, and τ is a time constraint ensuring that each step can be performed. This specification allows precise determination of not only the external actions of the protocol but also internal ones.

The message grammar is the same as the corresponding grammar in CL:

$$M \ : \ A \ | \ T \ | \ K \ | \ N \ | \ L \ | \ M, M \ | < M >_K \ .$$

In addition, we specify time constraints according to the following grammar:

$$\tau \ : \ false \ | \ true \ | \ t - T \leq L \ | \ \tau_1 \wedge \tau_2.$$

As an example of a full protocol specification in the PTL language, we give now a formal description of the Kerberos Protocol mentioned above.

Protocol specification in PTL is as follows:

1. The first step $(\alpha_1, \ \alpha_2)$, where

$$\alpha_1 = (A; S; A, B), \ \alpha_2 = (t_1, \{A, B\}, \emptyset, true).$$

2. The second step $(\beta_1, \ \beta_2)$, where

$$\beta_1 = (S; A; < T, L, K, B >_{K_{AS}}, < T, L, K, A >_{K_{BS}}),$$

$$\beta_2 = (t_2, \{T, L, K, A, B, K_{AS}, K_{BS}\}, \{T, K\}, t_2 - T \leq L).$$

3. The third step $(\gamma_1, \ \gamma_2)$, where

$$\gamma_1 = (A; B; < A, T >_K, < T, L, K, A >_{K_{BS}}),$$

$$\gamma_2 = (t3, \{T, K, A, < T, L, K, A >_{K_{BS}}\}, \emptyset, t_3 - T \leq L).$$

4. The fourth step $(\delta_1,\ \delta_2)$, where

$$\delta_1 = (B; A; < T >_K),\ \delta_2 = (t_4, T, K, \emptyset, t_4 - T \leq L).$$

Note that this specification gives precise information on both external and internal actions of the protocol. It also determines precisely the time conditions which users need to fulfill.

From the technical point of view, a PTL specification file contains only information needed for further steps of verification process. This file consists of two main parts. In the first one, we have basic information about numbers of considered users and protocol steps. The second one contains specification of all protocol steps in the way mentioned before. In next lines, we have the description of pairs specifying steps of the protocol.

5. Conclusion

Effective methods of specification and verification of cryptographic protocols are an important problem of applied cryptography. In this paper, we have discussed a few basic languages developed for CP specification. We have presented the Common Language – the simplest language for protocol specification, HLPSL – the language used in the European project Avispa, and the PTL language. Investigations in this area are still in progress. The next steps consists in expanding the expressive power of the PTL language to describe a larger class of protocols and delays in the network occurring during the transmission of information.

References

[1] *Security Protocols Open Repository* (SPORE): http://www.lsv.ens-cachan.fr/Software/spore/

[2] R. Needham, M. Schroeder. Using encryption for authentication in large networks of computers. *Comm. ACM*, **21** (12), 993–999, 1978.

[3] D. Dolev, A. Yao. On the security of public key protocols. *IEEE Trans. Information Theory*, **29** (2), 198–208, 1983.

[4] G. Lowe. Breaking and fixing the Needham-Schroeder public-key protocol using FDR. In: *Proc. TACAS*, pp. 147-166, Springer, 1996.

[5] P. Dembiński, A. Janowska, P. Janowski, W. Penczek, A. Półrola, M. Szreter, B. Woźna, A. Zbrzezny. VerICS: A tool for verifying timed automata and estelle specifications. In: *Proc. 9th Int. Conf. TACAS'03*, vol. 2619 of LNCS, pp. 278–283, Springer, 2003.

[6] A. Menezes, P. van Oorschot, S. Vanstone. *Kryptografia stosowana*, WNT, 2005.

[7] A. Armando, D. Basin, Y. Boichut, Y. Chevalier, L. Compagna, J. Cuellar, P. Hankes Drielsma, P.C. Heam, O. Kouchnarenko, J. Mantovani, S. Modersheim, D. von Oheimband M. Rusinowitch, J. Santiago, M. Turuani, L. Vigano, L. Vigneron. The AVISPA tool for the automated validation of internet security protocols and applications. In: *Proc. 17th Int. Conf. Computer Aided Verification (CAV'05)*, vol. 3576 of LNCS, pp. 281–285, Springer, 2005.

[8] M. Benerecetti, N. Cuomo, A. Peron. TPMC: A model checker for time-sensitive security protocols. In: *Proc. 2007 High Performance Computing and Simulation Conf. (HPCS 2007)*, pp. 742-749, Prague, 2007.

[9] M. Kurkowski, W. Penczek. Verifying security protocols modelled by networks of automata. *Fund. Informaticae*, **79** (3-4), 453–471, 2007.

[10] A. Armando, L. Compagna. Sat-based model-checking for security protocols analysis. *Int. J. Information Security*, **7** (1), 3–32, 2008.

[11] M. Kacprzak, W. Nabiałek, A. Niewiadomski, W. Penczek, A. Półrola, M. Szreter, A. Zbrzezny. Verics 2008 – a model checker for high-level languages. *Artificial Intelligence Studies* **5** (28), 131–140, 2008.

[12] M. Kurkowski, W. Penczek. Verifying timed security protocols via translation to timed automata. *Fund. Informaticae*, **93** (1-3), 245–259, 2009.

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XV, Częstochowa 2010

MODULAR NUMBER SYSTEMS IN THE COMPLEX PLANE

Mikhail Selyaninov

Institute of Technical Education and Safety
Jan Długosz University in Częstochowa
al. Armii Krajowej 13/15, 42-200 Częstochowa, Poland
e-mail: m.selianinov@ajd.czest.pl

Abstract. In the present paper, we consider methods of constructing modular number systems (MNS), named also as residue number systems, in the complex plane. The structure of complete sets of residues (CSR) with respect to complex modulo is investigated. For its creation, the effective constructive rule realizing isomorphism of the given CSR and an adequate ring of real integer residues is proposed.

1. Introduction

Procedures over the complex data in modern computer applications to digital signal processing, numerical methods, theoretical mechanics, physics, other sciences are of fundamental importance. Typical representatives of mentioned procedures are, for example, discrete Fourier transforms, spectral analysis, convolution and correlation of complex sequences, algorithms of linear algebra and differential equations etc. In view of exclusive complexity of this type of applied procedures, studies in the field of modular technique of high-speed parallel computations in the complex plane are among the most priority directions of modern computer science and its applications [1–4].

2. Some theoretical foundations

Let us consider the set of integer complex numbers (ICN) often named also as Gaussian: $\Gamma = \{X + iY \mid X,\ Y \in \mathbf{Z};\ i^2 = -1\}$. The set Γ represents a commutative ring without zero divisor and with four dividers of unity: 1; -1; i and $-i$ [1, 2].

Definition 1. The norm of ICN $U = X + iY$ is the square of its magnitude: $\|U\| = X^2 + Y^2$.

Definition 2. The Gaussians differing by multipliers which are the unit dividers are said to be associated.

Definition 3. The unit dividers of ring Γ and units associated with ICN are named trivial dividers.

Definition 4. The Gaussian having nontrivial zero divisor is called composite, otherwise it is named a Gauss prime number (GPN).

The necessary and sufficient condition of simplicity of the ICN $U = X + iY$ is primality of integer real number (IRN) U. This implies that the norm of GPN is a prime IRN or square of prime IRN. In the first case, the real and imaginary parts of the GPN are distinct from zero, and in the second case, the GPN coincides with the prime IRN accurate within unit dividers.

In the ring Γ the Euclidian lemma is formulated as follows [1].

Lemma. For any ICN A and m in Γ there are some q and r such that

$$A = q\ m; \quad \|r\| < \|m\|. \tag{1}$$

It should be noted that, unlike a real case, the condition $\|r\| < \|m\|$ does not ensure the uniqueness of an incomplete quotient q and a residual r in formula (1).

Definition 5. The common divisor d of the ICN $A_1, A_2, \ldots, A_k$ $(k > 1)$ dividing by another common divisor is named the greatest common divisor (GCD) and is designated as $d = (A_1, A_2, \ldots, A_k)$. For any collection of the ICN $A_1, A_2, \ldots, A_k$ a GCD exists accurate within unit dividers.

Definition 6. If $(A_1, A_2, \ldots, A_k) = 1$, then the ICN $A_1, A_2, \ldots, A_k$ are named coprime.

It is obvious that if $A = X + iY$ and $(X, Y) = 1$, then the conjugate ICN A and $\bar{A} = X - iY$ are coprime: $(A,\ \bar{A}) = 1$. For arbitrary ICN A and GPN p the following statement is valid: $(A,\ p)$ does not divide A. Also in the case when ICN $m_1, m_2, \ldots, m_k$ are pairwise prime, then the least common multiple is $[m_1, m_2 ..., m_k] = M_k = \prod_{i=1}^{k} m_i$.

Following the offered technique of constructing a modular number systems (MNS) [5–7] first of all we study the structure of adjanced classes of factor ring $\Gamma/(m)$, where (m) is a principal ideal generated by some Gaussian

$m = m' + im''$ $(m',\ m'' \in Z)$. We also will study the problem of choice of the complete set of residues (CSR) modulo m.

The given CSR (as distinct from CSR $|\cdot|_m$ in the real case) will be designated as $\langle\cdot|_m$, thus for the set member of CSR being a result of modulo operation over ICN $A \in \Gamma$ a denotation $\langle\ A|_m$ is used. Specifically, in the case whent A vary over Γ, a set of all various residuals r satisfying (1) can be selected as a ring $\langle\cdot|_m$.

Theorem 1. Let $m = m' + im''$ be an arbitrary module from Γ. Two ICN $A = A' + iA''$ and $B = B' + iB''$ belong to the same residue class of factor ring $\Gamma/(m)$ if and only if the same components of pairs ICN $(m'\ A' + m''\ A'';\ m'\ A'' - m''\ A')$ and $(m'\ B' + m''\ B'';m'\ B'' - m''\ B')$ belong to the same classes of factor ring $\Gamma/(\|m\|)$. In other words, the complex congruence

$$A \equiv B\ (mod\ m) \tag{2}$$

is equivalent to simultaneous real congruences

$$\begin{cases} m'A' + m''A'' \equiv m'B' + m''B'', \\ m'A'' - m''A' \equiv m'B'' - m''B'. \end{cases} \tag{3}$$

Proof. At first, we assume that $A,\ B \in (m) \subset \Gamma$. This implies validity of (2) and guarantees existence of some ICN $q = q' + iq''$ such that

$$A - B = q\ m. \tag{4}$$

Multiplying (4) by $\bar{m} = m' - im''$, we obtain

$$(m'(A' - B') + m''(A'' - B''))+$$

$$i((m'(A'' - B'') + m''(A' - B')) = (q' + iq'')\|m\|.$$

It follows that

$$\begin{cases} (m'A' + m''A'') - (m'B' + m''B'') \in (\|m\|) \subset \mathbf{Z}, \\ (m'A'' - m''A') - (m'B'' - m''B') \in (\|m\|) \subset \mathbf{Z}; \end{cases}$$

thus resulting in simultaneous congruences (3). The described operations realized in the reverse sequence from (3) lead to (2). The theorem is proved.

As for any ICN $A = A' + iA''$ the congruence

$$A \equiv \langle A|_m\ (mod\ m) \tag{5}$$

and also the simultaneous congruences

$$\begin{cases} (m'A' + m''A'') \equiv |(m'A' + m''A'')|_{\|m\|}, \\ (m'A'' - m''A') \equiv |(m'A'' - m''A')|_{\|m\|} \end{cases} \tag{6}$$

are true, then from the point of view of the theorem 1 it is natural to assume that the complex residue $\alpha = \langle A|_m$ can be uniquely defined by means of a pair of the real residues

$$(a';\ a'') = (|(m'A' + m''A'')|_{\|m\|};\ |(m'A'' - m''A')|_{\|m\|}).$$

Let us prove the hypothesis validity. Let $\alpha = \alpha' + i\alpha''$ ($\alpha',\ \alpha'' \in \mathbf{Z}$). By setting $B = \alpha$ in the theorem 1 according to (3), we have

$$\begin{cases} (m'A' + m''A'') \equiv |(m'\alpha' + m''\alpha'')|_{\|m\|}, \\ (m'A'' - m''A') \equiv |(m'\alpha'' - m''\alpha')|_{\|m\|}. \end{cases} \tag{7}$$

Let us demand that the right-hand members of the same congruences of simultaneous congruences (3) and (7) coincide. Then for the real and imaginary components of residue $\langle A|_m$ we obtain the simultaneous equations

$$\begin{cases} m'\alpha' + m''\alpha'' = a', \\ m'\alpha'' - m''\alpha' = a'', \end{cases} \tag{8}$$

which solution is

$$(\alpha',\ \alpha'') = \left(\frac{m'a' + m''a''}{\|m\|};\ \frac{m'a'' - m''a'}{\|m\|} \right). \tag{9}$$

Thus,

$$\alpha = \frac{m'a' + m''a''}{\|m\|} + i\ \frac{m'a'' - m''a'}{\|m\|} = \frac{(m' + im'')(a' + ia'')}{\|m\|}$$

or

$$\langle A|_m = \alpha = \frac{m}{\|m\|}\ \left(|(m'A' + m''A'')|_{\|m\|} + i\ |(m'A'' - m''A')|_{\|m\|} \right). \tag{10}$$

The rule (10) for constructing the CSR $\langle A|_m$ generated by one-to-one correspondence between the ICN $(\alpha';\ \alpha'')$ and $(a';\ a'')$ (see (8), (9)) can be represented in specified and more constructive form.

Theorem 2. In the case when
a) a module m is a nonnegative IRN ($m' > 1,\ m'' = 0$);
b) a module $m = m' + im''$ is a complex module satisfying the condition $(m',\ m'') = 1$
for the residue $\langle A|_m$ corresponding to the arbitrary ICN $A = A' + iA''$ the followig formulas are true, accordingly,

$$\langle A' + iA''|_m = \langle A'|_m + i\langle A''|_m, \tag{11}$$

$$\begin{aligned}\langle A' + iA''|_m = \frac{1}{\|m\|}\ \big((m'|m'R_m(A)|_{\|m\|} - m''|m''R_m(A)|_{\|m\|}\big)\\ + i(m'|-m''R_m(A)|_{\|m\|} + m''|m'R_m(A)|_{\|m\|})),\end{aligned} \tag{12}$$

where

$$R_m(A) = R_m(A',\ A'') = |A' + JA''|_{\|m\|} \quad \left(J = |\frac{m''}{m'}|_{\|m\|}\right). \tag{13}$$

Proof. Let $m = m'$. Then according to (9)

$$\alpha' = \frac{m'a'}{\|m\|} = \frac{1}{m}|mA'|_{m^2} = \frac{1}{m}\left(mA' - \lfloor\frac{mA'}{m^2}\rfloor\ m^2\right) = A' - \lfloor\frac{A'}{m}\rfloor m = |A'|_m,$$

where the integer part of a real number x is designated as $\lfloor x\rfloor$. Similar calculations for a'' give the equality $\alpha'' = |A''|_m$. Thus, in the case a) the equality (11) is true.

Consider now the case b). As $|(m')^2 + (m'')^2|_{\|m\|} = 0$, taking into account (13), we have

$$a' = |m'A' + m''A''|_{\|m\|} = |m'\left(A' + A''\frac{m''}{m'}\right)|_{\|m\|} = |m'R_m(A)|_{\|m\|}, \tag{14}$$

$$\begin{aligned}a'' = |m'A'' - m''A'|_{\|m\|} = |\frac{(m')^2}{m'}A'' - m''A'|_{\|m\|} =\\ |-\frac{(m'')^2}{m'}\ A'' - m''A'|_{\|m\|} = |-m''R_m(A)|_{\|m\|}.\end{aligned} \tag{15}$$

Correctness of expressions (14) and (15) is ensured by the condition $(m',\ \|m\|) = 1$ following from the theorem condition $(m',\ m'') = 1$. Substitution (14) and (15) into (9) leads to the required outcome (12).

Theorem 2 can be also generalized to a case of arbitrary complex module m. However, for computer applications classes of the modules considered in the theorem 2 are the most acceptable.

Theorem 3. The complete set of residues $\langle\cdot|_m$ is isomorphic, correspondingly:

a) to the combinatorial square $(|\cdot|_m)^2$ of the ring $|\cdot|_m$ in the case when a natural module $m > 1$;

b) to the ring $|\cdot|_{\|m\|}$ in the case when $m = m' + im''$ is a complex module satisfying the condition $(m',\ m'') = 1$ with $|\langle\cdot|| = \|m\|$.

Proof. According to the main theorem of modular arithmetic [6, 7] in the case of natural module m, the mapping $f\colon |\cdot|_m \times |\cdot|_m \to \langle\cdot|_m$, which for every $A = A' + iA'' \in \Gamma$ associates to a pair of IRN $(|A'|_m;\ |A''|_m) \in (|\cdot|_m)^2$ a complex residue $\langle A|_m \in \langle\cdot|_m$ defined by formula (11), is bijective. Then on account of surjectivity of mapping $A \to (|A'|_m;\ |A''|_m)$, the cardinality of ring $\langle\cdot|_m$ coincides with the cardinality of a set $|\cdot|_m \times |\cdot|_m$, i.e. $|\langle\cdot||_m = m^2 = \|m\|$.

This formula also takes place in the case when the complex module m satisfies the condition $(m',\ m'') = 1$, because of surjectivity of mapping $A \to R_m(A',\ A'')$ (see (13)) and bijectivity of mapping $f\colon |\cdot|_{\|m\|} \to \langle\cdot|_m$ which for every $A \in \Gamma$ associates the complex residue $\langle A|_m$ formed by a rule (12) with an IRN $R(A',\ A'') \in |\cdot|_{\|m\|}$.

The fact that the specified-above mapping f is an isomorphism of corresponding rings in the case of the real module m is proved by equations

$$\langle\langle A' + iA''|_m + \langle B' + iB''|_m|_m = ||A'|_m + |B'|_m|_m + i||A''|_m + |B''|_m|_m,$$

$$\langle\langle A' + iA''|_m \cdot \langle B' + iB''|_m|_m = ||A'|_m \cdot |B'|_m - |A''|_m \cdot |B''|_m|_m +$$

$$i||A'|_m \cdot |B''|_m + |A''|_m \cdot |B'|_m|_m,$$

and in the case of complex module $m = m' + im''$ such that $(m',\ m'') = 1$, is proved by equations

$$R_m(A + B) = R_m(Re(A + B),\ Im(A + B)) =$$

$$|A' + B' + J(A'' + B'')|_m = |R_m(A) + R_m(B)|_{\|m\|}, \tag{16}$$

$$R_m(A \cdot B) = R_m(Re(A \cdot B),\ Im(A \cdot B)) =$$

$$R_m(A'B' - A''B'',\ A'B'' + A''B') = |A'B' - A''B'' + J(A'B'' + A''B')|_{\|m\|} =$$

$$|A'(B' + JB'') + A''(-B'' + JB')|_{\|m\|} = |A'R_m(B) + A''(J^2B'' + JB')|_{\|m\|} =$$

$$|A'R_m(B) + JA''R_m(B)|_{\|m\|} = |(A' + JA'')R_m(B)|_{\|m\|} =$$

$$|R_m(A)R_m(B)|_{\|m\|}. \tag{17}$$

It follows from (8) that the considered CSR $\langle\cdot|$ is formed by all the ICN $\alpha = \alpha' + i\alpha''$ satisfying simultaneous inequalities

$$\begin{cases} 0 \le m'\alpha' + m''\alpha'' < \|m\|, \\ 0 \le m'\alpha'' - m''\alpha' < \|m\|. \end{cases}$$

Thus, the ring $\langle\cdot|_m$ includes all the ICN arranged in a square with vertexes $A_1 = (0;\ 0),\ A_2 = (m';\ m''),\ A_3 = (m' - m'';\ m' + m''),\ A_4 = (-m'';\ m')$.

For the computer applications of interest are those complex modular number systems (CMNS) for which the ranges $\langle\cdot|_{M_k}$, where $M_k = \prod_{i=1}^{k} m_i$ with $m_1, m_2 ..., m_k$ being pairwise coprime modules, are located in the specified-above squares with the sides parallel to the axes. This means that M_k is an IRN. However, the present condition does not eliminate a possibility of using the ICN as modules $m_i,\ i = 1, 2, \dots, k$.

In particular, for practical applications it is convenient to use the CMNS with modules $m_1, m_2, \dots, m_k$ such that all or part of them are considered as macromodules which represent products of two conjugated ICN, i.e. $m = p\bar{p}$ where $p = p' + ip'',\ \bar{p} = p' - ip'';\ p', p'' \in \mathbf{Z}$. In accordance with theorems 2 and 3, it is expedient to select the numbers p and $\bar{p}$ based on conditions $p' > 0,\ p'' > 0$ and $(p',\ p'') = 1$. The last condition guarantees relative primality of the ICN p and $\bar{p}$. Such systems are called quadratic MNS.

References

[1] I.J. Akushsky, V.M. Amerbaev, I.T. Pak. *Bases of Machine Arithmetics of Complex Numbers.* Nauka, Alma-Ata, 1970. (In Russian).

[2] V.M. Amerbaev, I.T. Pak. *Parallel Calculation in Complex Plane.* Nauka, Alma-Ata, 1985. (In Russian).

[3] A.F. Chernyavsky, V.V. Danilevich, A.A. Kolyada, M.Y. Selyaninov. *High-speed Methods and Systems of Digital Information Processing.* Belgosuniversitet, Minsk, 1996. (In Russian).

[4] M. Selyaninov. Modular technique of parallel information processing. *Scientific Issues of Jan Długosz University of Częstochowa, Ser. Mathematics*, **XIII**, 43–52, 2008.

[5] A.A. Kolyada, V.V. Revinsky, M.Y. Selyaninov *et al.* Theoretical bases of modular computing structures on the finite mathematical models. *Modern Problems of Optics, Radiation Materials Science, Informatics, Radiophysics and Electronics. Proc. Sci. Research Inst. Appl. Phys. Probl.* Belgosuniversitet, Minsk, vol. 2, pp. 4–9, 1996. (In Russian).

[6] M.Y. Selyaninov. Theoretical bases of modular codification of algebraic systems. *Proc. Nat. Acad. Sci. Belarus,* No. 1, 114–119, 2002. (In Russian).

[7] M. Selyaninov. Construction of modular number systems with arbitrary finite ranges. *Scientific Issues of Jan Długosz University in Częstochowa, Ser. Mathematics*, **XIV**, 105–115, 2009.

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XV, Częstochowa 2010

VALUATION GRAPHS FOR PROPOSITIONAL LOGIC

Lidia Stępień

Institute of Mathematics and Computer Science
Jan Długosz University in Częstochowa
al. Armii Krajowej 13/15, 42-200 Częstochowa, Poland
e-mail: l.stepien@ajd.czest.pl

Abstract. In this paper we present the proof system, called *the valuation graphs system*, which is a new version of two proof procedures: Davis-Putnam and Stålmarck. The novelty is that in the rules we note which propositional variable occurring in some propositional formula does not determine the logical value of that formula. Due to Stålmarck, we define a notion of *proof width*, corresponding to the width of structure of valuation graph which is a number of applications of *dilemma rule*. The dilemma rule considers two cases, so the time of proof grows up exponentially.

1. Introduction

In recent years, there has been considerable renewed interest in the SATisfiability problem of propositional logic. The SAT is the question whether a propositional formula has a satisfying valuation. The SAT problem is known to be difficult to solve – it is the first known NP-complete problem, as it was proved by Stephen Cook in 1971. Because the SAT problem is fundamental to many practical problems in mathematics, computer science, and electrical engineering, efficient methods that can solve a large subset of SAT problems are eagerly sought. There are many competing algorithms for it and many implementations, most of them have been developed over the last two decades as highly optimized versions of the DPLL procedure of [3] and [4]. As a motivation, we refer to the Stålmarck patented method [7] for solving the propositional satisfiability problem in practical applications. For instance, it has been used successfully for industrial-scale problems [2]. This is the algorithm which is acceptably efficient in a large number of important cases (proof width).

Moreover, the algorithm itself is not yet widely known, and it is interesting to investigate how it performs. We have tried to do this, and the valuation graphs system has been created as a new version of Stålmarck procedure [1, 5, 6, 8].

The paper is organised as follows. Next section provides the preliminary notions. The definitions and rules of building the valuation graph are shown in section 3. In section 4 we give proofs of soundness and completness of our system. Section 5 presents the complexity of valuations graphs procedure. Section 6 completes the paper with some conclusions and future work.

2. Preliminary notions

We define the valuation graphs for propositional formulas which are built with logical connective $\rightarrow$, called *implication*, and the false symbol $\mathbf{F}$. The true symbol $\mathbf{T}$ can be defined as a formula: $\mathbf{F} \rightarrow \mathbf{F}$. Every propositional formula can be equivalently translated, in linear time, to implication form applying the Stålmarck procedure [6] and the following equivalences:

$$\begin{aligned} p \vee q &\leftrightarrow \neg\, p \rightarrow q \\ p \wedge q &\leftrightarrow \neg\, (p \rightarrow \neg\, q) \\ \neg\, \neg\, p &\leftrightarrow p \\ \neg\, p &\leftrightarrow p \rightarrow \mathbf{F} \end{aligned}$$

Due to Stålmarck, every implication will be a triplet (c, β, γ), eventually with indices, where β and γ are subformulas and c is a new propositional variable which value is equivalent to the value of implication $\beta \rightarrow \gamma$; so $c \leftrightarrow (\beta \rightarrow \gamma)$. Each propositional formula α in implication form will be represented by a sequence $\bar{\mathbf{d}} = (\mathbf{d_1}, \ldots, \mathbf{d_n})$, where $\mathbf{d_i}$ is a triplet (c_i, β_i, γ_i) for $1 \leqslant i \leqslant n$, $\mathbf{d_n}$ is a main implication, n is the number of occurrences $\rightarrow$ in α. Then β_i and γ_i can be propositional variables $p, q, r, \ldots$, propositional constants $\mathbf{F}$ or $\mathbf{T}$ or a new triplet variable c_j. By $|\bar{\mathbf{d}}|$ we denote the number of triplets in the sequence $\bar{\mathbf{d}}$ ($|\bar{\mathbf{d}}| = n$). Our procedure inputs a propositional formula α in the following form:

$$\bar{\mathbf{d}} : \underbrace{c_1 \leftrightarrow (\beta_1 \rightarrow \gamma_1)}_{\mathbf{d_1}}, \underbrace{c_2 \leftrightarrow (\beta_2 \rightarrow \gamma_2)}_{\mathbf{d_2}}, \ldots, \underbrace{c_n \leftrightarrow (\beta_n \rightarrow \gamma_1)}_{\mathbf{d_n}}.$$

To check if the given formula is satisfiable, we construct its satisfying valuation. If constructing is failure, the propositional formula is unsatisfiable. Instead of the value of propositional variable p, we will say about substitution of propositional constant: $p{:=}\mathbf{F}$ or $p{:=}\mathbf{T}$. Similarly, by $p{:=}q$ or $p{:=}{-}q$ we denote

substitution of p. By "$-p$" we denote value of p contrary to its present value. For simplicity we will write "=" instead ":=". The substitution sets will be denoted by $\Sigma, \Delta, \ldots$, eventually with indices. By $\beta_{\surd}$ we denote a substitution of one of the constants $\{\mathbf{T}, \mathbf{F}\}$ in place of β. This means that we do not need a value of β to determine a value of whole propositional formula. If in Σ we have contradiction of one of the form: $x{=}{-}x$; or $x{=}\mathbf{T}$ and $x{=}\mathbf{F}$; or $x{=}y$ and $x{=}{-}y$ for some propositional variables x and y, then we denote this contradiction by $\bot_x$ and we say that Σ is contradictory (we replace Σ by $\bot$).

3. Rules and definitions

First, in this section we present the rules of substitution of propositional values (constant) according to the truth table of $\rightarrow$: the reduction rules (RR) and the dilemma rule (RD). Next we define a valuation graph.

In general, the reduction rule has the following form:

$$\frac{\mathbf{d}\ [\ \Sigma\ ]}{[\ \Sigma'\ ]},$$

where $\mathbf{d}$ is a some triplet in a sequence of triplets representing a propositional formula α which tautology/satisfiability we check, and Σ' comes from Σ by adding the conclusions of that rule.

The special cases of **THE REDUCTION RULES**:

$$\frac{\mathbf{F} \leftrightarrow (\beta \rightarrow \gamma)}{\beta = \mathbf{T}, \gamma = \mathbf{F}} \quad (RR1) \qquad \frac{c \leftrightarrow (\beta \rightarrow \mathbf{T})}{c = \mathbf{T}, \beta_{\surd}} \quad (RR2)$$

$$\frac{c \leftrightarrow (\mathbf{F} \rightarrow \gamma)}{c = \mathbf{T}, \gamma_{\surd}} \quad (RR3) \qquad \frac{c \leftrightarrow (\mathbf{T} \rightarrow \gamma)}{c = \gamma} \quad (RR4)$$

$$\frac{c \leftrightarrow (\beta \rightarrow \mathbf{F})}{c = -\beta} \quad (RR5) \qquad \frac{c \leftrightarrow (c \rightarrow \gamma)}{c = \mathbf{T}, \gamma = \mathbf{T}} \quad (RR6)$$

$$\frac{c \leftrightarrow (\beta \rightarrow \beta)}{c = \mathbf{T}, \beta_{\surd}} \quad (RR7)$$

In each reduction rule we have given some triplet representing subformula of propositional formula α, and a set of substitutions Σ. In particular, subformulas β and γ can be logical constants. Conclusions (substitutions) of each

reduction rule are added to the set Σ, and the sequence of triplets representing α is reduced by removing the given triplet. Moreover, the RR2, RR3 and RR7 rules say that the value of implication does not depend on its subformulas. The logical value of implication, so the value of a triplet variable representing this implication too, can be sometimes deduced from the partial valuation with logical constants occurring in implication. Thus, in all of those rules the logical constant **T** substitutes a triplet variable.

As can be seen above, the reduction rules are not limited to deduce conclusions of the form: β *is* **T** and γ *is* **F** (see RR1), but includes also conclusions of the form: c *has the same value as* γ (see RR4) or c *and* β *have different values* (see RR5).

THE DILEMMA RULE:

$$\frac{\bar{\mathbf{d}}[\Sigma]}{\bar{\mathbf{d}}[\Sigma \cup \{\mathbf{x} = \mathbf{T}\}] \quad \Big| \quad \bar{\mathbf{d}}[\Sigma \cup \{\mathbf{x} = \mathbf{F}\}]} \quad (RD)$$

where x occurs in a sequence $\bar{\mathbf{d}}$ and for x there does not exist substitution in Σ.

When we cannot apply any reduction rule in a sequence, we have to use the dilemma rule. Then we obtain two sets of substitutions which arise from the set Σ: the first one by adding x=**T** and the second by adding x=**F**. The dilemma rule is used to the variable from a sequence of a triplets if there does not exist a substitution in Σ for this variable. So, the sequence will not be reduced. Only one of substitutions of variable x is true, so we have the dilemma which set of substitutions is searched by us. Now, a merger of both sets of substitutions is necessary (see definition 1).

By $\bar{\mathbf{d}}[\Sigma]$ we denote the lable of a vertex of a valuation graph which is defined by induction on the length of propositional formula.

Definition 1. Let $\bar{\mathbf{d}}$ be a finite sequence of triplets $(d_1, d_2, \ldots, d_n)$ representing propositional formula α and $\Sigma = \emptyset$.

1. The single vertex labeled by $\bar{\mathbf{d}}[\Sigma \cup \{c_n = \mathbf{F}\}]$ is the *valuation graph* for α.

2. If $\mathcal{G}$ is a valuation graph for α, $\bar{\mathbf{d}}[\Sigma]$ is a label of a leaf, and $\mathcal{G}^*$ arises from $\mathcal{G}$ by adding a new vertex (and an edge from it to the leaf $\bar{\mathbf{d}}[\Sigma]$) labeled by $\bar{\mathbf{d}}'[\Sigma']$, which is deduced by applying one of the reduction rules, then $\mathcal{G}^*$ is the *valuation graph* for α.

3. If $\mathcal{G}$ is the valuation graph for α, $\bar{\mathbf{d}}[\Sigma]$ is a label of a leaf, and $\mathcal{G}^*$ arises from $\mathcal{G}$ by adding two new vertices (and two edges from them to the leaf $\bar{\mathbf{d}}[\Sigma]$, respectively) labeled by $\bar{\mathbf{d}}[\Sigma \cup \{\beta = \mathbf{T}\}]$ and $\bar{\mathbf{d}}[\Sigma \cup \{\beta = \mathbf{F}\}]$, which are deduced by applying dilemma rule, then $\mathcal{G}^*$ is *the valuation graph* for α.

4. If $\mathcal{G}$ is the valuation graph for α and $\bar{\mathbf{d}}_\mathbf{1}[\Delta_1]$ and $\bar{\mathbf{d}}_\mathbf{2}[\Delta_2]$ are labels of leaves in $\mathcal{G}$ obtained from vertices labeled by $\bar{\mathbf{d}}[\Sigma \cup \{\beta = \mathbf{T}\}]$ and $\bar{\mathbf{d}}[\Sigma \cup \{\beta = \mathbf{F}\}]$, respectively; $\bar{\mathbf{d}}_\mathbf{1}$ and $\bar{\mathbf{d}}_\mathbf{2}$ are empty sequences or in the set of substitutions there exists a contradictory, then $\mathcal{G}^*$ is the *valuation graph* for α obtained from $\mathcal{G}$ by adding a new vertex (and edge from it to those leaves) labeled by $\bar{\mathbf{d}}[\Delta]$, where

$$\Delta = \begin{cases} \bot & \text{when } \Delta_1 \text{ and } \Delta_2 \text{ are contradictory} \\ \Delta_1 & \text{when } \Delta_2 \text{ is contradictory and } \Delta_1 \text{ is not contradictory} \\ \Delta_2 & \text{when } \Delta_1 \text{ is contradictory and } \Delta_2 \text{ is not contradictory} \\ (\Delta_1 \cap \Delta_2) \cup \{\alpha = \gamma\} & \text{if } \Delta_1 \text{ and } \Delta_2 \text{ are not contradictory and } \{\alpha = \gamma\} \text{ occurs on one of the paths and } \alpha_{\surd} \text{ occurs on the second, simultaneously} \\ (\Delta_1 \cap \Delta_2) \cup \{\beta_{\surd}\} & \text{when } \Delta_1 \text{ and } \Delta_2 \text{ are not contradictory and } \Delta_1 \cap \Delta_2 = \Sigma \\ \Delta_1 \cap \Delta_2 & \text{otherwise, in particular, when } \Delta_1 \text{ and } \Delta_2 \text{ are not contradictory} \end{cases}$$

The set of substitutions Δ of path Θ of valuation graph $\mathcal{G}$ is a conjunction of substitutions, and the set of triplets $\bar{\mathbf{d}}$ is a conjunction of triplets.

Definition 2. The path Θ of valuation graph $\mathcal{G}$ is *closed* when the set of substitutions Δ of this path is contradictory.

Definition 3. The path Θ of valuation graph $\mathcal{G}$ is *maximal* when the set of triplets of this path is empty and the set of substitutions includes substitutions of all the propositional variables.

Definition 4. The valuation graph $\mathcal{G}$ is *closed* when all its paths are closed.

Definition 5. The valuation graph $\mathcal{G}$ is *maximal* when it is not closed.

In other words, if the maximal path exists in a valuation graph, then the valuation graph is maximal.

Definition 6. *Proof* of propositional formula α is the closed valuation graph for α.

To check if α is a tautology, we try to construct a model for $\neg\alpha$, i.e. we build a valuation v such that $v(\neg\alpha) = \mathbf{1}$. The valuation will be meant as a set of substitutions in which every propositional variable occurring in formula α has a substitution. Building the valuation, we start from substitution $\mathbf{F}$ into triplet variable representing the whole formula. Every next substitution is a conclusion of the reduction rules or the dilemma rule. If the set of substitutions includes contradictory, then α is a tautology, otherwise v is a model for $\neg\alpha$.

Example 1. The formula $(p \xrightarrow{b_1} p)$ in implication form is represented by the triplet (b_1, p, p). Building a valuation graph (Figure 1), we start from single vertex (I) labeled by a sequence of triplets, and the set of substitutions contains only the substitution $\mathbf{F}$ into b_1. The triplet and the substitution $b_1 = \mathbf{F}$ are premises of the reduction rule RR1. Thus, the conclusions of those rules are added to the set of substitutions, and the triplet (b_1, p, p) is removed from a sequence of triplets (II). Now we have empty sequence of triplets and contradictory $\perp_p$ in the set of substitution ($p = \mathbf{F}$ and $p = \mathbf{T}$). So, the set of substitutions is contradictory. Therefore, we have built the closed valuation graph for $p \to p$, so the formula $p \to p$ has the proof in the valuation graphs system.

$$\begin{array}{ll} \text{I} & ((b_1, p, p))[b_1 = \mathbf{F}] \\ & \qquad\qquad \Big| \; (RR1_{b_1}) \\ \text{II} & (\emptyset)[b_1 = \mathbf{F}, p = \mathbf{T}, p = \mathbf{F}] \\ \text{III} & (\emptyset)[b_1 = \mathbf{F}, \perp_p] \\ \text{IV} & (\emptyset)[\perp] \end{array}$$

Figure 1. The closed valuation graph.

Example 2. For the propositional formula:

$$(p \xrightarrow{b_4} p) \xrightarrow{b_5} (((p \xrightarrow{b_1} q) \xrightarrow{b_2} p) \xrightarrow{b_3} q)$$

we get the following sequence of triplets:

$$\big((b_1, p, q), (b_2, b_1, p), (b_3, b_2, q), (b_4, p, p), (b_5, b_4, b_3)\big).$$

Now we are constructing the valuation graph:

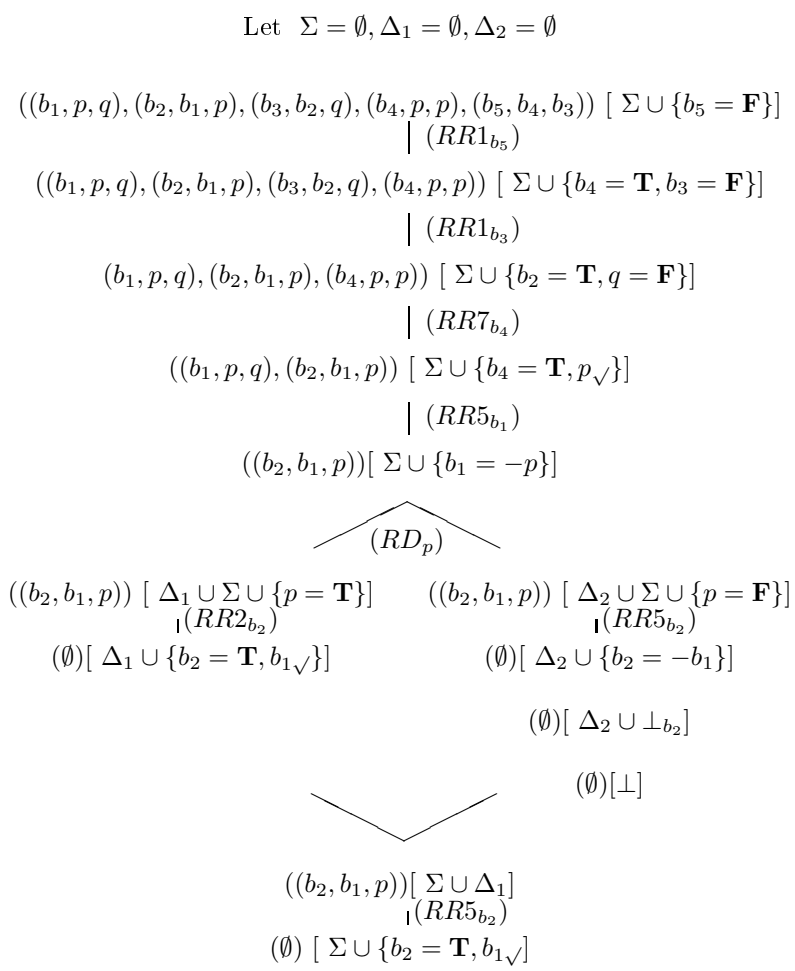


Figure 2. The maximal valuation graph.

4. Soundness and completness of valuation graphs system

Theorem 1 (Soundness). *If there exists a proof for α in the valuation graphs system, then α is a tautology.*

Proof: From the assumption that a proof for α exists in the valuation graphs system it follows that α has a closed valuation graph. Because the valuation graph representing all valuations is closed, every path in this graph is closed. Thus, contradictory occurs in the every path. So α is a tautology.

Theorem 2 (Completness). *If α is a tautology, then a proof of valuation graphs system exists for α.*

Proof: We assume that α is a tautology and there does not exist the proof for α in the valuation graphs system. This means that any valuation graph is not closed. Thus, in each valuation graph for α the contradictory does not occur at least in one path. Let $\mathcal{G}$ be one of valuation graphs for α. Because $\mathcal{G}$ has a path, which does not include a contradictory, hence the set of substitutions in label of leaf of this path contains the substitutions for all variables occurring in the triplet form of this formula. In particular, there are substitutions for propositional variables, subformulas and the whole formula α (because we have started building of valuation graph from $b_n = \mathbf{F}$ added to $\mathcal{G}$). Thus, a valuation v, constructed above, is a model for $\neg\alpha$. It is a contradictory to the assumption that α is a tautology.

The valuation graphs system with the reduction rules and the dilemma rule is sound and complete for propositional formulas built of propositional variable, implication and logical constant. Each propositional formula can be translated to implication form by applying the Stålmarck procedure (see [6]) in linear time. So the valuation graphs system is sound and complete for classical propositional logic.

Corollary 1. *A propositional formula α is satisfiable iff a valuation graphs system $\mathcal{G}$ for $\neg\alpha$ is maximal.*

Proof: A propositional formula α is satisfiable iff (by definition) there exists a Boolean valuation v such that α is true iff (by definition) a valuation graph for $\neg\alpha$ is maximal.

5. Complexity of procedure of valuation graphs system

The valuation graphs system allows searching for proofs and models for large class of formulas in linear or polynomial time with respect to the length of formula and width of its (maximal or closed) valuation graph. This estimation follows from analysis of branching valuation graphs of those formulas and depends exponentially on the width of valuation graph but not on the length of formula.

By *representation of structure* of valuation graph we mean its substructure consisted of vertices which labels are premises of the dilemma rule, its direct consequences and vertices of their merger (see item 4 of definition of valuation graph). Notice that for some valuation graph its representation of structure is

unique. Consider all the possible substructures of representation of structure for some valuation graph. We associate each substructures with the number of leaves.

Definition 7. A *width* of valuation graph is the maximal number of leaves by all substructures of its representation of structure.

Definition 8. A formula α is *i-hard* when there exists a maximal or closed valuation graph for α with the width equal to $i+1$.

By definition 8, a formula is 1-hard when in its valuation graph we apply the dilemma rule once (by definition 7, the width of this valuation graph is equal to 2).

The time of building of valuation graph is growing up when a valuation graph branches. Due to Stålmarck, we present the recursively function of complexity:

$$\begin{aligned} g(0,n) &= 2 \cdot n, \\ g(k,n) &= \sum_{i=1}^{n} 2 \cdot i \cdot g(k-1,n), \end{aligned}$$

where: n is a length of formula,
k is a width of valuation graph.

The function $g(k,n)$ is at most n^{2k+1}, so the complexity of the presented procedure is $O(n^{2k+1})$.

6. Conclusions and future work

In the valuation graphs system the time of finding a satisfying valuation depends exponentially on the width of valuation graph but not on a length of formula. A width of valuation graph depends on the number of applications of dilemma rule. Hence, the best place for optimisation is a place where we must choose a propositional variable (a triplet variable representing a propositional formula) as a premise of the dilemma rule.

The valuation graphs system was implemented and the prototype version is tested. We are working out at experimental results which will be presented soon.

References

[1] M. Björk. *A first order extension of Stålmarck's method.* Department of Computer Science and Engineering Chalmers University of Technology and Göteborg University, Göteborg, Sweden, 2006.

[2] A. Borälv. The industiral success of verification tools based on Stålmarck's method. *Proc.* 9^{th} *Int. Conf. on Computer Aided Verification*, LNCS, vol. 1254, pp. 7–10, Springer, 1997.

[3] M. Davis, H. Putnam. A computing procedure for quantification theory. *Journal of the ACM*, **7**, 201–215, 1960.

[4] M. Davis, G. Logemann, D. Loveland. A machine program for theorem proving. *Communications of the ACM*, **5**, 394–397, 1962.

[5] M. Sheeran, G. Stålmarck. A tutorial on Stålmarck's proof procedure for propositional logic. *Formal Methods in System Design*, **16** (1), 23–58, 2000.

[6] G. Stålmarck. A note on the computational complexity of the pure classical implication calculus. *Inf. Process. Lett.*, **31** (6), 277–278, 1989.

[7] G. Stålmarck. *System for Determining Propositional Logic Theorems by Applying Values and Rules to Triplets that are Generated from a Formula.* Swedish Patent No. 467 076, 1992; U.S. Patent No. 5 276 897, 1994; European Patent No. 0 403 454, 1995.

[8] G. Stålmarck, M. Säflund. Modelling and verifying systems and software in propositional logic. *Proc. SAFECOMP'90*, pp.31–36, Pergamon Press, 1990.

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XV, Częstochowa 2010

PROCESSOR SHARING QUEUEING SYSTEMS WITH NON-HOMOGENEOUS CUSTOMERS

Oleg Tikhonenko

Institute of Mathematics and Computer Science
Jan Długosz University in Częstochowa
al. Armii Krajowej 13/15, 42-200 Częstochowa, Poland
e-mail: oleg.tikhonenko@gmail.com

Abstract. We investigate processor sharing queueing systems with non-homogeneous customers having some random space requirements. Such systems have been used to model and solve various practical problems occurring in the design of computer or communicating systems. The above non-homogenity means that each customer (independently of others) has some random space requirement and his length (or amount of work for his service) generally depends on the space requirement. In real systems, a total sum of space requirements of customers presenting in the system is limited by some constant value (memory capacity) $V > 0$. We estimate loss characteristcs for such a system using queueing models with unlimited memory space.

1. Introduction

Egalitarian processor sharing (EPS) systems are used for modeling of computer and communicating networks [1]. Presently, they are applicable to situations where a common resource is shared by a varying number of concurrent users [2] (for example, to WEB-servers modeling [3]).

The EPS discipline was first introduced by Kleinrock [4] as a limiting case for modeling time sharing systems. The aim of the paper is to analyze classical and non-classical EPS systems. First, we shall analyze the classical EPS system notated by $M/G/1-EPS$. All the customers present in the classical $M/G/1-EPS$ system are served simultaneously. If there are $n > 1$ customers in the system at an arbitrary instant, then all of them are served at this instant n times slowly than in the case of $n = 1$.

Later on, the customer length means the amount of work necessary for customer's service, i.e. the service time under condition that there are no other customers in the system during his presence in it. Analogously, the residual length of the customer means his residual service time after some time instant under the same condition (see [2]).

We introduce the following additional assumption for the classical $M/G/1-EPS$ system. Assume that each customer is characterized by some non-negative random capacity. This random variable can be interpreted as a part of system's memory space used by the customer during his presence in the system. A total sum of customer capacities $\sigma(t)$ in the system at arbitrary time t is referred as the total customers capacity.

The random value $\sigma(t)$ can be limited by some constant value V $(0 < V < \infty)$, which is called the memory volume of the system. In this case we have a non-classical processor sharing system that will be notated by $M/G/1(V)-EPS$.

The purpose of the paper is

1) to obtain the non-stationary and stationary distribution of total customers capacity in the system $M/G/1-EPS$;

2) to determine some estimations of loss characteristics for systems $M/G/1(V)-EPS$ with limited memory space $(V < \infty)$ based on the model with unlimited one;

3) to compare processor sharing systems $M/G/1(V)-EPS$ and $M/G/1-EPS$ from the viewpoint of estimation of loss characteristics.

2. Classical processor sharing system

In this section we investigate the classical system $M/G/1-EPS$. Denote by $\eta(t)$ the number of customers present in the system at time t and by $\xi_i^*(t)$ the residual length of the ith customer at this time, $i=\overline{1,\eta(t)}$. Let

$$F(x,t)=\mathbf{P}\{\zeta<x,\xi<t\}$$

be the joint distribution function of the customer capacity ζ and his length ξ (we assume that customer's capacity and his length do not depend on his arrival time and on characteristics of other customers). Then $L(x)=F(x,\infty)$ and $B(t)=F(\infty,t)$ are the distribution functions of the random variables ζ and ξ, respectively. Let a be an arrival rate of entrance flow of customers,

$$\alpha(s,q)=\int_0^\infty\int_0^\infty e^{-sx-qt}dF(x,t)$$

be the double Laplace-Stieltjes transform (with respect to x and t) of the distribution function $F(x,t)$, $\varphi(s)=\alpha(s,0)$, and $\beta(q)=\alpha(0,q)$ be the Laplace-Stieltjes transform of the distributin functions $L(x)$ and $B(t)$, respectively.

$D(x,t) = \mathbf{P}\{\sigma(t) < x\}$ is the distribution function of total customers capacity at time t,

$$\delta(s,t) = \mathbf{E}e^{-s\sigma(t)} = \int_0^\infty e^{-sx} d_x D(x,t)$$

is the Laplace-Stieltjes transform of the function $D(x,t)$ with respect to x,

$$\overline{\delta}(s,q) = \int_0^\infty e^{-qt}\mathbf{E}e^{-s\sigma(t)}dt = \int_0^\infty e^{-qt}\delta(s,t)dt$$

is the Laplace transform of the function $\delta(s,t)$ with respect to t.

The mixed $(i+j)$th moments of the random variables ζ and ξ (if they exist) take the form:

$$\alpha_{i\,j} = \mathbf{E}(\zeta^i \xi^j) = (-1)^{i+j} \frac{\partial^{i+j}}{\partial s^i \partial q^j} \alpha(s,q)\Big|_{s=0,q=0}.$$

Assume that customers in the considered system at an arbitrary time t are numerated as random; i.e. if the number of customers is k, then there are $k!$ ways to enumerate them, and each enumeration can be chosen with the same probability $1/k!$.

One can easily show that the system under consideration is described by the Markov process

$$(\eta(t), \xi_i^*(t), i = \overline{1, \eta(t)}), \tag{1}$$

where components $\xi_i^*(t)$ are absent if $\eta(t) = 0$. In this case we also have $\sigma(t) = 0$.

In what follows, to simplify the notation, we denote $Y_k = (y_1, \ldots, y_k)$. Sometimes in the case $k = 1$, instead of Y_1 we write y_1 or the value that this component takes, and in the case $k = 2$, instead of Y_2 we write (y_1, y_2) or their values. In other words, we sometimes specify vectors of small dimensions by indicating their components. We also use the notation $(y_1, \ldots, y_k, u) = (Y_k, u)$.

We characterize the process (1) by functions with the following probabilistic sense:

$$P_0(t) = \mathbf{P}\{\eta(t) = 0\}; \tag{2}$$

$$\Theta_k(Y_k, t) = \mathbf{P}\{\eta(t) = k, \xi_j^*(t) < y_j, j = \overline{1,k}\},\ k = 1, 2, \ldots; \tag{3}$$

$$P_k(t) = \mathbf{P}\{\eta(t) = k\} = \Theta_k(\infty_k, t),\ k = 1, 2, \ldots, \tag{4}$$

where $\infty_k = (\infty, \ldots, \infty)$ is a k-component vector.

Note that the functions $\Theta_k(Y_k, t)$ are symmetric with respect to permutations of components of the vector Y_k due to our random enumeration of customers in the system.

Let us determine the function $\overline{\delta}(s,q)$ under zero initial condition $\eta(0) = \sigma(0) = 0$.

Denote by $\overline{p}_0(q) = \int_0^\infty e^{-qt} P_0(t)dt$ and $\overline{\theta}_k(Y_k,q) = \int_0^\infty e^{-qt}\Theta_k(Y_k,t)dt$ the Laplace transforms with respect to t of the functions $P_0(t)$ and $\Theta_k(Y_k,t)$, respectively. It is known (see [2]) that

$$\overline{p}_0(q) = [q + a - a\pi(q)]^{-1} \tag{5}$$

under zero initial condition, where $\pi(q)$ is the Laplace-Stieltjes transform of the busy period distribution function for the system under consideration. Note [2] that $\pi(q)$ is a unique solution of the functional equation $\pi(q) = \beta(q + a - a\pi(q))$ such that $|\pi(q)| \le 1$.

Lemma 1. *Under zero initial condition, the functions $\overline{\theta}_k(Y_k,q)$, where $k = 1,2,\ldots,$ have the following form:*

$$\overline{\theta}_k(Y_k,q) = \overline{p}_0(q)\prod_{i=1}^{k}\int_0^{y_i}[q + a - aB(u)]du.$$

Proof. Using the method of auxiliary variables [5] and taking into account the symmetric property of the functions $\Theta_k(Y_k,t)$, we can write out partial differential equations for functions (3):

$$\frac{\partial\Theta_1(y,t)}{\partial t} - \frac{\partial\Theta_1(y,t)}{\partial y} + \frac{\partial\Theta_1(y,t)}{\partial y}\bigg|_{y=0} = aP_0(t)B(y) - a\Theta_1(y,t) + \\ + \frac{\partial\Theta_2(y,u,t)}{\partial u}\bigg|_{u=0}; \tag{6}$$

$$\frac{\partial\Theta_k(Y_k,t)}{\partial t} - \frac{\partial\Theta_k(Y_k,t)}{\partial y_k} + \frac{\partial\Theta_k(Y_k,t)}{\partial y_k}\bigg|_{y_k=0} = a\Theta_{k-1}(Y_{k-1},t)B(y_k) - \\ - a\Theta_k(Y_k,t) + \frac{\partial\Theta_{k+1}((Y_k,u),t)}{\partial u}\bigg|_{u=0},\ k = 2,3,\ldots. \tag{7}$$

Passing to Laplace transform in the equations (6), (7), we obtain

$$-\frac{\partial\overline{\theta}_1(y,q)}{\partial y} = a\overline{p}_0(q)B(y) - (q+a)\overline{\theta}_1(y,q) - \frac{\partial\overline{\theta}_1(y,q)}{\partial y}\bigg|_{y=0} + \\ + \frac{\partial\overline{\theta}_2(y,u,q)}{\partial u}\bigg|_{u=0}; \tag{8}$$

$$-\frac{\partial\overline{\theta}_k(Y_k,q)}{\partial y_k} = a\overline{\theta}_{k-1}(Y_{k-1},q)B(y_k) - (q+a)\overline{\theta}_k(Y_k,q) -$$

$$-\frac{\partial\overline{\theta}_k(Y_k,q)}{\partial y_k}\bigg|_{y_k=0}+\frac{\partial\overline{\theta}_{k+1}((Y_k,u),q)}{\partial u}\bigg|_{u=0},\ k=2,3,\ldots. \tag{9}$$

By direct substitution, we can prove that the solution of Eqs. (8) and (9) has the form

$$\overline{\theta}_k(Y_k,q)=C(q)\prod_{i=1}^{k}\int_0^{y_i}[q+a-aB(u)]du, \tag{10}$$

where $C(q)$ is some function that can be determined if we substitute the relation (10) into Eq. (8). Then, we have $C(q)=\overline{p}_0(q)$.

The lemma is proved.

Let $\beta_i=\mathbf{E}\xi^i=(-1)^i\beta^{(i)}(0)$ be the ith moment of the customer length, $i=1,2,\ldots$.

Corollary 1. *If $\rho=a\beta_1<1$, then the limits $\theta_k(Y_k)=\lim\limits_{t\to\infty}\Theta_k(Y_k,t)$, $k=1,2,\ldots$, exist being independent of initial condition and have the form:*

$$\theta_k(Y_k)=(1-\rho)a^k\prod_{i=1}^{k}\int_0^{y_i}[1-B(u)]du.$$

Proof. If $\rho<1$, then the process (1) is regenerative with points of regeneration coinciding with epochs of termination of busy periods. It follows from the theory of regenerative processes [6] that the limit $\lim\limits_{t\to\infty}\Theta_k(Y_k,t)=$ $=\theta_k(Y_k)$ exists and

$$\theta_k(Y_k)=\lim_{q\to0}q\overline{\theta}_k(Y_k,q)=(1-\rho)a^k\prod_{i=1}^{k}\int_0^{y_i}[1-B(u)]du.$$

Corollary 2. *Let $\overline{p}_k(q)$ be the Laplace transform of the function $P_k(t)$, $k=0,1,\ldots$, under zero initial condition. Then we have*

$$\overline{p}_k(q)=\frac{a^k(1-\pi(q))^k}{(q+a-a\pi(q))^{k+1}}.$$

Proof. It is obvious that $\overline{p}_k(q)=\overline{\theta}_k(\infty_k,q)$. Let us prove the equality

$$\int_0^{\infty}(q+a-aB(y))dy=\frac{a(1-\pi(q))}{q+a-a\pi(q)}. \tag{11}$$

It follows from the normalization condition written in terms of Laplace transforms that

$$\overline{p}_0(q)+\sum_{k=1}^{\infty}\overline{\theta}_k(\infty_k,q)=1/q,$$

whence, taking into account the result of lemma 1, we obtain:

$$1+\sum_{k=1}^{\infty}\left[\int_0^{\infty}(q+a-aB(y))dy\right]^k=\frac{1}{q}[q+a-a\pi(q)].$$

From the last relation we have (11). Now, the statement of the corollary follows from formulae (5) and (10).

From corollary 1 we can obtain the known relation for the stationary distribution $\{p_k\}$ of the number of customers in the system $(\rho=a\beta_1<1)$[2]:

$$p_k=\theta_k(\infty_k)=(1-\rho)\rho^k,\ k=0,1,\dots.$$

Let $\chi(t)$ be the capacity of a customer being on service at the time t and $\xi^*(t)$ be the residual length of this customer at the time t. We shall use the notation $E_y(x)=\mathbf{P}\{\chi(t)<x|\,\xi^*(t)=y\}$. It is known [7] that the Laplace–Stieltjes transform of the conditional distribution function $E_y(x)$ has the form:

$$e_y(s)=[1-B(y)]^{-1}\int_{x=0}^{\infty}e^{-sx}\int_{u=y}^{\infty}dF(x,u). \tag{12}$$

We introduce the notation

$$d_{Y_k}\Theta_k(Y_k,t)=\mathbf{P}\{\eta(t)=k,\xi_i^*(t)\in[y_i,y_i+dy_i),i=\overline{1,k}\}=$$

$$=\frac{\partial^k\Theta_k(Y_k,t)}{\partial y_1\dots\partial y_k}dy_1\dots dy_k.$$

Later on, we use the notation $\overset{k}{\underset{i=1}{*}}R_i(x)$ for Stieltjes convolution of distribution functions $R_i(x)$, $i=1,2,\dots$, $R_i(x)=0$, if $x\le 0$.

Theorem 1. *For zero initial condition, the function $\overline{\delta}(s,q)$ is determined by the relation*

$$\overline{\delta}(s,q)=\{[q+a-a\pi(q)][1-I(s,q)]\}^{-1},$$

where

$$I(s,q)=\int_0^{\infty}(q+a-aB(y))e_y(s)dy$$

and $e_y(s)$ is determined by relation (12).

Proof. The distribution function $D(x,t)$ can be represented as

$$D(x,t)=P_0(t)+$$

$$+\sum_{k=1}^{\infty}\int_0^{\infty}\dots\int_0^{\infty}\mathbf{P}\{\sigma(t)<x|\,\eta(t)=k,\xi_i^*(t)=y_i,i=\overline{1,k}\}d_{Y_k}\Theta_k(Y_k,t).$$

From the random enumeration of components of the vector Y_k it is obvious that

$$\mathbf{P}\{\sigma(t) < x | \, \eta(t) = k, \xi_i^*(t) = y_i, i = \overline{1,k}\} = \underset{i=1}{\overset{k}{*}} E_{y_i}(x).$$

Then we get:

$$D(x,t) = P_0(t) + \sum_{k=1}^{\infty} \int_0^{\infty} \cdots \int_0^{\infty} \underset{i=1}{\overset{k}{*}} E_{y_i}(x) d_{Y_k} \Theta_k(Y_k, t).$$

Passing in the last relation to Laplace–Stieltjes transform with respect to x, we have:

$$\delta(s,t) = P_0(t) + \sum_{k=1}^{\infty} \int_0^{\infty} \cdots \int_0^{\infty} \prod_{i=1}^{k} e_{y_i}(s) d_{Y_k} \Theta_k(Y_k, t).$$

Passing to Laplace transform with respect to t, we obtain:

$$\overline{\delta}(s,q) = \overline{p}_0(q) + \sum_{k=1}^{\infty} \int_0^{\infty} \cdots \int_0^{\infty} \prod_{i=1}^{k} e_{y_i}(s) d_{Y_k} \overline{\theta}_k(Y_k, q),$$

where $d_{Y_k}\overline{\theta}_k(Y_k,q) = \overline{p}_0(q) \prod_{i=1}^{k} [q + a - aB(y_i)] dy_i$ (it follows from Eq. (8) and the relation $C(q) = \overline{p}_0(q)$). Then we get:

$$\overline{\delta}(s,q) = \overline{p}_0(q) + \overline{p}_0(q) \sum_{k=1}^{\infty} \int_0^{\infty} \cdots \int_0^{\infty} \prod_{i=1}^{k} e_{y_i}(s)[q + a - aB(y_i)] dy_i =$$

$$= \overline{p}_0(q) \left\{ 1 + \sum_{k=1}^{\infty} \left[\int_0^{\infty} (q + a - aB(y)) e_y(s) dy \right]^k \right\} =$$

$$= \overline{p}_0(q) \left[1 + \sum_{k=1}^{\infty} (I(s,q))^k \right], \tag{13}$$

where

$$I(s,q) = \int_0^{\infty} (q + a - aB(y)) e_y(s) dy.$$

Now, the statement of the theorem follows from formula (13).

Corollary 3. *If the random variables ζ and ξ are independent, we obtain:*

$$\overline{\delta}(s,q) = [q + a(1 - \pi(q))(1 - \varphi(s))]^{-1}. \tag{14}$$

Proof. In this case, taking into account Equation (12) and the relation $F(x,t) = L(x)B(t)$, we have that

$$I(s,q) = \int_{y=0}^{\infty} \frac{q + a - aB(y)}{1 - B(y)} \int_{x=0}^{\infty} e^{-sx} \int_{u=y}^{\infty} dF(x,u) =$$

$$= \varphi(s) \int_0^{\infty} [q + a - aB(y)]dy = \frac{a\varphi(s)(1 - \pi(q))}{q + a - a\pi(q)},$$

whence relation (14) follows.

Corollary 4. *Under zero initial condition, the Laplace transform $g(s,q)$ with respect to t of generation function $P(z,t) = \sum_{k=0}^{\infty} P_k(t)z^k$, $|z| \le 1$, of the customers number in the system at time t has the following form:*

$$g(z,q) = \int_0^{\infty} e^{-qt} P(z,t)dt = [q + a(1 - z)(1 - \pi(q))]^{-1}. \tag{15}$$

Proof. It follows from corollary 1 that in the case when the customer length does not depend on his capacity and the capacity is equal to 1, we have $\varphi(s) = e^{-s}$ and

$$\overline{\delta}(s,q) = [q + a(1 - \pi(q))(1 - e^{-s})]^{-1} =$$

$$= \int_0^{\infty} e^{-qt} \mathbf{E} e^{-s\sigma(t)} dt = \int_0^{\infty} e^{-qt} P(e^{-s}, t)dt,$$

whence Eq. (15) follows if we substitute e^{-s} by z.

Corollary 5. *Let $\rho = a\beta_1 < 1$. Then stationary mode exists. The Laplace–Stieltjes transform $\delta(s)$ of the stationary distribution function $D(x) = \lim_{t\to\infty} D(x,t)$ of customers total capacity has the form:*

$$\delta(s) = \frac{1 - \rho}{1 + a\alpha_q'(s,q)|_{q=0}}. \tag{16}$$

Note that relation (16) was first obtained by Sengupta [8].

Proof. It follows from the theory of regenerative processes [6] that the limit $\delta(s) = \lim_{t\to\infty} \delta(s,t)$ exists and

$$\delta(s) = \lim_{q\to 0} q\overline{\delta}(s,q) = (1 - \rho) \lim_{q\to 0} [1 - I(s,q)]^{-1},$$

where, as it follows from theorem 1,

$$\lim_{q\to 0} I(s,q) = a\int_0^\infty [1-B(y)]e_y(s)dy =$$

$$= a\int_{x=0}^\infty \int_{u=0}^\infty ue^{-sx}dF(x,u) = -a\alpha'_q(s,q)|_{q=0},$$

whence the statement of the corollary follows.

Corollary 6. *Let $\delta_1(t)$ be the first moment of the total customers capacity $\sigma(t)$ under zero initial condition, $\overline{\delta}_1(q)$ be the Laplace transform of the function $\delta_1(t)$. Then we have:*

$$\overline{\delta}_1(q) = \frac{a\alpha_{11} + q\int_0^\infty \int_0^\infty xS(t)dF(x,t)}{[q+a-a\pi(q)]\left[1-\rho-q\int_0^\infty S(t)dB(t)\right]^2},$$

where $S(t) = \int_0^t [1-B(y)]^{-1}dy$.

Let σ be a stationary total customers capacity ($\sigma(t) \Rightarrow \sigma$ in the sense of a weak convergence). The following known formulae [8]

$$\delta_1 = \mathbf{E}\sigma = -\delta'(0) = \frac{a\alpha_{11}}{1-\rho},\ \delta_2 = \mathbf{E}\sigma^2 = \delta''(0) = \frac{a\alpha_{21}}{1-\rho} + 2\delta_1^2 \qquad (17)$$

can be obtained from relation (16).

For some special cases we can get the distribution function $D(x)$ from formula (16). For example, consider the case when customer's capacity ζ and his length ξ are connected by the relation $\xi = c\zeta + \xi_1$, $c > 0$, where the random variables ζ and ξ_1 are independent (such dependence for customer's capacity and his length is true for many real information systems).

Denote by $\kappa_1 = \mathbf{E}\xi_1$ the first moment of the random variable ξ_1. In this case we have $\alpha(s,q) = \varphi(s+cq)\kappa(s)$, where $\kappa(s)$ is the Laplace–Stieltjes transform of the distribution function of the random variable ξ_1. Then relation (16) takes the following form:

$$\delta(s) = \frac{1-\rho}{1+a[c\varphi'(s) - \kappa_1\varphi(s)]}. \qquad (18)$$

Assume that customer capacity ζ has an exponential distribution with the parameter $f > 0$. Then from formula (18) we obtain:

$$\delta(s) = \frac{(1-\rho)(s+f)^2}{(s+f)^2 - \rho_1 f^2 - \rho_2 f(s+f)},$$

where $\rho_1 = ac/f$, $\rho_2 = a\kappa_1$, so that $\rho = a\beta_1 = \rho_1 + \rho_2$.

Now we can determine the inverse Laplace transform of $\delta(s)/s$, where $\delta(s)$ is defined by formula (18), and obtain the stationary distribution function $D(x)$:

$$D(x) = 1 - \frac{(1-\rho)e^{-fx}}{2b}\left[\frac{(\rho_2+b)^2 e^{(\rho_2+b)fx/2}}{2-\rho_2-b} - \frac{(\rho_2-b)^2 e^{(\rho_2-b)fx/2}}{2-\rho_2+b}\right], \quad (19)$$

where $b = \sqrt{\rho_2^2 + 4\rho_1}$.

3. Estimation of loss characteristics

The $M/G/1-EPS$ is a system without losing of customers ($V = \infty$). But with the help of this model we can estimate the memory capacity V in order to guarantee inexceeding of given loss probability.

Assume that we have a stationary queueing system Q_∞ with Poisson entrance flow without losses of customers. Let Q_V be a stationary system that differs from Q_∞ only with the fact that its total capacity is limited by the constant value V. We denote by $D(x)$ the distribution function of total customers capacity for the system Q_∞ and by $D_V(x)$ the distribution function of this random value for the system Q_V.

Theorem 2. *The inequality $D(x) \le D_V(x)$ takes place for all $x > 0$.*

Proof of the theorem can be found in [7].

It follows from theorem 2 that the loss probability P for the system Q_V satisfies the following inequality [7]:

$$P = 1 - \int_0^V D_V(V-x)dL(x) \le 1 - \int_0^V D(V-x)dL(x) = P^*. \quad (20)$$

Thus, the value P^* is an upper estimation of loss probability for the system Q_V. If we choose V under condition that P^* is given so that the equality

$$\int_0^V D(V-x)dL(x) = 1 - P^*$$

is satisfied, then the real loss probability P does not exceed P^*. If only very rare losses are permitted in the system under consideration, the difference between the values P and P^* is inessential.

Note that the loss probability is not exhaustive characteristic of losses, because its value shows a part of lost customers, not a part of lost capacity or, in other words, information being lost. Really, it is obvious that customers having large capacity will be lost more often. Therefore, more objective losses estimation is the value

$$Q = 1 - \frac{1}{\varphi_1}\int_0^V xD_V(V-x)dL(x).$$

The value Q is the probability of losing a unit of customer capacity. The next inequality follows from theorem 2:

$$Q = 1 - \frac{1}{\varphi_1}\int_0^V xD_V(V-x)dL(x) \le 1 - \frac{1}{\varphi_1}\int_0^V xD(V-x)dL(x) = Q^*.$$

If only very rare losses are permitted in the system under consideration, the difference between the values Q and Q^* is inessential.

For example, in the case of the distribution function (19) we obtain:

$$P^* = \left\{1 - \frac{1-\rho}{b}\left[a_1\frac{1-e^{-(1-b_1)fV}}{b+\rho_2} + a_2\frac{1-e^{-(1-b_2)fV}}{b-\rho_2}\right]\right\}e^{-fV},$$

where $a_1 = \dfrac{(\rho_2+b)^2}{2-\rho_2-b}$, $a_2 = \dfrac{(\rho_2-b)^2}{2-\rho_2+b}$, $b_1 = -1+\dfrac{\rho_2+b}{2}$, $b_2 = -1+\dfrac{\rho_2-b}{2}$;

$$Q^* = \left\{1 + fV - \frac{2(1-\rho)}{b}\left[\frac{(a_1+a_2)fV}{8\rho_1} + \right.\right.$$

$$\left.\left. + a_1\frac{1-e^{-(1-b_1)fV}}{(b+\rho_2)^2} - a_2\frac{1-e^{-(1-b_2)fV}}{(b-\rho_2)^2}\right]\right\}e^{-fV}.$$

Note that in the most cases the calculation and estimation of the probability Q is very complicated. Therefore, we often must restrict ourselves to the calculation and estimation of the loss probability P.

If it is impossible to determine the form of the distribution function $D(x)$, we can estimate the value P^* by approximation of the function

$$\Phi(x) = \int_0^x D(x-u)dL(u)$$

being the distribution function of the sum of independent random variables σ and ζ, with the distribution function of the gamma distribution $\Phi^*(x) = \gamma(h, rx)/\Gamma(h)$, where $\gamma(h, rx) = \displaystyle\int_0^{hx} t^{h-1}e^{-t}dt$ is the incomplete gamma function, $\Gamma(h) = \gamma(h, \infty)$ is the gamma function. The parameters h and r of the approximate distribution should be chosen so that its first and second moments $f_1^* = h/r$ and $f_2^* = h(h+1)/r^2$ should be equal to the first and second moments of the distribution function $\Phi(x)$, respectively. It is obvious that these moments have the form

$$f_1 = \delta_1 + \varphi_1,\ f_2 = \delta_2 + \varphi_2 + 2\delta_1\varphi_1. \tag{21}$$

Thus, the parameters of the distribution function $\Phi^*(x)$ should be chosen as follows:

$$h = \frac{f_1^2}{f_2 - f_1^2}, \quad r = \frac{f_1}{f_2 - f_1^2},$$

where f_1 and f_2 can be calculated from relations (17), (21). Hence, we have the approximate formula

$$P^* \cong 1 - \Phi^*(V).$$

Note that in the case of not very small permissible loss probabilities, using the estimation P^* instead of P leads to unjustifiably surplus choice of the capacity volume V. Therefore, the direct analysis of processor sharing systems with limited memory space is very important.

4. The case of limited total capacity

The system $M/G/1(V)-EPS$ with customers of different types was analyzed in detail in [9, 10]. We shall concider a special case of customers of the same type. Then, for stationary probabilities of number of customers present in the system we have:

$$p_0 = \left(\sum_{k=0}^{\infty} a^k A_*^{(k)}(V) \right)^{-1}, \; p_k = p_0 a^k A_*^{(k)}(V), \; k = 1, 2, \ldots,$$

where $A_*^{(k)}(x)$ is a kth order Stieltjes convolution of the function

$$A(x) = \int_{u=0}^{x} \int_{t=0}^{\infty} u dF(u,t).$$

The loss probability has the form:

$$P = 1 - p_0 \left[L(V) - \sum_{k=1}^{\infty} a^k A_*^{(k)}(V) \right].$$

Assume additionally that customer capacity has an exponential distribution with parameter f, and let the customer length be proportional to his capacity ($\xi = c\zeta$, $c > 0$). Then, after some calculations we obtain

$$p_0 = \begin{cases} \dfrac{1-\rho}{1 - \sqrt{\rho} e^{-fV} \left[\sinh(\sqrt{\rho} fV) + \sqrt{\rho} \cosh(\sqrt{\rho} fV) \right]}, & \text{if } \rho \neq 1, \\ \dfrac{1 + e^{-2fV}}{1 + fV}, & \text{if } \rho = 1; \end{cases}$$

$$p_k = p_0 \rho^k \left[1 - e^{-fV} \sum_{i=0}^{2k-1} \frac{(fV)^i}{i!} \right], \ k = 1, 2, \ldots;$$

$$P = p_0 e^{-fV} \cosh(\sqrt{\rho} fV),$$

where $\rho = ac/f$.

Table 1: Probabilities P and Q for $\rho = 0.6$

V	P^*	Q^*	P	Q
0.0	1.00000	1.00000	1.00000	1.00000
0.2	0.92721	0.99569	0.81994	0.98269
0.4	0.86622	0.98366	0.67754	0.94034
0.6	0.81392	0.96529	0.56700	0.88482
0.8	0.76815	0.94194	0.48156	0.82409
1.0	0.72735	0.91487	0.41516	0.76311
2.0	0.56855	0.75562	0.23586	0.51290
3.0	0.45178	0.60242	0.15775	0.35596
4.0	0.35651	0.47628	0.11281	0.25640
5.0	0.28750	0.37679	0.08340	0.18993
6.0	0.22947	0.29888	0.06291	0.14330
7.0	0.18316	0.23763	0.04811	0.10963
8.0	0.14620	0.18925	0.03716	0.08464
10.0	0.09314	0.12034	0.02263	0.05165
15.0	0.03018	0.03896	0.00697	0.01589
20.0	0.00978	0.01262	0.00222	0.00512
30.0	0.00103	0.00133	0.00023	0.00054
40.0	0.00011	0.00014	0.00002	0.01589
50.0	0.00001	0.00002	0.00000	0.00001

Now we can compare the values P^* and P or Q^* and Q using analytical results and simulation. Table 1 presents the dependence of loss characteristics upon the memory capacity V. Here we assume that $\rho = 0.6$, the customer length is proportional to his capacity ($\xi = c\zeta$), where $c = 1$, and capacity ζ has an exponential distribution with parameter $f = 1$.

The values P^*, Q^*, P were obtained by calculation from the above relations, whereas the value Q was estimated by simulation. The table shows that estimators P^*, Q^* are not very precise, and we can use them for the case when the proper loss characteristics are near zero.

References

[1] R. Litjens, H. Van der Berg, R. J. Boucherie. Throughputs in processor sharing models for integrated stream and elastic traffic. *Performance Evaluation*, **65**, 152–180, 2008.

[2] S. F. Yashkov. *Analysis of Queues in Computers.* Radio i Svyaz, Moskow, 1989. (In Russian).

[3] S. F. Yashkov, A. S. Yashkova. Processor sharing: a survey of the mathematical theory. *Autom. Remote Control*, **68** (9), 1662–1731, 2007.

[4] L. Kleinrock. Time-shared system: a theoretical treatment. *J. Assoc. Comput. Mach.*, **14** (2), 242–251, 1967.

[5] V. F. Matveev, V. G. Ushakov. *Queueing Systems.* Moscow University, 1984. (In Russian).

[6] D. R. Cox, W. L. Smith. *Renewal Theory.* Methunen, London, 1962.

[7] O. Tikhonenko. *Metody probabilistyczne analizy systemów informacyjnych.* Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2006.

[8] B. Sengupta. The spatial requirement of an $M/G/1$ queue or: how to design for buffer space. *Lect. Notes Contr. Inf. Sci.*, **60**, 547–564, 1984.

[9] O. M. Tikhonenko. Queueing systems with processor sharing and limited resources. *Autom. Remote Control*, **71** (5), 803–815, 2010.

[10] O. Tikhonenko. Classical and non-classical processor sharing systems with non-homogeneous customers. *Scientific Issues of Jan Długosz University in Częstochowa, Ser. Mathematics*, **XIV**, 133–150, 2009.

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XV, Częstochowa 2010

SAT-BASED SEARCHING FOR k-QUASI-OPTIMAL RUNS IN WEIGHTED TIMED AUTOMATA

Bożena Woźna-Szcześniak, Andrzej Zbrzezny

Institute of Mathematics and Computer Science
Jan Długosz University of Częstochowa
al. Armii Krajowej 13/15, 42-200 Częstochowa, Poland
e-mail: {b.wozna, a.zbrzezny}@ajd.czest.pl

Abstract. In the paper we are concerned with an optimal cost reachability problem for weighted timed automata, and we use a translation to SAT to solve the problem. In particular, we show how to find a run of length $k \in \mathbb{N}$ that starts at the initial state and terminates at a state containing a target location, its total cost belongs to the interval $[c, c+1)$, for some natural number $c \in \mathbb{N}$, and the cost of each other run of length k, which also leads from the initial state to a state containing the target location, is greater or equal to c. This kind of runs is called *k-quasi-optimal*. We exemplify the use of our solution to the mentioned problem by means of the *air traffic control problem*, and we provide some preliminary experimental results.

1. Introduction

In automatic verification of hardware and software systems, the *reachability* problem is a core decision problem. This is because it can be used to detect *deadlocks*, or a violation of a *safety* property, which means that nothing bad will ever happen. For real-time systems like, for example, an air traffic control, or process controllers in manufacturing plants, it is also reasonable to ask questions about the minimum cost of reaching a desirable state of the system. Therefore, in the paper, we deal with the *k-optimal cost reachability* problem for *weighted timed automata* [3], in particular, we are interested in using SAT-methods to solve the problem.

A timed automaton [2] is a formalism that can be used to model the behaviour of a real-time system. It extends a finite automaton by adding a finite

set of variables that are able to measure real-time, and express timing constraints; these variables are called clocks. The semantics of a timed automaton is given in terms of an infinite labelled transition system with two kinds of transitions: a discrete transition and a time transition. The first one correspond to a change of a location, and the second one to the passage of time. However, in order to define the *k-optimal cost reachability* problem for timed automata we need to associate costs with transitions and locations. The costs assigned to transitions (switch costs) will give the cost of discrete transitions, and the costs assigned to locations (duration costs) will define the cost of time spent in these locations. Such timed automata augmented with costs are known as *weighted timed automata* [3], or *priced timed automata* [5].

Our solution to the k-optimal cost reachability problem relies on combining the well-know *forward reachability* analysis and the *bounded model checking* (BMC) method [6, 13, 14]. The forward reachability algorithm searches the state space using the breadth first mode, whereas the BMC performs a verification on a part of the automata model exploiting SAT solvers.

The rest of the paper is organised as follows. In the next section we provide the main formalisms used throughout the paper, i.e. weighted timed automata. In Section 3 we define and solve the k-optimal cost reachability problem for weighted timed automata. In Section 4 we show how our solution to the considered reachability problem works by means of the *air traffic control problem*. We conclude in Section 5 by discussing related work.

2. Weighted Timed Automata

Let us start by fixing names of the sets of numbers used in the rest of the paper. By $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ we denote the set of natural numbers, by $\mathbb{Q}$ the set of non-negative rational numbers, and by PV a set of propositional variables.

To define weighted timed automata formally, we need to say what type of clock constraints are allowed as guards and invariants, and what are the cost functions. This is introduced in the following subsection.

2.1. Clocks and clock valuation

For a finite set $\mathcal{X}$ of real variables, called *clocks*, the set $\mathcal{C}(\mathcal{X})$ of all the *clock constraints* over $\mathcal{X}$ is defined by the following grammar:

$$\mathbf{cc} ::= \mathit{true} \mid x \sim c \mid x - y \sim c \mid \mathbf{cc} \wedge \mathbf{cc},$$

where $x, y \in \mathcal{X}$, $c \in \mathbb{N}$, and $\sim \in \{\leq, <, =, >, \geq\}$.

A *clock valuation* is a total mapping $\mathbf{c} : \mathcal{X} \to \mathbb{Q}$. Satisfiability of a clock constraint $\mathbf{cc} \in \mathcal{C}(\mathcal{X})$ by a clock valuation $\mathbf{c}$ ($\mathbf{c} \models \mathbf{cc}$) is defined inductively as follows:

- $\mathbf{c} \models \mathit{true}$,
- $\mathbf{c} \models (x \sim c)$ iff $\mathbf{c}(x) \models c$,
- $\mathbf{c} \models (x - y \sim c)$ iff $\mathbf{c}(x) - \mathbf{c}(y) \sim c$,
- $\mathbf{c} \models \mathbf{cc}_1 \wedge \mathbf{cc}_2$ iff $\mathbf{c} \models \mathbf{cc}_1$ and $\mathbf{c} \models \mathbf{cc}_2$.

In what follows, the set of all the clock valuations satisfying a clock constraint $\mathbf{cc}$ is denoted by $[\![\mathbf{cc}]\!]$. Given a clock valuation $\mathbf{c}$ and $\delta \in \mathbb{Q}$, by $\mathbf{c}+\delta$ we denote a clock valuation $\mathbf{c}'$ such that $\mathbf{c}'(x) = \mathbf{c}(x) + \delta$, for all $x \in \mathcal{X}$. Moreover, for a subset of clocks $X \subseteq \mathcal{X}$, $\mathbf{c}[X := 0]$ denotes the valuation $\mathbf{c}'$ such that for all $x \in X$, $\mathbf{c}'(x) = 0$ and for all $x \in \mathcal{X} \setminus X$, $\mathbf{c}'(x) = \mathbf{c}(x)$. Finally, by $\mathbf{c}^0$ we denote the *initial* clock valuation, i.e. the valuation such that $\mathbf{c}^0(x) = 0$ for all $x \in \mathcal{X}$.

2.2. Syntax and semantics

We assume the definition of weighted timed automata from [3] but augmented to include a special rational variable z.

Definition 1. (Weighted timed automaton). *A* weighted timed automaton *is a tuple* $\mathcal{A} = (\Sigma, L, l^0, \mathcal{X}, E, \mathcal{I}, J_s, J_d, z, \mathcal{V})$, *where* Σ *is a finite set of labels (actions),* L *is a finite set of locations,* l^0 *is an initial location,* $\mathcal{X}$ *is a finite set of clocks,* $E \subseteq L \times \Sigma \times \mathcal{C}(\mathcal{X}) \times 2^{\mathcal{X}} \times L$ *is a transition relation,* $\mathcal{I} : L \to \mathcal{C}(\mathcal{X})$ *is an invariant function,* $J_s : E \to \mathbb{N}$ *is a switch cost function,* $J_d : L \to \mathbb{N}$ *is a duration cost function,* z *is a rational variable, and* $\mathcal{V} : L \to 2^{PV}$ *is a valuation function assigning to each location a set of atomic propositions true in that location.*

The switch cost function assigns to each transition a cost expressing the price of taking the transition. The duration cost function assigns to each location a cost expressing the price of staying in this location for one time unit. The invariant function assigns to each location a clock constraint expressing the condition under which $\mathcal{A}$ can stay in this location. Each element $t = (l, \sigma, \mathbf{cc}, X, l') \in E$ represents a transition from the location l to the location l', where σ is the label of the transition t, $\mathbf{cc}$ defines the enabling conditions for t, and X is a set of clocks to be reset.

The semantics of the weighted timed automaton is defined by associating to it a *dense model* as defined below.

Definition 2. *Let* $\mathcal{A} = (\Sigma, L, l^0, \mathcal{X}, E, \mathcal{I}, J_s, J_d, z, \mathcal{V})$ *be a weighted timed automaton,* $\mathbf{z} : \{z\} \to \mathbb{Q}$ *a* valuation for z, *and* $\mathbf{z}^0$ *denote the* initial *valuation for* z, *i.e.,* $\mathbf{z}^0(z) = 0$. *A* dense model *for* $\mathcal{A}$ *is a tuple* $M(\mathcal{A}) = (\Sigma \cup \mathbb{Q}, S, s^0, \to, \mathcal{V}')$, *where* $\Sigma \cup \mathbb{Q}$ *is a set of labels,* $S = \{(l, \mathbf{c}, \mathbf{z}) \mid l \in L,\ \mathbf{c} \in \mathbb{Q}^{|\mathcal{X}|},\ \mathbf{c} \models \mathcal{I}(l), \mathbf{z} \in \mathbb{Q}\}$ *is a set of states,* $s^0 = (l^0, \mathbf{c}^0, \mathbf{z}^0)$ *is the initial state,* $\mathcal{V}' : S \to 2^{PV}$ *is a valuation function such that* $\mathcal{V}'((l, \mathbf{c}, \mathbf{z})) = \mathcal{V}(l)$, *and* $\to \subseteq S \times \Sigma \cup \mathbb{Q} \times S$ *is the smallest transition relation defined by the following rules:*

- *for* $\sigma \in \Sigma$, $(l, \mathbf{c}, \mathbf{z}) \xrightarrow{\sigma} (l', \mathbf{c}', \mathbf{z}')$ *iff there exists a transition* $t = (l, \sigma, \mathbf{cc}, X, l') \in E$ *such that* $\mathbf{c} \models \mathbf{cc}$, $\mathbf{c} \models \mathcal{I}(l)$, $\mathbf{c}[X := 0] \models \mathcal{I}(l')$, *and* $\mathbf{z}' = \mathbf{z} + J_s(t)$ (action transition),
- *for* $\delta \in \mathbb{Q}$, $(l, \mathbf{c}, \mathbf{z}) \xrightarrow{\delta} (l, \mathbf{c} + \delta, \mathbf{z}')$ *iff* $\mathbf{c}, \mathbf{c} + \delta \models \mathcal{I}(l)$, *and* $\mathbf{z}' = \mathbf{z} + J_d(l) \cdot \delta$ (time transition).

Intuitively, an action transition corresponds to an action performed by the automaton under consideration. The action can be performed only if the underlying enabling condition is satisfied. Moreover, all the clocks that are associated with the action are set to zero, its locations change accordingly, and the value of the variable z is increased by the switch cost. A time transition causes an equal increase in the value of all the clocks, and does not involve a location change. Obviously, the new clock valuations have to still satisfy all the location invariants, and the value of the variable z is increased by the duration cost.

Let $(l, \mathbf{c}, \mathbf{z}) \xrightarrow{\delta,\sigma} (l', \mathbf{c}', \mathbf{z}')$ denote that $(l, \mathbf{c}, \mathbf{z}) \xrightarrow{\delta} (l'', \mathbf{c}'', \mathbf{z}'')$ and $(l'', \mathbf{c}'', \mathbf{z}'') \xrightarrow{\sigma} (l', \mathbf{c}', \mathbf{z}')$, where $\sigma \in \Sigma$ and $\delta \in \mathbb{Q}$. A *run* ρ of a weighted timed automaton $\mathcal{A}$ is a finite sequence of states:

$$(l_0, \mathbf{c}_0, \mathbf{z}_0) \xrightarrow{\delta_1,\sigma_1} (l_1, \mathbf{c}_1, \mathbf{z}_1) \xrightarrow{\delta_2,\sigma_2} \dots \xrightarrow{\delta_{k-1},\sigma_{k-1}} (l_{k-1}, \mathbf{c}_{k-1}, \mathbf{z}_{k-1}) \xrightarrow{\delta_k,\sigma_k} (l_k, \mathbf{c}_k, \mathbf{z}_k)$$

such that $(l_i, \mathbf{c}_i, \mathbf{z}_i) \in S$, $\sigma_i \in \Sigma$, and $\delta_i \in \mathbb{Q}$ for each $i \in \{1, \dots, k\}$. Hereafter, we refer to a run ρ of length k as k-run.

Given a k-run ρ of $\mathcal{A}$ and cost functions J_s and J_d, we associate cost to ρ as follows:

- $J_s(\rho) = \sum_{i=0}^{k-1} J_s(t_i)$, where $t_i := (l_i, \mathbf{c}_i, \mathbf{z}_i) \xrightarrow{\delta_{i+1},\sigma_{i+1}} (l_{i+1}, \mathbf{c}_{i+1}, \mathbf{z}_{i+1})$,
- $J_d(\rho) = \sum_{i=1}^{k} \delta_i \cdot J_d(l_i)$.

The *total cost* associated to a k-run ρ is defined as $J(\rho) = J_d(\rho) + J_s(\rho)$.

The *k-optimal cost* for k-runs that start at a state containing location l and end at a state containing location l' is defined as $J_k^*(l, l') = inf\{J(\rho) | \rho$ is a k-run from a state containing location l to a state containing location $l'\}$.

A k-run ρ from a state containing location l to a state containing location l' such that $\lfloor J(\rho) \rfloor = \lfloor J_k^*(l, l') \rfloor$ is called *k-quasi-optimal*.

In this paper, for given two locations l and l' we are interested in finding the greatest integer lower bound (g.i.l.b. for short) of the k-optimal cost for k-runs starting at a state s containing location l and terminating at a state t containing location l', where k is the length of a shortest run from s to t. Moreover, we are interested in finding *k-quasi-optimal* runs. Therefore, in Section 3 we define k-optimal cost reachability problem, and we show how to solve it using SAT-methods.

2.3. Discrete semantics

In real-time systems modeled by (weighted) timed automata, in order to use SAT-techniques to test reachability or other properties, it is customary to discretise the set of all the clocks valuations. Here we take the discretisation scheme that is based on the one introduced in [15], but here we use the discretisation step that depends not only on the length of considered runs, but also on the maximal duration cost. It uses the following set of discretised clock's values and labels as primitives. Let c_{max} be the largest constant c appearing in all the invariants and guards of a weighted timed automaton $\mathcal{A}$. For every $m \in \mathbb{N}$ we define $A_m = \{a \in \mathbb{Q} \mid (\exists j \in \mathbb{N})\ a \cdot 2^m = j\}$ and $B_m = \{b \in \mathbb{Q} \mid (\exists j \in \mathbb{N})\ b \cdot 2^m = j \text{ and } b < c_{max} + 1\}$. Then, $A = \bigcup_{m=0}^{\infty} A_m$ defines the set of discretised clock's values, and $B = \bigcup_{m=1}^{\infty} B_m$ defines the set of labels. We use this technique to define a *discretised model* for a weighted timed automaton. This model is crucial for the translation of the k-optimal cost reachability problem to the SAT-problem as described in the next section.

To give a definition of a discretised model that supports clock constraints of the form $x - y \sim c$, we first recall the notion of *weak region equivalence* [15].

Definition 3. (Weak region equivalence). *Assume a set of clocks $\mathcal{X}$, and for any $t \in \mathbb{Q}$ let $\langle t \rangle$ denote the fractional (respectively integral) part of t (respectively $\lfloor t \rfloor$). The* weak region equivalence *is a relation $\cong \subseteq \mathbb{Q}^{\mathcal{X}} \times \mathbb{Q}^{\mathcal{X}}$ defined as follows. For two clock valuations u and v in $\mathbb{Q}^{\mathcal{X}}$, $u \cong v$ iff all the following conditions hold:*

[1] $\lfloor u(x) \rfloor = \lfloor v(x) \rfloor$, *for all* $x \in \mathcal{X}$.

[2] $\langle u(x) \rangle = 0$ *iff* $\langle v(x) \rangle = 0$, *for all* $x \in \mathcal{X}$.

[3] $\langle u(x) \rangle < \langle u(y) \rangle$ *iff* $\langle v(x) \rangle < \langle v(y) \rangle$, *for all* $x, y \in \mathcal{X}$.

Definition 4. (Discretised model). *Let $\mathcal{A} = (\Sigma, L, l^0, \mathcal{X}, E, \mathcal{I}, J_s, J_d, z, \mathcal{V})$ be a weighted timed automaton. A* discretised model *for $\mathcal{A}$ is a tuple*

$M_d(\mathcal{A}) = (\Sigma \cup B, S_d, s_d^0, \rightarrow_d, \mathcal{V}_d)$, where $S_d = L \times A^{\mathcal{X}} \times B$ is a set of states, $s_d^0 = (l^0, \mathbf{c}^0, \mathbf{z}^0)$ is the initial state, $\mathcal{V}_d : S_d \rightarrow 2^{PV}$ is a valuation function defined by $\mathcal{V}_d((l, \mathbf{c}, \mathbf{z})) = \mathcal{V}(l)$, and $\rightarrow_d \subseteq S_d \times (\Sigma \cup B) \times S_d$ is a time/action transition relation defined by:

- *Time transition: for any $\delta \in B$, $(l, \mathbf{c}, \mathbf{z}) \xrightarrow{\delta}_d (l, \mathbf{c} + \delta, \mathbf{z}')$ iff $(l, \mathbf{c}, \mathbf{z}) \xrightarrow{\delta} (l, \mathbf{c} + \delta, \mathbf{z}')$ in $M(\mathcal{A})$ and $(\forall \delta' \leq \delta)$ $\mathbf{c} + \delta' \cong \mathbf{c}$ or $\mathbf{c} + \delta' \cong \mathbf{c} + \delta$,*
- *Action transition: for any $\sigma \in \Sigma$, $(l, \mathbf{c}, \mathbf{z}) \xrightarrow{\sigma}_d (l', \mathbf{c}', \mathbf{z}')$ iff $(l, \mathbf{c}, \mathbf{z}) \xrightarrow{\sigma} (l', \mathbf{c}', \mathbf{z}')$ in $M(\mathcal{A})$.*

The theorem below shows that the k-optimal cost reachability problem for a weighted timed automaton $\mathcal{A}$ can be solved using the discretised model $M_d(\mathcal{A})$ instead of the dense model $M(\mathcal{A})$.

In what follows, we denote by $\rho(s,t)$ a run that starts at state s and ends at state t. Moreover, for two states $s = (l, \mathbf{c}, \mathbf{z})$ and $t = (l', \mathbf{c}', \mathbf{z}')$, we write $s \cong t$ if and only if $l = l'$, $\mathbf{c} \cong \mathbf{c}'$, $\lfloor \mathbf{z}(z) \rfloor = \lfloor \mathbf{z}'(z) \rfloor$ and $\langle \mathbf{z}(z) \rangle = 0 \iff \langle \mathbf{z}'(z) \rangle = 0$.

Theorem 1. *Let $\mathcal{A}$ be a weighted timed automaton, s and t two states in $M(\mathcal{A})$, and $\rho(s,t)$ a k-quasi-optimal run in $M(\mathcal{A})$, where $k \in \mathbb{N}$ is the length of a shortest run that starts at s and ends at t. Then, there exist two states s' and t' in $M_d(\mathcal{A})$ and there exists a k-quasi-optimal run $\rho'(s', t')$ in $M_d(\mathcal{A})$ such that $s \cong s'$ and $t \cong t'$.*

Proof (Idea). The proof is an extension of the proof of Theorem 3.1 in [15], and it is conducted by means of induction on k. The induction step consists in showing that for each $q = (l, \mathbf{c}_q, \mathbf{z}_q), r = (l', \mathbf{c}_r, \mathbf{z}_r) \in M(\mathcal{A})$, $q' = (l, \mathbf{c}_{q'}, \mathbf{z}_{q'}), r' = (l', \mathbf{c}_{r'}, \mathbf{z}_{r'}) \in M_d(\mathcal{A})$, $\delta \in \mathbb{Q}$, $\delta' \in B$, if $q \cong q'$, $\delta \cong \delta'$ and there exist transitions $q \xrightarrow{\delta,\sigma} r$, $q' \xrightarrow{\delta',\sigma} r'$, then $r \cong r'$. The crucial part of the induction step is rather tedious, and relies on showing that $\mathbf{z}_q + J_d(l) \cdot \delta \cong \mathbf{z}_{q'} + J_d(l) \cdot \delta'$, what requires using some techniqal facts concerning the underling discretisation.

3. k-optimal cost reachability problem

In this section we formally define the k-optimal cost reachability problem for weighted timed automata, and we present a solution to the problem which uses SAT-solvers. We start by defining the problem, then we describe our solution informally, and finaly we show our algorithm.

The k-optimal cost reachability problem for weighted timed automata is defined as follows.

Definition 5. (k-optimal cost reachability). *Given a weighted timed automaton $\mathcal{A} = (\Sigma, L, l^0, \mathcal{X}, E, \mathcal{I}, J_s, J_d, z, \mathcal{V})$, and a desirable location $l^p \in L$ satisfying a property p.* k-optimal cost reachability *consists in finding a k-quasi-optimal run ρ starting at $s_d^0 \in M_d(\mathcal{A})$ and terminating at a state in $M_d(\mathcal{A})$ containing location l^p.*

Note that if ρ is a k-quasi-optimal run, then there exists $c \in \mathbb{N}$ such that: $c \leq J(\rho) < c + 1$, and for all the $k-$runs ρ' that starts at s_d^0 and terminates at a state in $M_d(\mathcal{A})$ containing location l^p, $J(\rho') \geq c$ holds.

3.1. Our solution – an informal explanation

We begin with an informal explanation of our solution to the k-optimal cost reachability problem, which will help to understand the formal description presented later on in this section.

To solve the k-optimal cost reachability problem we proceed as follows. We first encode by propositional formulae both the property p, and the unfolding of the transition relation of $M_d(\mathcal{A})$ up to the depth k (for $k \in \mathbb{N}$). Let φ_k be the conjunction of the two above formulae. We test φ_k for the propositional satisfiability using a SAT-solver. If the test for φ_k is positive, we calculate the cost $r_0 \in \mathbb{Q}$ of the resulting witness ρ_0, and we know that $J(\rho_0) < \lceil r_0 \rceil$. Next, we set $c_0 = \lceil r_0 \rceil - 1$, and we run the propositional satisfiability test once again, but for the formula $\phi_k(c_0) = \varphi_k \wedge (z < c_0)$[1]. If the test for $\phi_k(c_0)$ is positive, we calculate the cost $r_1 \in \mathbb{Q}$ of the resulting witness ρ_1, and we know that $r_1 < c_0$. Next, we set $c_1 = \lceil r_1 \rceil - 1$, and we run the propositional satisfiability test once again, but for the formula $\phi_k(c_1) = \varphi_k \wedge (z < c_1)$, and so on. We stop testing if the test for $\phi_k(c_i)$ is negative or $r_i = 0$.

Notice, that if the test for $\phi_k(c_i)$ is negative, we can perform one more test for the formula $\psi_k(c_i) = \varphi_k \wedge (z = c_i)$. If the test for $\psi_k(c_i)$ is positive, we can conclude that k-optimal cost is equal to c_i. Otherwise, we can only conclude that the g.i.l.b. of the k-optimal cost is equal to c_i.

3.2. Translation to propositional formulae

Let $\mathcal{A} = (\Sigma, L, l^0, \mathcal{X}, E, \mathcal{I}, J_s, J_d, z, \mathcal{V})$ be a weighted timed automaton, $M_d(\mathcal{A}) = (\Sigma \cup B, S_d, s_d^0, \rightarrow_d, \mathcal{V}_d)$ a discretised model, and $k \in \mathbb{N}$. Each state s of $M_d(\mathcal{A})$ reachable on a k-run can be encoded by a bit-vector whose length, say n, depends on the number of locations, the constant c_{max}, the maximal duration cost, and the number k. Thus, each state s of $M_d(\mathcal{A})$ can be represented by a vector $\mathtt{w} = (\mathtt{w}[1], \ldots, \mathtt{w}[n])$ of propositional variables (usually

[1] The notation $z < c_i$, for $i = 0, 1, 2, \ldots$ appearing in this section, denotes a propositional formula encoding the fact that the value of the variable z is less than c_i.

called *state variables*) to which we refer to as a *global state variable* [2]. A finite sequence $(\mathtt{w}_0, \ldots, \mathtt{w}_k)$ of global state variables is called a *symbolic k-path*.

For two global state variables $\mathtt{w}, \mathtt{w}'$, we define the following propositional formulae:

- $I_s(\mathtt{w})$ is a formula over $\mathtt{w}$ that is true for a valuation s_w of $\mathtt{w}$ iff $s_w = s$.
- $p(\mathtt{w})$ is a formula over $\mathtt{w}$ that is true for a valuation s_w of $\mathtt{w}$ iff $p \in \mathcal{V}(s_w)$ (encodes a set of states of $M_d(\mathcal{A})$ in which $p \in PV$ holds).
- $T(\mathtt{w}, \mathtt{w}')$ is a formula over $\mathtt{w}$ and $\mathtt{w}'$ that is true for two valuations s_w of $\mathtt{w}$ and $s_{w'}$ of $\mathtt{w}'$ iff $(s_w, s_{w'}) \in \rightarrow_d$ (encodes the transition relation of $M_d(\mathcal{A})$).

  The definition of the formula T involves the Boolean encoding of addition and multiplication of rational numbers, which has been described in [16].

We can now define the propositional formula φ_k, introduced in Subsection 3.1. As it was mentioned in Subsection 3.1, φ_k is a conjunction of two formulae. The first one, denoted by $p(\mathtt{w})$, is a translation of a propositional variable p that represents a location in question. The second one, denoted by $[M_d^{s_d^0}]_k$, encodes the unfolding of the transition relation of $M_d(\mathcal{A})$ up to depth $k \in \mathbb{N}$.

The formula $[M_d^{s_d^0}]_k$ is defined over global state variables w_i, for $0 \leq i \leq k$, and it constrains the symbolic k-path to be valid k-run of $M_d(\mathcal{A})$. Namely,

$$[M_d^{s_d^0}]_k := I_{s_d^0}(w_0) \wedge \forall_{i=0}^{k-1} T(w_i, w_{i+1})$$

3.3. Our solution – a formal algorithm

Now we give an algorithm that formalises the method for finding the greatest integer lower bound of k-optimal cost informally described above.

In Algorithm 1 we use the procedure $checkSAT(\gamma)$ that for any given propositional formula γ returns a pair (X, W), where W denotes the valuation returned by a SAT solver, and X can be one of the following three values: $TRUE$, $FALSE$, and $UNKNOWN$. The meanings of the values $TRUE$ and $FALSE$ are self-evident. The value $UNKNOWN$ is returned either if the procedure $checkSAT$ is not able to decide satisfiability of its argument

[2] Notice that we distinguish between states s encoded as sequences of 0's and 1's and their representations in terms of propositional variables $w[i]$.

within some preset timeout period, or has to terminate itself due to exhaustion of available memory. We also use the procedure $getCOST(W)$ that for the valuation W, which represents a $k-$run ρ, returns a natural number c such that the cost of ρ is less than c. Further, for a given propositional formula φ_k, we denote by $\phi_k(c)$ the formula $\varphi_k \wedge (z < c)$, and by $\psi_k(c)$ the formula $\varphi_k \wedge (z = c)$.

Algorithm 1 An algorithm for finding g.i.l.b. of k-optimal cost

1: $k \leftarrow 0$
2: **repeat**
3: $(result, W) \leftarrow checkSAT(\varphi_k)$
4: **if** $result = FALSE$ **then**
5: $k \leftarrow k + 1$
6: **else if** $result = UNKNOWN$ **then**
7: **return** $UNKNOWN$
8: **end if**
9: **until** $result = TRUE$
{there exists a witness of the length k for a desirable property}
10: $c \leftarrow getCOST(W)$
11: **repeat**
12: **if** $c = 0$ **then**
13: **return** k-optimal cost is equal to 0
14: **end if**
15: $(result, W) \leftarrow checkSAT(\phi_k(c-1))$
16: **if** $result = TRUE$ **then**
17: $c \leftarrow getCOST(W)$
18: **else if** $result = UNKNOWN$ **then**
19: **return** $UNKNOWN$
20: **end if**
21: **untill** $result = FALSE$
{optimal cost of any k-run is greater or equal to c}
22: $(result, W) \leftarrow checkSAT(\psi_k(c))$
23: **if** $result = TRUE$ **then**
24: **return** k-optimal cost is equal to c
25: **else**
26: **return** g.i.l.b. of k-optimal cost is equal to c
27: **end if**

4. Case study

4.1. An air traffic control problem

Weighted timed automata with the rational variable are suitable formalism for modelling several optimisation problems, for example, scheduling problems or air traffic control problems. In this section we take a closer look at the later problem.

Assume a situation in which two aircrafts send a landing request to an airport, and they are approaching the same runway. The goal is to allow both the aircrafts to land safely and at minimum cost. Safety requires that only one aircraft at a time must be acknowledged for landing, thus there are two possible choices: aircraft 1 waits for the landing of aircraft 2 to be completed, or vice versa. This waiting can be implemented either by slowing down an aircraft (this concerns a situation, in which the aircrafts share the same trajectory, and the aircraft that is following is faster), or by forcing one of them to change its trajectory (this concerns a situation, in which the aircrafts reach the joining point of their trajectories almost at the same time).

Consider the automaton in Figure 1 [3]. It models the above scenario, i.e. the discrete values c_1 and c_2 are the costs of the choice of forcing, respectively, aircraft 1 and aircraft 2 to wait. These costs label the transitions, respectively, from location *Start* to location W_1, and from location *Start* to location W_2. The cost w_i, attached to location W_i, is related to the time spent on waiting by aircraft i. For the aircraft that has to wait for the clearance, we model two possible manoeuvres. A first one is to reduce the speed, and in this case the aircraft stays in location W_i. Another possibility is to change the original trajectory, which is modelled by the loop trough location W_i'. Doing this manoeuvre requires a fixed cost c_i', takes at least one time unit, and allows to pay w_i' instead of w_i per each time unit. Since it is realistic to reduce the time a runway stays unused, we penalise this event by a cost c_0 per time unit. Finally, we assume that the landing of each aircraft takes at least one time unit since the related acknowledgement was issued by the control tower.

4.2. Experimental results

All of the experiments have been performed on a computer equipped with the processor Intel Core 2 Duo (2 GHz), 2 GB main memory and the operating system Linux.

In Tables 1 and 2 we present experimental results for the air traffic control problem modeled by the automaton on Figure 1 with the following costs: $c_0 = 20$, $w_2 = 20$, $w_2' = 40$, $w_1 = 60$, $w_1' = 40$, $c_1 = 20$, $c_1' = 20$, $c_2 = 20$, $c_2' = 20$; we refer to this automaton as Automaton 1.

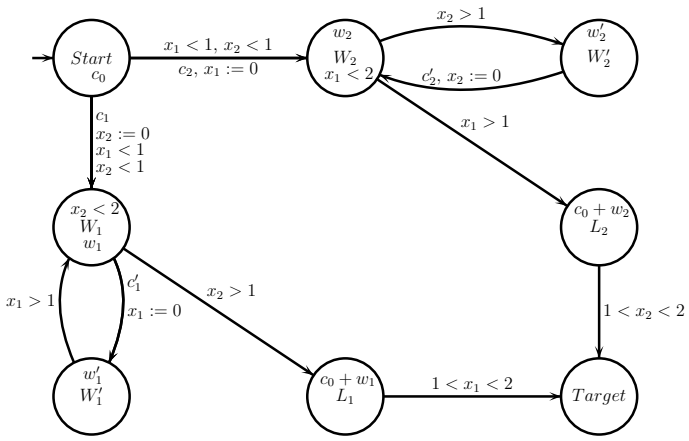


Figure 1: A weighted timed automaton for an air traffic control problem.

Table 1 shows how we get a shortest run of Automaton 1 that leads from the initial state $s_d^0 = (Start, < 0,0 >, 0)$ to a state containing the location $Target$. The cost of the 6-run is equal to $64\frac{12}{512}$, i.e., is less or equal than 65. Table 2 shows how we get the 6-quasi-optimal run of Automaton 1 such that the g.i.l.b. of 6-optimal cost is equal to 40. Table 3 shows the 6-quasi-optimal run of Automaton 1 that leads from the initial state to a state containing the location $Target$ with the g.i.l.b. of 6-optimal cost equal to 40.

BMC4WTA					**RSat**		
k	**variables**	**clauses**	**sec**	**MB**	**sec**	**MB**	**satisfiable**
0	133	190	0.00	1.9	0.0	1.3	NO
2	2278	6881	0.12	2.4	0.0	1.8	NO
4	4901	15057	0.12	3.1	0.0	2.4	NO
6	7275	22480	0.19	3.7	0.0	2.9	YES

Table 1: The shortest run of Automaton 1 on which the $Target$ location is reachable. Its total cost is $64\frac{12}{512}$.

5. Conclusions and related work

In this paper we have defined the k-optimal cost reachability problem for weighted timed automata, and presented a SAT-based method consisting in reducing this problem to the SAT-problem. In particular, we have shown how to find a k-quasi-optimal run that starts at the initial state and terminates at a desirable target state, and how to calculate g.i.l.b. of k-optimal cost for it.

Experimental results, which we have performed, show that the proposed algorithm can be very useful in finding g.i.l.b. of k-optimal cost. Obviously,

our method allows for finding only lower and upper bounds on the cost, to which the k-quasi-optimal run belongs (an unit interval $[c, c+1)$, for $c \in \mathbb{N}$), but in many real-time settings such a cost optimal approximation is sufficient.

BMC4WTA						**RSat**		
$z < c$	**cost**	**variables**	**clauses**	**sec**	**MB**	**sec**	**MB**	**satisfiable**
$z < 64$	$59\frac{132}{512}$	8194	25412	0.22	3.8	0.0	3.2	YES
$z < 59$	$48\frac{444}{512}$	8304	25784	0.21	3.9	0.1	3.2	YES
$z < 48$	$46\frac{508}{512}$	8238	25558	0.22	3.9	0.0	3.2	YES
$z < 46$	$43\frac{504}{512}$	8297	25756	0.22	3.9	0.0	3.2	YES
$z < 43$	$42\frac{36}{512}$	8318	25826	0.25	3.9	0.0	3.2	YES
$z < 42$	$41\frac{368}{512}$	8311	25798	0.22	3.9	0.0	3.2	YES
$z < 41$	$40\frac{480}{512}$	8339	25889	0.22	3.9	0.0	3.2	YES
$z < 40$	-	8267	25652	0.22	3.9	0.0	3.2	NO
$z = 40$	-	7552	23311	0.20	3.7	0.0	3.0	NO

Table 2: Searching for 6-quasi-optimal run of Automaton 1 that leads from the initial state to a state containing the location $Target$. The g.i.l.b. of 6-optimal cost is equal to 40.

k:	**location**	value of **z**	**delay**	values of **x1**, **x2**
0:	$Start$	$< 0 + \frac{0}{512} >$	$< 0 + \frac{0}{512} >$	$< 0 + \frac{0}{512}, 0 + \frac{0}{512} >$
1:	$Start$	$< 0 + \frac{20}{512} >$	$< 0 + \frac{1}{512} >$	$< 0 + \frac{1}{512}, 0 + \frac{1}{512} >$
2:	W_2	$< 20 + \frac{20}{512} >$	$< 0 + \frac{0}{512} >$	$< 0 + \frac{0}{512}, 0 + \frac{1}{512} >$
3:	W_2	$< 40 + \frac{480}{512} >$	$< 1 + \frac{23}{512} >$	$< 1 + \frac{23}{512}, 1 + \frac{24}{512} >$
4:	L_2	$< 40 + \frac{480}{512} >$	$< 0 + \frac{0}{512} >$	$< 1 + \frac{23}{512}, 1 + \frac{24}{512} >$
5:	L_2	$< 40 + \frac{480}{512} >$	$< 0 + \frac{0}{512} >$	$< 1 + \frac{23}{512}, 1 + \frac{24}{512} >$
6:	$Target$	$< 40 + \frac{480}{512} >$	$< 0 + \frac{0}{512} >$	$< 1 + \frac{23}{512}, 1 + \frac{27}{512} >$

Table 3: A 6-quasi-optimal run of Automaton 1 leading to a state containing the location $Target$.

The optimal reachability problem was considered by many researchers and several approaches treating the problem in the context of timed or hybrid automata have been described in the literature, but none of them used SAT-methods. In particular, in [9] the problem of computing lower and upper bounds on time delays in timed automata was addressed. In [1] a *duration-bounded reachability* problem for timed automata augmented to include the duration cost function is considered. This problem asks if there is a run of the timed automaton from the initial state to the given final state such that the duration of the run satisfies an arithmetic constraint (an optimal cost).

The *duration-bounded reachability* problem has been also analysed in [10]. This is because the problem can be reduced to checking whether a duration formula, which defines an optimal cost, is satisfied by a integer computation of an integration graph (a kind of a timed automaton). The solution is based on constructing a set of equations that characterises the length of time a computation spends in each automaton location.

The work [4] also tackles the optimal (minimum-time) reachability problem for timed automata. In particular, here, the problem is formulated in terms of a timed game automaton (TGA), and solved by constructing an optimal strategy using a backward fixed-point calculation on the state-space of the TGA. Minimum-time reachability problem for timed automata is also solved in [12]. However here, the solution is based on the forward fixed-point algorithm that generates on-the-fly a forward reachability graph for a given timed automaton.

The paper [5] introduces *priced timed automata* as an extension of timed automata with prices on both transitions and locations, and shows how to solve the minimum cost reachability problem; this sort of automata we have used in the paper. In [3] such reachability problem is called as the single-source optimal reachability problem, and it is solved by a reduction of the problem to a parametric shortest-path problem. The methods presented in both papers [5] and [3] are based on clock region graphs; in [3] the authors refer to priced timed automata as weighted timed automata.

Further, the paper [7] addresses the optimal reachability problem for weighted timed automata with cost functions allowing for both positive and negative costs on edges and locations, and apply the proposed method to timed games. In [11] the decidability of the optimal (minimum and maximum cost) reachability problems for multi-priced timed automata (an extension of timed automata with multiple cost variables evolving according to given rates for each location) is proved, and in [8] cost-optimal infinite schedules in terms of minimal (or maximal) cost per time ratio in the limit is considered.

References

[1] R. Alur, C. Courcoubetis, T. Henzinger. Computing accumulated delays in real-time systems. *Formal Methods in System Design*, **11** (2), 137–155, 1997.

[2] R. Alur, D. Dill. A theory of Timed Automata. *Theoretical Computer Science*, **126** (2), 183–235, 1994.

[3] R. Alur, S. La Torre, G. J. Pappas. Optimal paths in weighted timed automata. *Theoretical Computer Science*, **318** (3), 297–322, 2004.

[4] E. Asarin, O. Maler. As soon as possible: Time optimal control for timed automata. In: *Proc. 2nd Int. Workshop on Hybrid Systems: Computation and Control*, vol. 1569 of *LNCS*, pp. 19–30. Springer, 1999.

[5] G. Behrmann, A. Fehnker, T.S. Hune, K. G. Larsen, P. Pettersson, J. M. T. Romijn, F.W. Vaandrager, F. W. Va, G. Behrmann, A. Fehnker, T. Hune, K. Larsen, P. Pettersson, J. Romijn. Minimum-cost reachability for priced timed automata. In: *Proc. HSCC'01*, vol. 2034 of *LNCS*, pp. 147–161. Springer, 2001.

[6] A. Biere, A. Cimatti, E. Clarke, O. Strichman, Y. Zhu. Bounded model checking. In: *Highly Dependable Software*, vol. 58 of *Advances in Computers*. Academic Press, 2003. Pre-print.

[7] P. Bouyer, T. Brihaye, V. Bruyère, J. Raskin. On the optimal reachability problem of weighted timed automata. *Formal Methods System Design*, **31** (2), 135–175, 2007.

[8] P. Bouyer, E. Brinksma, K. G. Larsen. Optimal infinite scheduling for multi-priced timed automata. *Formal Methods in System Design*, **32** (1), 3–23, 2008.

[9] C. Courcoubetis, M. Yannakakis. Minimum and maximum delay problems in real-time systems. *Formal Methods in System Design*, **1** (4), 385–415, 1992.

[10] Y. Kesten, A. Pnueli, J. Sifakis, S. Yovine. Decidable integration graphs. *Information and Computation*, **150** (2), 209–243, 1999.

[11] K. G. Larsen, J. I. Rasmussen. Optimal reachability for multi-priced timed automata. *Theoretical Computer Science*, **390**, 197–213, 2008.

[12] P. Niebert, S. Tripakis, S. Yovine. Minimum-time reachability for Timed Automata. In: *Proc. 8th IEEE Mediterranean Conf. on Control and Automation (MED'2000)*, Patros, Greece, July 2000.

[13] W. Penczek, B. Woźna, A. Zbrzezny. Bounded model checking for the universal fragment of CTL. *Fund. Informaticae*, **51** (1-2), 135–156, 2002.

[14] B. Woźna, A. Zbrzezny, W. Penczek. Checking reachability properties for Timed Automata via SAT. *Fund. Informaticae*, **55**, 223–241, 2003.

[15] A. Zbrzezny. SAT-based reachability checking for timed automata with diagonal constraints. *Fund. Informaticae*, **67** (1-3), 303–322, 2005.

[16] A. Zbrzezny. A boolean encoding of arithmetic operations. In: *Proc. Int. Workshop on Concurrency, Specification and Programming*, vol. 170 of *Informatik-Berichte*, pp. 536–547. Humboldt University, 2008.

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XV, Częstochowa 2010

A BOOLEAN ENCODING OF ARITHMETIC OPERATIONS

Andrzej Zbrzezny

Institute of Mathematics and Computer Science
Jan Długosz University in Częstochowa
al. Armii Krajowej 13/15, 42-200 Częstochowa, Poland
e-mail: a.zbrzezny@ajd.czest.pl

Abstract. In this paper we present algorithms for a Boolean encoding of four basic arithmetic operations on integer numbers: addition, subtraction, multiplication and division. Integer numbers are encoded in two's complement system as vectors of Boolean formulae, and arithmetic operations are faithfully encoded as operations on vectors of Boolean formulae.

1. Introduction

Boolean encoding of arithmetic operations is an important issue in some areas of symbolic model checking, for example, in SAT-based model checking for timed automata with discrete data (TADD), i.e. timed automata augmented with integer variables. The first attempt to develop bounded model checking for TADD was undertaken in [9]. However, the set of arithmetic operations considered in this paper was limited to addition and subtraction of integer variables.

In Saturn [7], the system for static analysis of programs that was developed at Stanford University, Boolean encoding of arithmetic operations is also limited. In an unpublished technical report [1] there are listed many operations for constructing and manipulating vectors of Boolean formulae, among others, addition and subtraction. But for multiplication and division there are mentioned only restricted versions of operations on Boolean formulae: namely, multiplication and division of a Boolean vector by a constant integer number and division of a Boolean vector by a constant integer number.

There are many tools which make use of a Boolean encoding of arithmetic operations. One of them is C32SAT – a tool for checking C expressions by means of satisfiability testing [2]. C32SAT parses the input expression and builds a parse tree, which is transformed into an And-Inverter Graph. Afterwards, the graph is transformed into conjunctive normal form and passed to a SAT solver.

In this paper we show how to encode faithfully the four basic arithmetic operations for integer numbers: addition, subtraction, multiplication and division. Our algorithms for Boolean encoding of the operations in question are based on standard algorithms well-known in the theory of computer arithmetic.

In the technical report [8] we have also provided algorithms for a Boolean encoding of the operation of calculating integer square root and of the operation of exponentiation with nonnegative integer exponent.

2. Basic notions and notations

Definition 1. *Let $\mathcal{V}$ be a nonempty set of propositional variables. The set $\mathcal{F}(\mathcal{V})$ of* Boolean formulae *over $\mathcal{V}$ is defined by the following grammar:*

$$f ::= \boldsymbol{false} \mid \boldsymbol{true} \mid p \mid \neg f \mid f \vee f \mid f \wedge f$$

The propositional variables and the constants $\boldsymbol{false}$ and $\boldsymbol{true}$ are called *atomic Boolean formulae.* In order to enhance readability of Boolean encoding of arithmetic operations we shall use two auxiliary propositional connectives: $\oplus$ (*exclusive disjunction*) and $\equiv$ (*biconditional*), defined in the standard way:

$$f \oplus g = (f \wedge \neg g) \vee (\neg f \wedge g) \qquad f \equiv g = (f \wedge g) \vee (\neg f \wedge \neg g)$$

We assume that, from greatest to lowest priority, the priority order is as follows: $\neg$, $\wedge$, $\vee$, $\oplus$, $\equiv$.

Definition 2. *Let $\mathcal{B}_2 = \langle \{0, 1\}, -, \cup, \cap, 0, 1 \rangle$ be the two element Boolean algebra. A* valuation *v is a mapping from the set of atomic Boolean formulae to the universe of the Boolean algebra $\mathcal{B}_2$ satisfying the condition $v(\boldsymbol{false}) = 0$ and $v(\boldsymbol{true}) = 1$. The set of all the valuations will be denoted by $\mathcal{V}al(\mathcal{V})$.*

It is well known that each valuation v can be uniquely extended to a homomorphism h^v from the algebra of formulae $\langle \mathcal{F}(\mathcal{V}), \neg, \vee, \wedge \boldsymbol{false}, \boldsymbol{true} \rangle$ to the Boolean algebra $\mathcal{B}_2$.

From now on we shall write $\mathcal{F}$ and $\mathcal{V}al$ instead of $\mathcal{F}(\mathcal{V})$ and $\mathcal{V}al(\mathcal{V})$ respectively, as we assume that the set $\mathcal{V}$ of the propositional variables is fixed.

Definition 3. *A* vector of Boolean values *is a finite, nonempty sequence of Boolean values* 0 *and* 1.

As the Boolean values 0 and 1 can be identified with binary digits, from now on, vectors of Boolean values will be called *bit vectors*. Every bit vector will be interpreted as an integer encoded in the two's-complement system. Namely, let $\mathtt{a} = \langle \mathtt{a}_{n-1}, \ldots, \mathtt{a}_0 \rangle$ be a bit vector of length n. Define the interpretation $\mathcal{I}(\mathtt{a})$ in the following standard way:

$$\mathcal{I}(\mathtt{a}) \;=\; \left(\sum_{i=0}^{n-1} \mathtt{a}_i \cdot 2^i \right) - (\mathtt{a}_{n-1} \cdot 2^n)\,.$$

Definition 4. *A* vector of Boolean formulae *(a* Boolean vector *for short) is a finite, nonempty sequence of Boolean formulae. A set of all the Boolean vectors of length n will be denoted by* $\mathcal{BV}_n$.

Let $\mathtt{x} = \langle \mathtt{x}_{n-1}, \ldots, \mathtt{x}_0 \rangle$ be a Boolean vector and v be a valuation. Then a sequence $H^v(\mathtt{x}) = \langle h^v(\mathtt{x}_{n-1}), \ldots, h^v(\mathtt{x}_0) \rangle$ is a bit vector that will be interpreted as a number $\mathcal{I}(H^v(\mathtt{x}))$. From now on we shall write $\mathcal{I}^v(\mathtt{x})$ instead of $\mathcal{I}(H^v(\mathtt{x}))$.

It it well known from computer arithmetic that in two's complement representation of a number b the most significant bit is equal to 1 if and only if the number b is negative. Recall also that for every bit vector $\mathtt{a}$ of length n, the following hold:

$$-2^{n-1} \;\leq\; \mathcal{I}(\mathtt{a}) \;\leq\; 2^{n-1} - 1. \tag{1}$$

3. Encoding of arithmetic relations and operations

We start with an obvious observation that the result of an arithmetic operation may not fit in the two's complement representation of a given length n. This is clear for addition, subtracting and multiplication. There is also one particular case for division. Namely, when a dividend is equal to $-2^{n-1}-1$ and a divisor is equal to -1, the result, which is equal to 2^{n-1}, does not fit into n bits. Such a situation is called an *overflow*. This motivates the following notion of faithful encoding.

Let $\circ$ be a binary arithmetic operation and let $\boxdot$ be a binary operation on Boolean vectors. We say that the operation $\boxdot$ *encodes* the operation $\circ$ *faithfully* if and only if for every $\mathtt{x}, \mathtt{y} \in \mathcal{BV}_n$ and every $v \in \mathcal{V}al$,

$$-2^{n-1} \leq \mathcal{I}^v(\mathtt{x}) \circ \mathcal{I}^v(\mathtt{y}) \leq 2^{n-1} - 1 \implies \mathcal{I}^v(\mathtt{x} \boxdot \mathtt{y}) \;=\; \mathcal{I}^v(\mathtt{x}) \circ I^v(\mathtt{y}).$$

The definition of the faithfully encoding of a unary arithmetic operation is analogous.

In what follows we assume that there is a global variable `overflow` initially set to ***false*** in which the Boolean formula expressing a possible overflow is computed.

Let $\sim$ be a two argument arithmetic relation. We say that a two argument operation $\bowtie : \mathcal{BV}_n \times \mathcal{BV}_n \longrightarrow \mathcal{F}$ *faithfully encodes* the relation $\sim$ if and only if for every $\mathtt{x}, \mathtt{y} \in \mathcal{BV}_n$ and every $v \in \mathcal{V}al$,

$$\mathtt{h}^v(\mathtt{x} \bowtie \mathtt{y}) = 1 \iff \mathcal{I}^v(\mathtt{x}) \sim I^v(\mathtt{y}).$$

3.1. Encoding of the relation "equal to"

In order to find a Boolean formula that faithfully encodes the equality relation assume that v is an arbitrary but fixed valuation, and observe that $\mathcal{I}^v(\mathtt{x}) = \mathcal{I}^v(\mathtt{y})$ iff $h^v\left(\forall_{j=0}^{n-1}(\mathtt{x}_j \equiv \mathtt{y}_j)\right) = 1$. Thus, Algorithm 1 constructs a Boolean formula EQUAL$(\mathtt{x}, \mathtt{y})$ that is the conjunction of all the formulae of the form $\mathtt{x}_j \equiv \mathtt{y}_j$.

Algorithm 1 EQUAL

Input: Boolean vectors `x`, `y` of length n.

Output: A Boolean formula `f` such that $\forall v \in \mathcal{V}al, \mathcal{I}^v(\mathtt{f}) = 1 \iff \mathcal{I}^v(\mathtt{x}) = \mathcal{I}^v(\mathtt{y})$.

```
1: function EQUAL(x, y)
2:     f ← true
3:     for j ← 0 to n − 1 do
4:         f ← f ∧ (x[j] ≡ y[j])
5:     end for
6:     return f
7: end function
```

3.2. Addition

To define the addition of two Boolean vectors we adapt the method of the addition of two bit vectors known from computer arithmetic. Let `x`, `y` be two Boolean vectors of length n, i.e. let $\mathtt{x} = \langle \mathtt{x}_{n-1}, \ldots, \mathtt{x}_0 \rangle$ and $\mathtt{y} = \langle \mathtt{y}_{n-1}, \ldots, \mathtt{y}_0 \rangle$, where for every $0 \leq k < n$, $\mathtt{x}_k$ and $\mathtt{y}_k$ are Boolean formulae. Define an ordered pair of Boolean vectors $\langle \mathtt{w}, \mathtt{c} \rangle \in \mathcal{BV}_n \times \mathcal{BV}_{n+1}$ as follows: first, let $\mathtt{c}_0 = 0$; next, for $0 \leq k < n$, let

$$\langle \mathtt{w}_k,\ \mathtt{c}_{k+1} \rangle = \langle \mathtt{x}_k \oplus \mathtt{y}_k \oplus \mathtt{c}_k,\ (\mathtt{x}_k \wedge \mathtt{y}_k) \vee (\mathtt{x}_k \wedge \mathtt{c}_k) \vee (\mathtt{y}_k \wedge \mathtt{c}_k) \rangle .$$

The vector $\mathtt{w}$ represents the sum of $\mathtt{x}$ and $\mathtt{y}$, and the vector $\mathtt{c}$ represents the succeeding carry bits. Clearly, the sum of two bit vectors of length n may not fit into n bits. By (1), this happens if and only if the sum is less than -2^n or greater than $2^n - 1$. It is known from computer arithmetic that adding two integers cause an overflow exactly when the carry bits $\mathtt{c}_n$ and $\mathtt{c}_{n+1}$ are different.

Algorithm 2 ADD

Input: Boolean vectors $\mathtt{x}$, $\mathtt{y}$ of length n.

Output: A Boolean vector $\mathtt{w}$ of length n such that $\forall v \in Val$, if $-2^{n-1} \leq \mathcal{I}^v(\mathtt{x}) + I^v(\mathtt{y}) \leq 2^{n-1} - 1$, then $\mathcal{I}^v(\mathtt{w}) = \mathcal{I}^v(\mathtt{x}) + I^v(\mathtt{y})$.

```
1: function ADD(x, y)
2:     c[0] ← false
3:     for k ← 0 to n − 1 do
4:         w[k] ← x[k] ⊕ y[k] ⊕ c[k]
5:         c[k + 1] ← (x[k] ∧ y[k]) ∨ (x[k] ∧ c[k]) ∨ (y[k] ∧ c[k])
6:     end for
7:     overflow ← overflow ∨ (c[n] ⊕ c[n + 1])
8:     return w
9: end function
```

3.3. Subtraction

Notice that in order to subtract two integers it is enough to add to the first number the additive inverse of the second number. Therefore, we need an operation on Boolean vectors that encodes additive inverse.

Recall that computing additive inverse for a two's complement number involves complementing each bit and then adding 1. It follows that we need an operation for creating a Boolean vector that represents the number 1. It is obvious that the number 1 is represented by the Boolean vector of the form $\langle \mathit{false}, \ldots, \mathit{false}, \mathit{true} \rangle$, and an algorithm for creating this vector is trivial. Nevertheless, it will be useful to provide a more general Algorithm 3 that for a given integer creates a Boolean vector representing that number.

In this algorithm we use the operation $>>$ of arithmetic right shift also known as signed shift. Recall that in `Java` the operator $>>$ designates signed shift, whereas in `C++` a meaning of the operator $>>$ is implementation-defined. Note that in `gcc` compiler, i.e the compiler we use, the operator $>>$ is implemented as signed shift. In order to ensure that an implementation of Algorithm 3 in the language `C++` is independent of an used compiler, one should use a proper implementation of signed shift instead of the operator $>>$. Now

Algorithm 3 BOOLVEC

Input: A number of bits n and an integer a.

Output: A Boolean vector $\mathtt{w}$ of length n such that $\forall v \in \mathcal{V}al$, if $-2n-1 \leq a \leq 2^{n-1}-1$, then $\mathcal{I}^v(\mathtt{w}) = a$.

```
1: function BoolVec(n, a)
2:     if a < 0 then
3:         w[n − 1] ← true
4:     else
5:         w[n − 1] ← false
6:     end if
7:     for k ← 0 to n − 2 do
8:         if a mod 2 = 0 then
9:             w[k] ← false
10:        else
11:            w[k] ← true
12:        end if
13:        a ← a >> 1
14:    end for
15:    overflow ← overflow ∨ (a ≠ 0 ∧ a ≠ −1)
16:    return w
17: end function
```

we are able to write down an algorithm that computes the additive inverse of a Boolean vector $\mathtt{x}$, also called the opposite of $\mathtt{x}$.

Note that in two's complement arithmetic adding the number 1 to a n-bit number b generates overflow if and only if $b = 2^n - 1$. As the two's complement representation of the number $2^n - 1$ is of the form $\langle 0, 1, \ldots, 1\rangle$, which is the result of complementing each bit in the vector $\langle 1, 0, \ldots, 0\rangle$ that represents the number -2^{n-1}, it follows that taking additive inverse of a given number a generates overflow exactly when $a = -2^{n-1}$.

In the algorithm for subtracting we need some auxiliary operations on Boolean vectors.

For a given Boolean vector $\mathtt{x}$ of length n and an integer $m \geq n$, the auxiliary operation EXTEND implemented in Algorithm 5 creates a Boolean vector $\mathtt{w}$ of length m that represents the same integers as the vector $\mathtt{x}$. This is done by copying all the elements of $\mathtt{x}$ to the corresponding elements of y and then copying the sign bit of x to the most significant $m - n$ elements of y. The algorithm described reflects the known operation of extension that consists in increasing the number of bits of a binary number while preserving the number's sign and value. For example, if 8 bits are used to represent the value

Algorithm 4 OPP

Input: A Boolean vector $\mathtt{x}$ of length n.

Output: A Boolean vector $\mathtt{w}$ of length n such that $\forall v \in \mathcal{V}al$, if $-2^{n-1} < \mathcal{I}^v(\mathtt{x}) \leq 2^{n-1} - 1$, then $\mathcal{I}^v(\mathtt{w}) = -\mathcal{I}^v(\mathtt{x})$.

```
1: function OPP(x)
2:     for k ← 0 to n − 1 do
3:         w[k] ← ¬x[k]
4:     end for
5:     w ← ADD(w, BOOLVEC(n, 1))
6:     return w
7: end function
```

Algorithm 5 EXTEND

Input: A Boolean vector $\mathtt{x}$ of length n and a positive number $m \geq n$.

Output: A Boolean vector $\mathtt{w}$ of length m such that $\forall v \in \mathcal{V}al$, $\mathcal{I}^v(\mathtt{w}) = \mathcal{I}^v(\mathtt{x})$.

```
1: function EXTEND(x, m)
2:     for k ← 0 to n − 1 do
3:         w[k] ← x[k]
4:     end for
5:     for k ← n to m − 1 do
6:         w[k] ← x[n − 1]
7:     end for
8:     return w
9: end function
```

-15 using two's complement $\langle 1,1,1,1,0,0,0,1\rangle$, and sign extend to 16 bits is used, the new representation would be $\langle 1,1,1,1,1,1,1,1,1,1,1,1,0,0,0,1\rangle$.

For a given Boolean vector $\mathtt{x}$ of length m, the auxiliary operation REDUCE implemented in Algorithm 6 creates a Boolean vector $\mathtt{w}$ of length $n \leq m$ such that every integer represented by the Boolean $\mathtt{w}$ is also represented by the Boolean vector $\mathtt{x}$.

Now we are in a position to write down an algorithm for subtraction of two Boolean vectors representing integer numbers. At the beginning the algorithm enlarges both the arguments by one bit in order to avoid a possible overflow that may occur in the operation of taking the additive inverse. Then the algorithm adds the enlarged first argument to the additive inverse of the enlarged second argument and puts the result in an auxiliary Boolean vector $\mathtt{w}$. Eventually, the overflow is computed and as the result of subtraction the algorithm returns the Boolean vector REDUCE($\mathtt{w}$, n).

Algorithm 6 REDUCE

Input: A Boolean vector $\mathtt{x}$ of length m and a positive number $n \leq m$.
Output: A Boolean vector $\mathtt{w}$ of length n such that $\forall v \in \mathcal{V}al$, if $-2^n \leq \mathcal{I}^v(\mathtt{x}) \leq 2^n - 1$, then $\mathcal{I}^v(\mathtt{w}) = \mathcal{I}^v(\mathtt{x})$.

```
1: function REDUCE(x, n)
2:     for k ← 0 to n − 2 do
3:         w[k] ← x[k]
4:     end for
5:     w[n − 1] ← x[n − 1]
6:     return w
7: end function
```

Algorithm 7 SUBTRACT

Input: Boolean vectors $\mathtt{x}$, $\mathtt{y}$ of length n.
Output: A Boolean vector $\mathtt{w}$ of length n such that $\forall v \in \mathcal{V}al$, if $-2^{n-1} \leq \mathcal{I}^v(\mathtt{x}) - I^v(\mathtt{y}) \leq 2^{n-1} - 1$, then $\mathcal{I}^v(\mathtt{w}) = \mathcal{I}^v(\mathtt{x}) - I^v(\mathtt{y})$.

```
1: function SUBTRACT(x, y)
2:     p ← EXTEND(x, n + 1); q ← EXTEND(y, n + 1)
3:     w ← ADD(p, OPP(q))
4:     overflow ← overflow ∨ ¬(w[n − 1] ≡ w[n])
5:     return REDUCE(w, n)
6: end function
```

3.4. Multiplication

In the algorithm for multiplication we need some additional auxiliary operations on Boolean vectors, namely, SHIFTLEFT, CONJUNCTION and ABS.

The auxiliary operation SHIFTLEFT, which is implemented in Algorithm 8, is the operation of shifting a Boolean vector one bit left, and after shifting, filling in the least significant position of the vector with the Boolean formula $\boldsymbol{false}$. This operation corresponds to the well known operation called a logical shift. The operation of shifting is also used in the algorithm for dividing nonnegative integers.

The operation CONJUNCTION implemented in Algorithm 9 creates the bitwise conjunction of a Boolean formula and a Boolean vector. This simple operation enables simulating a conditional execution in the algorithms of multiplying and dividing. The reason for this is the following obvious property:

$$(\forall v \in \mathcal{V}al)\ \mathcal{I}^v(\textsc{Conjunction}(f, \mathtt{x}) = \begin{cases} 0, & \text{if } \mathcal{I}^v(f) = 0 \\ \mathcal{I}^v(\mathtt{x}), & \text{if } \mathcal{I}^v(f) = 1. \end{cases}$$

Algorithm 8 SHIFTLEFT

Input: A Boolean vector `x` of length n.
Output: A Boolean vector `x` logically shifted left by one.

```
1: procedure ShiftLeft(x)
2:     for k ← n − 1 down to 1 do
3:         x[k] ← x[k − 1]
4:     end for
5:     x[0] ← false
6: end procedure
```

Algorithm 9 CONJUNCTION

Input: A Boolean formula f and a Boolean vector `x` of length n.
Output: A Boolean vector `w` of length n such that for every $0 \leq k < n$, $\mathtt{w}[k] = f \wedge \mathtt{x}[k]$.

```
1: function Conjunction(f, x)
2:     for k ← 0 to n − 1 do
3:         w[k] ← f ∧ x[k]
4:     end for
5:     return w
6: end function
```

For a given Boolean vector `x` of length n, the auxiliary operation ABS implemented in Algorithm 10 creates a Boolean vector `w` of length n such that `w` represents the absolute value of `x`.

Now we are in a position to write down the algorithm for multiplication of two Boolean vectors representing nonnegative integers. Algorithm 11 creates a Boolean vector that represents the result of multiplication of two Boolean vectors that represent nonnegative integers. We adapted the simplest method that computes the product one bit at a time, and is a symbolic version of the paper-and-pencil method.

Note that at the beginning of the algorithm some preparatory steps are needed. First, both the arguments are copied to auxiliary variables `p` and `q`; next, the most significant bit of each of the auxiliary variables is set to ***false***; eventually, both the auxiliary variables are enlarged to size $2 \cdot n$, and `w` is set to $\langle \boldsymbol{false}, \ldots, \boldsymbol{false} \rangle$.

After these preparatory steps, the algorithm proceeds as follows: for every k from 0 to $n-1$ the conjunction of the multiplicand and the kth element of multiplier is added to `w`. This last step simulates the conditional addition of the multiplicand to the product: the multiplicand is added in the kth step if and only if the kth element of multiplier represents the binary value 1.

Algorithm 10 ABS

Input: A Boolean vector $\mathtt{x}$ of length n.

Output: A Boolean vector $\mathtt{w}$ of length n such that $\forall v \in \mathcal{V}al$, if $-2^{n-1} < \mathcal{I}^v(\mathtt{x}) \leq 2^{n-1}-1$, then $\mathcal{I}^v(\mathtt{w}) = |\mathcal{I}^v(\mathtt{x})|$.

1: **function** ABS($\mathtt{x}$)
2: $\quad\mathtt{y} \leftarrow$ OPP(x)
3: $\quad$**for** $k \leftarrow 0$ **to** $n-2$ **do**
4: $\quad\quad\mathtt{w}[k] \leftarrow (\mathtt{x}[n-1] \wedge \mathtt{y}[k]) \vee (\neg\mathtt{x}[n-1] \wedge \mathtt{x}[k])$
5: $\quad$**end for**
6: $\quad\mathtt{w}[n-1] \leftarrow \boldsymbol{false}$
7: $\quad$**return** $\mathtt{w}$
8: **end function**

Algorithm 11 MULTIPLYNONNEG

Input: Boolean vectors $\mathtt{x}$, $\mathtt{y}$ of length n

Output: A Boolean vector $\mathtt{w}$ of length $2 \cdot n$ such that $\forall v \in \mathcal{V}al$, if $\mathcal{I}^v(\mathtt{x}) \geq 0$ and $\mathcal{I}^v(\mathtt{y}) \geq 0$, then $\mathcal{I}^v(\mathtt{w}) = \mathcal{I}^v(\mathtt{x}) \cdot I^v(\mathtt{y})$.

1: **function** MULTIPLYNONNEG($\mathtt{x}$, $\mathtt{y}$)
2: $\quad\mathtt{p} \leftarrow \mathtt{x}$; $\mathtt{p}[n-1] \leftarrow \boldsymbol{false}$; $\mathtt{p} \leftarrow$ EXTEND($\mathtt{p}$, $2 \cdot n$)
3: $\quad\mathtt{q} \leftarrow \mathtt{y}$; $\mathtt{q}[n-1] \leftarrow \boldsymbol{false}$; $\mathtt{q} \leftarrow$ EXTEND($\mathtt{q}$, $2 \cdot n$)
4: $\quad\mathtt{w} \leftarrow$ BOOLVEC($2 \cdot n$, 0)
5: $\quad$**for** $k \leftarrow 0$ **to** $n-2$ **do**
6: $\quad\quad\mathtt{w} \leftarrow$ ADD($\mathtt{w}$, CONJUNCTION($\mathtt{q}[k]$, $\mathtt{p}$)
7: $\quad\quad$SHIFTLEFT($\mathtt{p}$)
8: $\quad$**end for**
9: $\quad$**return** $\mathtt{w}$
10: **end function**

The following Algorithm 12 creates a Boolean vector that represents the result of multiplication of two Boolean vectors that represent signed integers. At the beginning, the algorithm enlarges both the arguments by one bit. Then, two cases are considered: the arguments are of the same sign (f_0) and the arguments have different signs (f_1). In each of the cases the algorithm symbolically converts the arguments to be nonnegative, does an unsigned multiplication, and for the case when the original arguments have different signs, negates the result. Next, from the two symbolic results, named $\mathtt{w_0}$ and $\mathtt{w_1}$, the final result is created in the following way: for every k such that $0 \leq k < 2 \cdot (n+1)$, the kth bit of the product is set to $f_0 \wedge \mathtt{w_0}[k] \vee f_1 \wedge \mathtt{w_1}[k]$. Eventually, the overflow is computed and the result is reduced to n bits.

Algorithm 12 MULTIPLY

Input: Boolean vectors $\mathtt{x}$, $\mathtt{y}$ of length n

Output: A Boolean vector $\mathtt{w}$ of length n such that $\forall v \in \mathcal{V}al$, if $-2^{n-1} \leq \mathcal{I}^v(\mathtt{x}) \cdot I^v(\mathtt{y}) \leq 2^{n-1} - 1$, then $\mathcal{I}^v(\mathtt{w}) = \mathcal{I}^v(\mathtt{x}) \cdot I^v(\mathtt{y})$.

1: **function** MULTIPLY($\mathtt{x}$, $\mathtt{y}$)
2: $\quad \mathtt{p} \leftarrow$ EXTEND($\mathtt{x}$, $n+1$); $\mathtt{q} \leftarrow$ EXTEND($\mathtt{y}$, $n+1$)
3: $\quad \mathtt{w_0} \leftarrow$ MULTIPLYNONNEG(ABS($\mathtt{p}$), ABS($\mathtt{q}$))
4: $\quad \mathtt{w_1} \leftarrow$ OPP($\mathtt{w_0}$)
5: $\quad f_0 \leftarrow (\neg \mathtt{x}[n-1] \wedge \neg \mathtt{y}[n-1]) \vee (\mathtt{x}[n-1] \wedge \mathtt{y}[n-1])$
6: $\quad f_1 \leftarrow (\neg \mathtt{x}[n-1] \wedge \mathtt{y}[n-1]) \vee (\mathtt{x}[n-1] \wedge \neg \mathtt{y}[n-1])$
7: $\quad m \leftarrow 2 \cdot (n+1)$
8: $\quad$ **for** $k \leftarrow 0$ **to** $m-1$ **do**
9: $\quad\quad \mathtt{w}[k] \leftarrow f_0 \wedge \mathtt{w_0}[k] \vee f_1 \wedge \mathtt{w_1}[k]$
10: $\quad$ **end for**
11: $\quad of \leftarrow \boldsymbol{false}$
12: $\quad$ **for** $k \leftarrow n-1$ **to** $m-2$ **do**
13: $\quad\quad \mathtt{of} \leftarrow \mathtt{of} \vee \neg(\mathtt{w}[k] \equiv \mathtt{w}[m-1])$
14: $\quad$ **end for**
15: $\quad \mathtt{overflow} \leftarrow \mathtt{overflow} \vee of$
16: $\quad$ **return** REDUCE($\mathtt{w}$, n)
17: **end function**

3.5. Division

There are many possible algorithms for dividing nonnegative integers. We adapted the so called restoring radix-2 division algorithm described in Appendix H of [3]. Algorithm 13 is done by shifts, subtractions, additions and testing whether the number is negative. The algorithm needs four registers: one for the dividend $\mathtt{x}$, one for the divisor $\mathtt{y}$, one for the quotient $\mathtt{q}$, and one for the remainder $\mathtt{r}$. The registers $\mathtt{r}$ and $\mathtt{q}$ form a double-length register pair. The register $\mathtt{q}$ is initially set to the value of $\mathtt{x}$ and the register $\mathtt{q}$ is initially set to 0.

Algorithm 14 creates a Boolean vector that represents the result of division of two Boolean vectors that represent signed integers. There are the same cases to consider for arguments as in Algorithm 12. Also the method of computing the final result is nearly the same. There are only two differences. The first one is that the result is not reduced to the length of arguments, as in all the cases considered the results are of length n. The second one is the method of setting the overflow.

Algorithm 13 DIVIDENONNEG

Input: Boolean vectors $\mathtt{x}$, $\mathtt{y}$ of length n

Output: Boolean vectors $\mathtt{r}, \mathtt{q}$ such that $\forall v \in \mathcal{V}al$, if $\mathcal{I}^v(\mathtt{x}) \geq 0$ and $\mathcal{I}^v(\mathtt{y}) > 0$, then $\mathcal{I}^v(\mathtt{x}) = \mathcal{I}^v(\mathtt{q}) \cdot I^v(\mathtt{y}) + \mathcal{I}^v(\mathtt{r})$.

```
1:  function DivideNonNeg(x, y)
2:      q ← x; r ← BoolVec(n, 0)
3:      for k ← 0 to n − 1 do
4:          ShiftLeft(r)
5:          r[0] ← q[n − 1]
6:          ShiftLeft(q)
7:          r ← Subtract(r, y)
8:          q[0] ← ¬r[n − 1]
9:          r ← Add(r, Conjunction(r[n − 1], y))
10:     end for
11:     return ⟨q, r⟩
12: end function
```

We would also point out that the signs of the quotient and of the remainder for negative dividends and/or negatives divisors are computed in accordance with the following rules of `C++` and `Java`: the quotient is negative if and only if both the dividend and the divisor have different signs, and the remainder is negative if and only if the dividend is negative.

3.6. Encoding of the relation "less than"

Let us note that the relation "less than" can be encoded by using the operation of subtraction. The algorithm enlarges both the arguments by one bit in order to avoid a possible overflow that may occur in the operation of subtraction and then returns the most significant element of the Boolean vector representing the difference.

4. Implementation

We have implemented the described algorithms in the programming language `C++` by designing the following classes: the class `BoolForm` that implements basic logical operations on Boolean formulae; the class `BoolFormVect` that implements basic operations on Boolean vectors; and the class `Integer`, derived from `BoolFormVect`, that implements the Boolean encoding of arithmetic relations and operations as described in this paper.

Algorithm 14 DIVIDE

Input: Boolean vectors $\mathtt{x}$, $\mathtt{y}$ of length n

Output: Boolean vectors $\mathtt{q}$, $\mathtt{r}$ such that $\forall v \in \mathcal{V}al$, if $\mathcal{I}^v(\mathtt{y}) \neq 0$, then $\mathcal{I}^v(\mathtt{x}) = \mathcal{I}^v(\mathtt{q}) \cdot I^v(\mathtt{y}) + \mathcal{I}^v(\mathtt{r})$, $sgn(\mathcal{I}^v(\mathtt{q})) = sgn(\mathcal{I}^v(\mathtt{x})) \cdot sgn(\mathcal{I}^v(\mathtt{y}))$, and $sgn(\mathcal{I}^v(\mathtt{r})) = sgn(\mathcal{I}^v(\mathtt{x}))$.

1: **function** DIVIDE($\mathtt{x}$, $\mathtt{y}$)
2: $\quad\mathtt{p} \leftarrow$ EXTEND($\mathtt{x}$, $n+1$); $\mathtt{q} \leftarrow$ EXTEND($\mathtt{y}$, $n+1$)
3: $\quad\langle \mathtt{q}_0, \mathtt{r}_0 \rangle \leftarrow$ DIVIDENONNEG(ABS($\mathtt{p}$), ABS($\mathtt{q}$))
4: $\quad\mathtt{q}_1 \leftarrow$ OPP($\mathtt{q}_0$); $\mathtt{r}_1 \leftarrow$ OPP($\mathtt{r}_0$)
5: $\quad\mathtt{f}_{00} \leftarrow \neg\mathtt{x}[n-1] \wedge \neg\mathtt{y}[n-1])$; $\mathtt{f}_{01} \leftarrow \neg\mathtt{x}[n-1] \wedge \mathtt{y}[n-1])$
6: $\quad\mathtt{f}_{10} \leftarrow \mathtt{x}[n-1] \wedge \neg\mathtt{y}[n-1])$; $\mathtt{f}_{11} \leftarrow \mathtt{x}[n-1] \wedge \mathtt{y}[n-1])$
7: $\quad$**for** $k \leftarrow 0$ **to** n **do**
8: $\quad\quad\mathtt{q}[k] \leftarrow ((\mathtt{f}_{00} \vee \mathtt{f}_{11}) \wedge \mathtt{q}_0[k]) \vee ((\mathtt{f}_{01} \vee \mathtt{f}_{10}) \wedge \mathtt{q}_1[k])$
9: $\quad\quad\mathtt{r}[k] \leftarrow ((\mathtt{f}_{00} \vee \mathtt{f}_{01}) \wedge \mathtt{r}_0[k]) \vee ((\mathtt{f}_{10} \vee \mathtt{f}_{11}) \wedge \mathtt{r}_1[k])$
10: $\quad$**end for**
11: $\quad\mathtt{a} \leftarrow$ BOOLVEC($n+1$, 0)
12: $\quad\mathtt{b} \leftarrow$ BOOLVEC($n+1$, 1)
13: $\quad\mathtt{z} \leftarrow$ BOOLVEC($n+1$, 2^{n-1})
14: $\quad\mathtt{of} \leftarrow$ EQUAL($\mathtt{p}$, OPP($\mathtt{z}$)) $\wedge$ EQUAL($\mathtt{q}$, OPP($\mathtt{b}$)) $\vee$ EQUAL($\mathtt{y}$, $\mathtt{a}$)
15: $\quad\mathtt{overflow} \leftarrow \mathtt{overflow} \vee \mathtt{of}$
16: $\quad$**return** $\langle \mathtt{q}, \mathtt{r} \rangle$
17: **end function**

In order to test the above algorithms we have created testing programs for all the arithmetic operations considered. In every program some suitable formula φ is tested in the following way: at first, φ is converted to a set of clauses C in a way such that although the set C is not logically equivalent to the formula φ, it preserves satisfiability, i.e. C is satisfiable if and only if φ is satisfiable; then, we check satisfiability of C by using `MiniSat`. Some of experimental results for the programs mentioned above are provided in [8].

5. Final remarks

As a result of implementing our Boolean encoding of arithmetic operations we were able to extend the module `BMC4TADD` of the model checker Verics [4] in order to include multiplication and division in the set of the allowed operations. The module `BMC4TADD` serves for verification of properties of timed automata with discrete data. The formalism of timed automata with discrete data and basic arithmetic operations is now used in verification of Java programs (see [6, 10]). The Boolean encoding of arithmetic operations was also used in a new approach to model checking of systems specified in UML (see [5]).

References

[1] A. Aiken, S. Bugrara, I. Dillig, T. Dillig, B. Hackett, P. Hawkins. The Saturn program analysis system. Technical Report, Stanford University, 2006.

[2] R. Brummayer, A. Biere. C32SAT: Checking C expressions. In: *Proc. CAV'2007*, LNCS 4590, pp. 294–297, Springer, Berlin, 2007.

[3] J.L. Hennessy, D.A. Patterson. *Computer Architecture: A Quantitative Approach*, 3rd edition. Morgan Kaufmann Publishers, San Francisco, CA, 2003.

[4] M. Kacprzak, W. Nabiałek, A. Niewiadomski, W. Penczek, A. Pólrola, M. Szreter, B. Woźna, A. Zbrzezny. VerICS 2007 - a model checker for knowledge and real-time. *Fund. Informaticae*, **85** (1-4), 313–328, 2008.

[5] A. Niewiadomski, W. Penczek, M. Szreter. A new approach to model checking of UML state machines. *Fund. Informaticae*, **93** (1-3), 289–303, 2009.

[6] A. Rataj, B. Woźna, A. Zbrzezny. A translator of Java programs to TADDs. *Fund. Informaticae*, **93** (1-3), 305–324, 2009.

[7] Y. Xie, A. Aiken. Saturn: A SAT-based tool for bug detection. In: *Proc. CAV'2005*, LNCS 3576, pp. 139–143. Springer, Berlin, 2005.

[8] A. Zbrzezny. A boolean encoding of arithmetic operations. Technical Report 999, ICS PAS, 2007.

[9] A. Zbrzezny, A. Pólrola. SAT-based reachability checking for timed automata with discrete data. *Fund. Informaticae*, **79** (3–4), 579–593, 2007.

[10] A. Zbrzezny, B. Woźna. Towards verification of Java programs in VerICS. *Fund. Informaticae*, **85** (1-4), 533–548, 2008.

Jan Długosz University in Częstochowa
Scientific Issues, Mathematics XV, Częstochowa 2010

AN INFLUENCE OF SERVICE DISCIPLINE ON CHARACTERISTICS OF A SINGLE-SERVER QUEUE WITH NON-HOMOGENEOUS CUSTOMERS

Oleg Tikhonenko, Artur Gola, Marcin Ziółkowski

Institute of Mathematics and Computer Science
Jan Długosz University in Częstochowa
al. Armii Krajowej 13/15, 42-200 Częstochowa, Poland
e-mail: oleg.tikhonenko@gmail.com

Abstract. For single-server queueing systems with non-homogeneous customers having some random space requirements we compare processor-sharing and FIFO disciplines and investigate their influence on the total sum of space requirements characteristics (when this sum is not limited, i.e. $V = \infty$) and customers loss probability (when this sum is limeted, i.e. $V < \infty$), using analytical modeling and simulation.

1. Introduction

In the present work we investigate single-server queueing systems with non-homogeneous customers. This means that

1) each customer is characterized by some non-negative random capacity ζ;

2) customer's length ξ and his capacity ζ are generally dependent.

Note that we shall use the notion "customer length" instead of "service time". The difference between these notions is essential for processor sharing systems. The amount of work necessary for customer's service is called the customer length [5], i.e. the customer service time under condition that there are no other customers on service during this time period. Analogously, the residual length of the customer is referrred as his residual service time after some time instant under the same condition.

The total sum $\sigma(t)$ of capacities of all the customers present in the system at arbitrary time t may be limited by some constant value V $(0 < V \leq \infty)$ that is called the capacity of the system.

Such systems are used to model and solve various problems occurring in the design of computer and communicating nets and systems. It is clear that they differ from usual classical queueing systems in the case $V < \infty$. For example, we can analyze the non-classical system $M/G/1/(\infty, V)$ with limited capacity that differs from the classical system $M/G/1/\infty$.

Let

$$F(x,t) = \mathbf{P}\{\zeta < x, \xi < t\}$$

be the distribution function of the random vector (ζ, ξ). Then

$$L(x) = \mathbf{P}\{\zeta < x\} = F(x,\infty) \quad \text{and} \quad B(t) = \mathbf{P}\{\xi < t\} = F(\infty, t)$$

are the distribution functions of customer's capacity and length, respectively. The part of system capacity is occupied by a customer at the epoch he arrives and is released entirely at the epoch he completes service. The process $\sigma(t)$ is called the total customers capacity.

Total capacity limitation (in the case $V < \infty$) leads to losses of customers. A customer arriving at the epoch τ and having capacity x will be admitted to the system if $\sigma(\tau - 0) + x \leq V$. Otherwise $(\sigma(\tau - 0) + x > V)$, the customer will be lost.

Various single-server queueing systems with non-homogeneous (in the sense of assumptions 1, 2) customers were analyzed in [1–4].

The purpose of this paper is to compare processor sharing and FIFO or other conservative, not depending on customers capacity disciplines and investigate their influence on the stationary first moment of the total sum of customers capacities (when $V = \infty$) and customers loss probability (when $V < \infty$). To realize this purpose we use analytical modeling and simulation.

2. The case of unlimited system capacity

Suppose that customers intrance flow is Poisson. Let a be an arrival rate of entrance flow of customers. Assume that $V = \infty$. Then we have the classical $M/G/1/\infty$ and $M/G/1/\infty - EPS$ (processor sharing) systems without losses of customers. For such a system we can obtain the stationary characteristics of total customers capacity (see e.g. [2, 3]).

We shall use the following notation. Denote by

$$\alpha(s,q) = \int_0^\infty \int_0^\infty e^{-sx-qt} dF(x,t)$$

the double Laplace-Stieltjes transfom (LST) of the function $F(x,t)$. Let $\varphi(s) = \alpha(s,0)$ and $\beta(q) = \alpha(0,q)$ be the LST of the functions $L(x)$ and $B(t)$, respectively. Let $D(x) = \mathbf{P}\{\sigma < x\}$ be the distribution function of stationary total customers capacity σ. Let $\varphi_i = \mathbf{E}\zeta^i$, $\beta_i = \mathbf{E}\xi^i$ and $\alpha_{ij} = \mathbf{E}(\zeta^i\xi^j)$ be the ith moments of the random variables ζ, ξ and the mixed $(i+j)$th moment of the random variables ζ and ξ, respectively. $i,j = 1,2,\ldots,$ $\rho = a\beta_1 < 1$. Denote by $\delta(s) = \int_0^\infty e^{-sx}dD(x)$ the LST of the function $D(x)$ and by $\delta_i = \mathbf{E}\sigma^i$ the ith moment of total customers capacity σ, $i = 1,2,\ldots.$

Then for the system $M/G/1/\infty$ (or for the discipline FIFO) we have [4]:

$$\delta_1^{\mathrm{FIFO}} = \mathbf{E}\sigma^{\mathrm{FIFO}} = a\alpha_{11} + \frac{a^2\beta_2\varphi_1}{2(1-\rho)}. \tag{1}$$

For the system $M/G/1/\infty - EPS$ (or for the discipline EPS) we get [2]:

$$\delta_1^{\mathrm{EPS}} = \mathbf{E}\sigma^{\mathrm{EPS}} = \frac{a\alpha_{11}}{1-\rho}. \tag{2}$$

From the simple relations (1) and (2), we obtain that $\delta_1^{\mathrm{FIFO}} < \delta_1^{\mathrm{EPS}}$ if the inequality $2\beta_1\alpha_{11} > \beta_2\varphi_1$ takes place. For example, if the random variables ζ and ξ are independent, i.e. $\alpha_{11} = \varphi_1\beta_1$, the last inequality takes the form $2\beta_1^2 > \beta_2$. Note that for exponential distributed customer length we have $2\beta_1^2 = \beta_2$. So, in this case for independent ζ and ξ we obtain that $\delta_1^{\mathrm{FIFO}} = \delta_1^{\mathrm{EPS}}$. If the customer length distribution is characterized by variation which is less than for exponential one, we always have $\delta_1^{\mathrm{FIFO}} < \delta_1^{\mathrm{EPS}}$. Evidently, this will be true for the case of positive correlated ζ and ξ (when $\alpha_{11} > \varphi_1\beta_1$).

For many real computer systems (for example, for communicating centers) the customer length can be defined by the relation $\xi = c\zeta + \xi_1$, where $c \geq 0$ and the random variables ζ and ξ_1 are independent.

Denote by κ_i the ith moment of the random variable ξ_1, $i = 1,2,\ldots.$ Then the first moments of the random variables σ^{FIFO} and σ^{EPS} can be calculated from relations (1) and (2), respectively, where [3]

$$\alpha_{11} = \varphi_1\kappa_1 + c\varphi_2,\ \beta_1 = c\varphi_1 + \kappa_1,\ \beta_2 = c^2\varphi_2 + 2c\varphi_1\kappa_1 + \kappa_2.$$

In this case we have that $\delta_1^{\mathrm{FIFO}} < \delta_1^{\mathrm{EPS}}$ if the following inequality takes place:

$$c^2\varphi_1\varphi_2 + 2\kappa_1(\varphi_1\kappa_1 + c\varphi_2) > \varphi_1\kappa_2. \tag{3}$$

In particular, if a customer length is proportional to his capacity, i.e. $\kappa_1 \equiv 0$, $\kappa_2 \equiv 0$, we have from (3) that $c^2\varphi_1\varphi_2 > 0$. Evidently, this inequality is always

true. For example, if we assume additionally that the customer length ζ has an exponential distribution with parameter f, we obtain:

$$\delta_1^{\text{FIFO}} = \frac{1}{f} \cdot \frac{\rho(2-\rho)}{1-\rho}, \qquad \delta_1^{\text{EPS}} = \frac{1}{f} \cdot \frac{2\rho}{1-\rho}.$$

Intuitively this is clear, because in the case of EPS discipline short (or having small capacity) customers are for a small time in the system, while FIFO service organization does not depend on the customer capacity.

3. The case of limited system capacity

In this case, it is interesting to compare loss characteristics for EPS and FIFO disciplines.

If customer's length does not depend on his capacity and has an exponential distribution with parameter f, we obtain [6] for systems $M/M/1/(\infty, V)$ and $M/G/1/(\infty, V) - EPS$ with the same $\rho = a\beta_1$ that the loss probability P has the form:

$$P^{\text{FIFO}} = P^{\text{EPS}} = \begin{cases} \dfrac{1-\rho}{e^{(1-\rho)fV} - \rho} & \text{if } \rho \neq 1, \\ \\ (1+fV)^{-1} & \text{if } \rho = 1. \end{cases}$$

Note that $\beta_1 = 1/\mu$ for the system $M/M/1/(\infty, V)$, where μ is the parameter of customer length.

Later on, we shall compare loss probabilities P and probabilities Q that unit of customer's capacity will be lost (see [7]) for cases of FIFO and EPS disciplines. It is clear (in this case) that probability Q is also the same for both systems under consideration. This fact can be confirmed by results of simulation (see Appendix, tables 1 and 2, where $f = 1$, $\mu = 1$). In our notation, we shall use the low indexes "an" or "sim" to demonstrate that an apropriate characteristic was obtained analytically or by simulation, respectively.

It can be conformed analytically and by simulation that we have the same results for loss characteristics P and Q in the systems $M/M/1/(\infty, V)$ and $M/G/1/(\infty, V) - EPS$ with the same ρ, when customer's length does not depend on his capacity and customer's capacity has the same distribution for both systems.

But if customer's length depends on his capacity, then service discipline has an influence on loss characteristics of the system. This influence depends on the character of this dependence and the value of ρ, but is inessential for small ρ. We demonstrate this fact in tables 3, 4, and 5 (in Appendix), when customer's length is proportional to his capacity and the capacity has an exponential distribution ($\xi = c\zeta$, $c = 1$, $\varphi_1 = \mathbf{E}\zeta = 1$).

It is interesting to compare the last results with those for the case of non-exponential customer volume and length distribution. We present them in tables 6–9 for independent random variables ζ and ξ having the uniform distribution on $[0; 2]$ (see tables 6, 7) and for the case when customer's length ξ is proportional to his capacity ζ (having the same distribution) with coefficient $c = 1$ (see tables 8 and 9).

4. Conclusion

In this paper we have analyzed the influence of service discipline on the first moment of total customers capacity in single-server queueing system with unlimited system capacity and on the loss characteristics for the system with limited total capacity. It was shown that

1) the discipline FIFO is better than EPS from the viewpoint of capacity occupied by customers in the system and loss characteristics;

2) the loss characteristics P and Q depend on service discipline and character of dependence between customer's capacity and his length.

However, the last dependence is inessential for rather small system capacities and small ρ; more precisely, in this case the influence of ζ and ξ dependence is inessential for loss characteristics calculation. Therefore, in practice we often need not to pay attention on this dependence and can use analytical methods to calculate the loss probability for queueing systems with customer length not depending on his capacity.

References

[1] A. M. Alexandrov, B. A. Kaz. Non-homogeneous demands flow service. *Izvestiya AN SSSR. Tekhnicheskaya Kibernetika*, No 2, 47–53, 1973. (In Russian).

[2] B. Sengupta. The spatial requirement of an $M/G/1$ queue or: how to design for buffer space. *Lect. Notes Contr. Inf. Sci.*, **60**, 547–562, 1984.

[3] O. M. Tikhonenko. *Queueing Models in Computer Systems.* Universitetskoe, Minsk, 1990. (In Russian).

[4] O. Tikhonenko. *Metody probabilistyczne analizy systemów informacyjnych.* Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2006.

[5] S. F. Yashkov. *Analysis of Queues in Computers.* Radio i Svyaz, Moskow, 1989. (In Russian).

[6] O. Tikhonenko. Classical and non-classical processor sharing systems with non-homogeneous customers. *Scientific Issues of Jan Długosz University in Częstochowa. Ser. Mathematics*, **XIV**, 133–150, 2009.

[7] O. Tikhonenko, A. Gola, M. Ziółkowski. Estimations of loss characteristics of single-server queueing systems with non-homogeneous customers. *Scientific Issues of Jan Długosz University in Częstochowa. Ser. Mathematics*, **XIII**, 53–66, 2008.

Appendix

Table 1: Probabilities Q for $\rho = 0.2$

V	Q_{sim}^{FIFO}	Q_{sim}^{ESP}
0.0	1.0000	1.0000
1.0	0.7526	0.7526
2.0	0.4454	0.4451
2.5	0.3290	0.3289
3.0	0.2402	0.2401
4.0	0.1230	0.1231
5.0	0.0610	0.0607
6.0	0.0293	0.0295
8.0	0.0067	0.0069
10.0	0.0014	0.0014
12.0	0.0003	0.0003

Table 2: Probabilities Q for $\rho = 0.8$

V	Q_{sim}^{FIFO}	Q_{sim}^{EPS}
0.0	1.0000	1.0000
2.0	0.5774	0.5771
4.0	0.3083	0.3077
6.0	0.1776	0.1772
8.0	0.1085	0.1085
10.0	0.0688	0.0678
15.0	0.0233	0.0232
20.0	0.0085	0.0085
25.0	0.0031	0.0031
30.0	0.0012	0.0011
35.0	0.0004	0.0004

Table 3: Probabilities P and Q for $\rho = 0.2$

V	P_{sim}^{FIFO}	Q_{sim}^{FIFO}	P_{an}^{EPS}	Q_{sim}^{EPS}
0.0	1.0000	1.0000	1.0000	1.0000
1.0	0.3846	0.7454	0.3847	0.7454
2.0	0.1711	0.4445	0.1718	0.4455
3.0	0.0850	0.2531	0.0866	0.2549
4.0	0.0446	0.1419	0.0467	0.1448
5.0	0.0240	0.0790	0.0260	0.0824
6.0	0.0129	0.0435	0.0147	0.0469
8.0	0.0038	0.0130	0.0048	0.0153
10.0	0.0011	0.0038	0.0016	0.0051
12.0	0.0003	0.0011	0.0005	0.0017
15.0	0.0001	0.0002	0.0001	0.0003

Table 4: Probabilities P and Q for $\rho = 0.8$

V	P_{sim}^{FIFO}	Q_{sim}^{FIFO}	P_{an}^{EPS}	Q_{sim}^{EPS}
0.0	1.0000	1.0000	1.0000	1.0000
2.0	0.2570	0.5371	0.2641	0.5424
4.0	0.1272	0.2851	0.1475	0.3111
6.0	0.0715	0.1642	0.0964	0.2036
8.0	0.0429	0.0997	0.0676	0.1426
10.0	0.0267	0.0624	0.0493	0.1042
15.0	0.0090	0.0212	0.0248	0.0525
20.0	0.0032	0.0076	0.0135	0.0285
25.0	0.0012	0.0028	0.0076	0.0160
30.0	0.0004	0.0010	0.0044	0.0092
35.0	0.0002	0.0004	0.0025	0.0054
40.0	0.0001	0.0002	0.0015	0.0031
50.0	0.0000	0.0000	0.0005	0.0011

Table 5: Probabilities P and Q for $\rho = 1.0$

V	P_{sim}^{FIFO}	Q_{sim}^{FIFO}	P_{an}^{EPS}	Q_{sim}^{EPS}
0.0	1.0000	1.0000	1.0000	1.0000
2.0	0.2803	0.5616	0.2902	0.5687
4.0	0.1550	0.3303	0.1819	0.3624
5.0	0.1229	0.2654	0.1539	0.3070
10.0	0.0551	0.1220	0.0870	0.1736
15.0	0.0344	0.0764	0.0606	0.1212
20.0	0.0256	0.0572	0.0465	0.0929
30.0	0.0132	0.0290	0.0317	0.0635
35.0	0.0105	0.0246	0.0248	0.0548
40.0	0.0076	0.0217	0.0241	0.0481
50.0	0.0070	0.0164	0.0195	0.0389
60.0	0.0065	0.0144	0.0163	0.0325
70.0	0.0056	0.0120	0.0140	0.0279
80.0	0.0049	0.0104	0.0122	0.0250

Table 6: Probabilities P and Q for $\rho = 0,2$ when ζ and ξ are independent

V	P_{sim}^{FIFO}	Q_{sim}^{FIFO}	P_{an}^{EPS}	Q_{sim}^{EPS}
0.0	1.0000	1.0000	1.0000	1.0000
0.5	0.7565	0.9397	0.7565	0.9397
1.0	0.5240	0.7658	0.5241	0.7658
1.5	0.3031	0.4904	0.3039	0.4908
2.0	0.0923	0.1226	0.0938	0.1240
2.5	0.0577	0.0846	0.0603	0.0873
3.0	0.0314	0.0492	0.0344	0.0528
4.0	0.0051	0.0076	0.0079	0.0113
5.0	0.0012	0.0018	0.0023	0.0034
6.0	0.0002	0.0003	0.0006	0.0009
7.0	0.0000	0.0001	0.0002	0.0002

Table 7: Probabilities P and Q for $\rho = 0,8$ when ζ and ξ are independent

V	P_{sim}^{FIFO}	Q_{sim}^{FIFO}	P_{an}^{EPS}	Q_{sim}^{EPS}
0.0	1.0000	1.0000	1.0000	1.0000
1.0	0.5854	0.8065	0.5878	0.8075
2.0	0.3008	0.3944	0.3111	0.4016
3.0	0.1857	0.2622	0.2083	0.2854
4.0	0.1153	0.1605	0.1435	0.1936
5.0	0.0759	0.1068	0.1044	0.1418
6.0	0.0515	0.0723	0.0779	0.1057
8.0	0.0250	0.0352	0.0455	0.0618
10.0	0.0127	0.0178	0.0280	0.0380
15.0	0.0024	0.0034	0.0090	0.0122
20.0	0.0005	0.0007	0.0030	0.0040
25.0	0.0001	0.0001	0.0010	0.0014
30.0	0.0000	0.0000	0.0003	0.0004
35.0			0.0001	0.0001

Table 8: Probabilities P and Q for $\rho = 0,2$ when ξ is proportional to ζ

V	P_{sim}^{FIFO}	Q_{sim}^{FIFO}	P_{an}^{EPS}	Q_{sim}^{EPS}
0.0	1.0000	1.0000	1.0000	1.0000
1.0	0.5158	0.7597	0.5158	0.7596
2.0	0.1141	0.1423	0.1150	0.1431
3.0	0.0438	0.0686	0.0484	0.0740
4.0	0.0079	0.0112	0.0138	0.0192
5.0	0.0020	0.0031	0.0049	0.0072
6.0	0.0003	0.0005	0.0015	0.0022
7.0	0.0001	0.0001	0.0005	0.0007
8.0	0.0000	0.0000	0.0002	0.0002
9.0			0.0001	0.0001

Table 9: Probabilities P and Q for $\rho = 0,8$ when ξ is proportional to ζ

V	P_{sim}^{FIFO}	Q_{sim}^{FIFO}	P_{an}^{EPS}	Q_{sim}^{EPS}
0.0	1.0000	1.0000	1.0000	1.0000
0.5	0.7582	0.9400	0.7582	0.9400
1.0	0.5570	0.7855	0.5573	0.7857
2.0	0.3179	0.3963	0.3252	0.4023
3.0	0.1832	0.2621	0.2142	0.2956
4.0	0.1109	0.1536	0.1558	0.2084
5.0	0.0703	0.0997	0.1168	0.1578
6.0	0.0464	0.0658	0.0901	0.1221
8.0	0.0219	0.0310	0.0569	0.0770
10.0	0.0110	0.0155	0.0376	0.0508
15.0	0.0021	0.0030	0.0148	0.0200
20.0	0.0004	0.0006	0.0062	0.0084
25.0	0.0001	0.0001	0.0027	0.0036
30.0	0.0000	0.0000	0.0011	0.0015
35.0			0.0005	0.0007
40.0			0.0002	0.0003
45.0			0.0001	0.0001