This is a Special issue
of our journal devoted
to the 60th anniversary
of the birth of
Efim Zelmanov

The aim of the journal **"Algebra and Discrete Mathematics"** (as $\mathcal{ADM}$ below) is to present timely the state-of-the-art accounts on modern research in all areas of algebra (general algebra, semigroups, groups, rings and modules, linear algebra, algebraic geometry, universal algebras, homological algebra etc.) and discrete mathematics (combinatorial analysis, graphs theory, mathematical logic, theory of automata, coding theory, cryptography etc.)

Languages:

Papers to be considered for publication in the $\mathcal{ADM}$ journal must be written in English.

Preparing papers:

Papers submitted to the $\mathcal{ADM}$ journal should be prepared in LaTeX. Authors are strongly encourages to prepare their papers using the $\mathcal{ADM}$ author package containing template, instructions, and $\mathcal{ADM}$ journal document class. $\mathcal{ADM}$ author package is available from the journal web-site `http://adm.luguniv.edu.ua`.

Graphical items should be prepared as eps (encapsulated PostScript) files and included by using the `graphicx` package. To avoid distortion from rescaling, figures must not be wider than 120 mm.

Submitting papers:

Authors who wish to submit their papers should send paper in PDF format directly to anyone of the Editors via electronic mail. E-mails of the Editors are listed on the Editorial Board page.

The source TeX-file of the paper will be needed if it accepted for publication.

Submission of a manuscript implies that the work described has not been published before and that it is not under consideration for publication elsewhere.

Required information:

The following information is required with the submission (note that all contact information, particularly email addresses, must be supplied to avoid delay):

1) *full postal address of each author;*
2) *e-mail of each author;*
3) *abstract (no more than 15 lines);*
4) *2010 Mathematics subject classification (can be accessible from* `http://www.ams.org/msc`*);*
5) *key words and phrases.*

Proof-sheets:

Authors receive only one set of proof-sheets in PDF format via e-mail for corrections. Only correction of misprints and minor changes can be made during proofreading.

Efim Zelmanov

To the 60th anniversary

On September 7, 2015, a distinguished mathematician, one of the founders of our journal, Efim Isaakovich Zelmanov, turned 60. He was born in Khabarovsk, Soviet Union (now Russian Federation), while his mother grew up in the Ukrainian city Zhytomyr.

Efim's impressive mathematical abilities appeared in his school time. He attended Novosibirsk State University, obtaining his Master's degree in 1977. He received his Ph.D. from Novosibirsk State University in 1980 having had his research supervised by the prominent algebraists Professors L.A. Bokut and A.I. Shirshov. He defended his Doctor of Sciences dissertation (habilitation) at Leningrad (St. Petersburg) State University in 1985. In 1980–1989, Efim Zelmanov held research positions (by increasing levels: Junior, Senior, and Leading Researcher) at the Institute of Mathematics of the USSR Academy of Science at Novosibirsk (Academgorodok). In 1989–1992, he worked for different universities in the USA, Canada, Germany, and UK. In 1990, Zelmanov was appointed a professor at the University of Wisconsin-Madison in the United States. In 1994, he was appointed to the University of Chicago. In 1995–2002, he held a professorship at Yale University. In 2002, Efim Zelmanov was appointed as the Rita L. Atkinson Chair in Mathematics at University of California, San Diego. His honors include: Fields Medal (1994), College de France Medal (1991), and Andre Aizenstadt Prize (1996). Efim Zelmanov was elected to American Academy of Arts and Sciences (1996), the U.S. National Academy (2001), he is a Fellow of the American Mathematical Society (2012). He is Foreign Member of the Spanish Royal Academy of Sciences (1997), of the Korean Academy of Sciences and Technology (2008), and of the Brazilian Academy of Sciences (2012). He was awarded by the Honorary Doctor degree in Hagen (Germany), Oviedo (Spain), and Kyiv (Ukraine) Universities.

Professor Zelmanov was invited to speak at the International Congresses of Mathematicians in Warsaw (1983), Kyoto (1990), and Zurich (1994).

The thesis Efim Zelmanov presented for his Ph.D. was on nonassociative algebra, namely on Jordan algebras in the infinite dimensional case. He showed that Glennie's identity generates (in a certain sense) all identities that hold in the algebra. This and his consequent works completely changed the entire content of Jordan algebras. He was able to extend the known results from the classical theory of finite dimensional Jordan algebras to infinite dimensional Jordan algebras. Zelmanov's results on Jordan algebras were presented in his invited lecture at the International Congress of Mathematicians at Warsaw in 1983.

Lie rings were the next step in the study of non-associative rings. In 1987 Zelmanov solved one of the most famous open questions in the theory of Lie algebras at that time. He proved that the Engel identity $\mathrm{ad}^n(y) = 0$ implies that the algebra is necessarily nilpotent. Similar to the case of Jordan algebras, Zelmanov was able to extend important properties of finite dimensional Lie algebras to the infinite dimensional case.

The mentioned results (and the results obtained by Zelmanov later) on Lie and Jordan algebras dramatically changed the theory of non-associative algebras. They made Efim Zelmanov a leading expert in non-associative algebras. He and his coauthors were able to also make a fundamental contribution to associative algebras, super-algebras, associated modules and representations. Self-similar algebras and growth of algebras are among of the topics of Zelmanov's recent research.

In 1991, Zelmanov made one more significant step in his mathematical career by solving the famous Restricted Burnside Problem. This problem has its roots in one of the most remarkable mathematical problems known as the Burnside Problem introduced by Burnside in 1902. A version of this problem, formulated by Magnus in the 1930's is called the Restricted Burnside Problem. Prominent mathematicians such as Hall, Higman, Kostrikin and many others put significant efforts toward solving this problem. Using previously known results and his own results on Lie and Jordan algebras, Efim Zelmanov obtained a complete solution to the problem, which made a significant impact on the subsequent development of group theory. This constitutes a remarkable example of the effectiveness of the applications of ring theory and, more generally, of purely algebraic methods, to group theory. The result of Zelmanov yields that if a group G is finitely generated, residually finite, and satisfies the identity $X^n = 1$,

then it is finite. Thus in the residually finite case (which is one of the most important cases in applications) the situation is completely opposite to the situation in the case of arbitrary finitely generated groups (the Burnside Problem was solved by Adjan and Novikov in 1967).

In 1994, Zelmanov was awarded the Fields Medal for his works on Lie and Jordan algebras and on the solution of the Restricted Burnside Problem.

In 1991, Efim Zelmanov began his work on pro-p-groups. These groups play a crucial role in Number Theory because the Galois groups of field extensions are profinite groups and the primary p-case is the most significant via its relation to p-adic fields. In this area, very soon Efim Zelmanov obtained his remarkable result. He solved the Platonov Problem, a version of the Burnside Problem for compact topological groups. He proved that a compact torsion group is locally finite. After this work, pro-p groups began playing a significant role in Zelmanov's research.

Graded algebras, constructions of Golod-Shavarevich type, Kac-Moody algebras and their subalgebras, superalgebra versions of Lie, Jordan and other type of algebras, modules over them, representations, growth, etc., is a broad spectrum of the topics included in the research interest of E. Zelmanov. In 2010 he came up with the idea of self-similarity in theory of rings and successfully implemented it. The field of his interests goes far beyond algebra. It includes theoretical physics, random processes, discrete mathematics and much more.

Additionally to his research and teaching, Efim Zelmanov performs tremendous outreach activity. The broader impact of his dedication to the international mathematics community is difficult to overestimate. He served and is serving numerous national and international important committees (including those for the assignment of the Field Medal and the Abel Prize). He plays an important role in mathematical life in many countries, including USA, Germany, China, France, UK, South Korea, Brazil, and many others. Efim Zelmanov plays a tremendous role in supporting mathematical life in Ukraine. He regularly participates in and helps to organize mathematics conferences that take place in Ukraine, and in various ways supports many Ukrainian mathematicians.

E. Zelmanov's contribution to mathematics goes far beyond his remarkable research, teaching and outreach achievements. In different periods of time, as an editor of many major mathematics journals including 'The Annals of Mathematics', 'The Journal of Algebra', 'The Journal of the American Mathematical Society', 'The Bulletin of Mathematical Science' for which he is the Editor in Chief, 'Groups Geometry and Dynamics',

and 'Algebra and Discrete Mathematics', he strongly raised the bar of the quality of publications.

Professor Zelmanov is a great speaker and lecturer. He is one of the most popular presenters not only in the USA but in the world. He has a rare and excellent type of humor which makes his presentations and communication with him a great pleasure for everybody. He is a very caring person and wonderful friend, always ready to extend his help and support to his numerous friends and colleagues.

Efim Zelmanov has a wonderful family. He and his lovely wife Lena are loving and caring parents and grandparents.

Professor Zelmanov is one of the top researchers in the world, distinguished leader of the world mathematics community, and one of the most pleasant and nice people.

We most warmly congratulate him and wish him Siberian health, much happiness, new great discoveries, and wonderful students.

The Editorial Board
of Algebra and Discrete
Mathematics Journal

Algebra and Discrete Mathematics
Volume **20** (2015). Number 1, pp. 1–12

RESEARCH ARTICLE

Universal property of skew PBW extensions

Juan Pablo Acosta and Oswaldo Lezama

Communicated by V. A. Artamonov

Abstract. In this paper we prove the universal property of skew PBW extensions generalizing this way the well known universal property of skew polynomial rings. For this, we will show first a result about the existence of this class of non-commutative rings. Skew PBW extensions include as particular examples Weyl algebras, enveloping algebras of finite-dimensional Lie algebras (and its quantization), Artamonov quantum polynomials, diffusion algebras, Manin algebra of quantum matrices, among many others. As a corollary we will give a new short proof of the Poincaré-Birkhoff-Witt theorem about the bases of enveloping algebras of finite-dimensional Lie algebras.

1. Introduction

Most of constructions in algebra are characterized by universal properties from which it is easy to prove important results about the constructed object. This is the case of the universal property of the tensor product; another well known example is the universal property for the localization of rings and modules by multiplicative subsets. A key example in non-commutative algebra is the skew polynomial ring $R[x;\sigma,\delta]$; the universal property in this case says that if B is a ring with a ring homomorphism $\varphi : R \to B$ and in B there exists and element y such that $y\varphi(r) = \varphi(\sigma(r))y + \varphi(\delta(r))$ for every $r \in R$, then there exists an

2010 MSC: Primary: 16S10, 16S80; Secondary: 16S30, 16S36.
Key words and phrases: skew polynomial rings, skew PBW extensions, PBW bases, quantum algebras.

unique ring homomorphism $\widetilde{\varphi} : R[x;\sigma,\delta] \to B$ such that $\widetilde{\varphi}(x) = y$ and $\widetilde{\varphi}(r) = \varphi(r)$ (see [9]). In this paper we prove the universal property of skew PBW extensions generalizing the universal property of skew polynomial rings. For this, we will prove first a theorem about the existence of skew PBW extensions similar to the corresponding result on skew polynomial rings. As application we will get the Poincaré-Birkhoff-Witt theorem about the bases of enveloping algebras of finite-dimensional Lie algebras. This famous theorem says that if K is a field and $\mathcal{G}$ is a finite-dimensional Lie algebra with K-basis $\{y_1, \ldots, y_n\}$, then a K-basis of the universal enveloping algebra $\mathcal{U}(\mathcal{G})$ is the set of monomials $y_1^{\alpha_1} \cdots y_n^{\alpha_n}$, $\alpha_i \geqslant 0$, $1 \leqslant i \leqslant n$ (see [4], [6]).

Skew PBW extensions were defined firstly in [7], and their homological and ring-theoretic properties have been studied in the last years (see [1], [3], [8], [10]). Skew polynomial rings of injective type, Weyl algebras, enveloping algebras of finite-dimensional Lie algebras (and its quantization), Artamonov quantum polynomials, diffusion algebras, Manin algebra of quantum matrices, are particular examples of skew PBW extensions (see [8]). In this first section we recall the definition of skew PBW extensions and some very basic properties needed for the proof of the main theorem.

Definition 1.1. Let R and A be rings. We say that A is a *skew PBW extension of R* (also called a $\sigma - PBW$ extension of R) if the following conditions hold:

(i) $R \subseteq A$.

(ii) There exist finite elements $x_1, \ldots, x_n \in A$ such A is a left R-free module with basis

$$\text{Mon}(A) := \{x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n} \mid \alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n\}.$$

In this case it says also that A *is a left polynomial ring over R* with respect to $\{x_1, \ldots, x_n\}$ and $Mon(A)$ is the set of standard monomials of A. Moreover, $x_1^0 \cdots x_n^0 := 1 \in Mon(A)$.

(iii) For every $1 \leqslant i \leqslant n$ and $r \in R - \{0\}$ there exists $c_{i,r} \in R - \{0\}$ such that

$$x_i r - c_{i,r} x_i \in R. \tag{1.1}$$

(iv) For every $1 \leqslant i, j \leqslant n$ there exists $c_{i,j} \in R - \{0\}$ such that

$$x_j x_i - c_{i,j} x_i x_j \in R + Rx_1 + \cdots + Rx_n. \tag{1.2}$$

Under these conditions we will write $A := \sigma(R)\langle x_1, \ldots, x_n\rangle$.

The following proposition justifies the notation and the alternative name given for the skew PBW extensions.

Proposition 1.2. Let A be a skew PBW extension of R. Then, for every $1 \leqslant i \leqslant n$, there exists an injective ring endomorphism $\sigma_i : R \to R$ and a σ_i-derivation $\delta_i : R \to R$ such that

$$x_i r = \sigma_i(r)x_i + \delta_i(r),$$

for each $r \in R$.

Proof. See [7], Proposition 3. □

Observe that if σ is an injective endomorphism of the ring R and δ is a σ-derivation, then the skew polynomial ring $R[x; \sigma, \delta]$ is a trivial skew PBW extension in only one variable, $\sigma(R)\langle x\rangle$.

Some extra notation will be used in the rest of the paper.

Definition 1.3. Let A be a skew PBW extension of R with endomorphisms σ_i, $1 \leqslant i \leqslant n$, as in Proposition 1.2.

(i) For $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, $\sigma^\alpha := \sigma_1^{\alpha_1} \cdots \sigma_n^{\alpha_n}$, $|\alpha| := \alpha_1 + \cdots + \alpha_n$. If $\beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}^n$, then $\alpha + \beta := (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n)$.

(ii) For $X = x^\alpha \in \mathrm{Mon}(A)$, $\exp(X) := \alpha$ and $\deg(X) := |\alpha|$.

(iii) If $f = c_1X_1 + \cdots + c_tX_t$, with $X_i \in Mon(A)$ and $c_i \in R - \{0\}$, then $\deg(f) := \max\{\deg(X_i)\}_{i=1}^t$.

The skew PBW extensions can be characterized in a similar way as was done in [5] for PBW rings.

Theorem 1.4. Let A be a left polynomial ring over R w.r.t. $\{x_1, \ldots, x_n\}$. A is a skew PBW extension of R if and only if the following conditions hold:

(a) For every $x^\alpha \in \mathrm{Mon}(A)$ and every $0 \neq r \in R$ there exist unique elements $r_\alpha := \sigma^\alpha(r) \in R - \{0\}$ and $p_{\alpha,r} \in A$ such that

$$x^\alpha r = r_\alpha x^\alpha + p_{\alpha,r}, \tag{1.3}$$

where $p_{\alpha,r} = 0$ or $\deg(p_{\alpha,r}) < |\alpha|$ if $p_{\alpha,r} \neq 0$. Moreover, if r is left invertible, then r_α is left invertible.

(b) For every $x^{\alpha}, x^{\beta} \in \text{Mon}(A)$ there exist unique elements $c_{\alpha,\beta} \in R$ and $p_{\alpha,\beta} \in A$ such that

$$x^{\alpha}x^{\beta} = c_{\alpha,\beta}x^{\alpha+\beta} + p_{\alpha,\beta}, \tag{1.4}$$

where $c_{\alpha,\beta}$ is left invertible, $p_{\alpha,\beta} = 0$ or $\deg(p_{\alpha,\beta}) < |\alpha + \beta|$ if $p_{\alpha,\beta} \neq 0$.

Proof. See [7], Theorem 7. □

2. Existence theorem for skew PBW extensions

If $A = \sigma(R)\langle x_1, \ldots, x_n\rangle$ is a skew PBW extension of the ring R, then as was observed in the previous section, A induces unique endomorphisms $\sigma_i : R \to R$ and σ_i-derivations $\delta_i : R \to R$, $1 \leqslant i \leqslant n$. Moreover, by (1.2), there exist $c_{ij}, d_{ij}, a_{ij}^{(k)} \in R$ such that $x_jx_i = c_{ij}x_ix_j + a_{ij}^{(1)}x_1 + \cdots + a_{ij}^{(n)}x_n + d_{ij}$, with $1 \leqslant i, j \leqslant n$. However, note that if $i < j$, since $Mon(A)$ is a R-basis, then $1 = c_{j,i}c_{i,j}$, i.e., for every $1 \leqslant i < j \leqslant n$, c_{ji} is a right inverse of $c_{i,j}$ univocally determined. In a similar way, we can check that $a_{ji}^{(k)} = -c_{ji}a_{ij}^{(k)}$, $d_{ji} = -c_{ji}d_{ij}$. Thus, given A there exist unique parameters $c_{ij}, d_{ij}, a_{ij}^{(k)} \in R$ such that

$$x_jx_i = c_{ij}x_ix_j + a_{ij}^{(1)}x_1 + \cdots + a_{ij}^{(n)}x_n + d_{ij}, \text{ for every } 1 \leqslant i < j \leqslant n. \tag{2.1}$$

Definition 2.1. Let $A = \sigma(R)\langle x_1, \ldots, x_n\rangle$ be a skew PBW extension. $\sigma_i, \delta_i, c_{ij}, d_{ij}, a_{ij}^{(k)}$, $1 \leqslant i < j \leqslant n$, defined as before, are called the parameters of A.

Conversely, given a ring R and parameters $\sigma_i, \delta_i, c_{ij}, d_{ij}, a_{ij}^{(k)}$, $1 \leqslant i < j \leqslant n$, we will construct in this section a skew PBW extension with coefficient ring R and satisfying the following equations

1) For $i < j$ in I and k in I, $x_jx_i = c_{ij}x_ix_j + \Sigma_k a_{ij}^{(k)}x_k + d_{ij}$,
2) For $i \in I$ and $r \in R$, $x_ir = \sigma_i(r)x_i + \delta_i(r)$,

where $I := \{1, \ldots, n\}$.

Definition 2.2. Let R be a ring and W be the free monoid in the alphabet $X \cup R$, with $X := \{x_i : i \in I\}$. Let w be a word of W, the complexity of w, denoted $c(w)$, is a triple of nonnegative integers (a, b, c), where a is the number of x's in w, b is is the number of inversions involving only x's, and c is the number of inversions of the type (x_i, r).

These triples are ordered with the lexicographic order, i.e., $(a,b,c) \leqslant (d,e,f)$ if and only if $a < d$, or, $a = d$ and $b < e$, or, $a = d$, $b = e$ and $c \leqslant f$. This is a well order. Let T be the set of elements of W such that $c(w) = (a,0,0)$ and $\mathbb{Z}T$ be the linear extension of T in $\mathbb{Z}\langle X \cup R\rangle$ (the $\mathbb{Z}$-free algebra in the alphabet $X \cup R$).

Definition 2.3. Let R be a ring, $\{c_{ij}\}_{i<j}$, $\{d_{ij}\}_{i<j}$ and $\{a_{ij}^{(k)}\}_{i<j,k}$ be elements of R indexed by i,j,k in I. Let $\sigma_i, \delta_i : R \to R$ be two functions for each $i \in I$. Suppose that c_{ij} is left invertible and that $\sigma_i(r) \neq 0$ for $r \neq 0$. We define the function p

$$p : W \to \mathbb{Z}\langle X \cup R\rangle, \quad \text{with } X := \{x_i : i \in I\},$$

by induction in the complexity, as follows:

1) If $w \in T$ then $p(w) = w$.
2) If $w = v_1 x_i r v_2$, with $r \in R$, $v_1 \in W$ and $rv_2 \in T$ then

$$p(w) = p(v_1 \sigma_i(r) x_i v_2) + p(v_1 \delta_i(r) v_2).$$

3) If $w = v_1 x_j x_i v_2$, where $v_1 \in W$, $x_i v_2 \in T$ with $i < j$, then

$$p(w) = p(v_1 c_{ij} x_i x_j v_2) + \Sigma_k p(v_1 a_{ij}^{(k)} x_k v_2) + p(v_1 d_{ij} v_2).$$

The linear extension of p to $\mathbb{Z}\langle X \cup R\rangle \to \mathbb{Z}\langle X \cup R\rangle$ is also denoted p. The image of p is contained in $\mathbb{Z}T$. Let $Mon := \{\Pi_{k=1}^{n} x_{i_k} : i_1 \leqslant \cdots \leqslant i_n, n \geqslant 0\}$, and $F_R(Mon)$ be the left free $R-$module with basis Mon. We define $q : \mathbb{Z}T \to F_R(Mon)$ as the bilinear extension of $q(r_1 \ldots r_m x_{i_1} \ldots x_{i_n}) := (\Pi_{k=1}^{m} r_k) x_{i_1} \ldots x_{i_n}$. Finally, we define $h : \mathbb{Z}\langle X \cup R\rangle \to F_R(Mon)$ as $h := qp$.

Theorem 2.4 (Existence)**.** Let $R, I, X, a_{ij}^{k}, c_{ij}, \sigma_i, \delta_i, h, p, q$ be as in Definition 2.3. Then, there exists a skew PBW extension A of R with variables $X := \{x_i : i \in I\}$ such that

(a) $x_i r = \sigma_i(r) x_i + \delta_i(r)$.
(b) $x_j x_i = c_{ij} x_i x_j + \Sigma_k a_{ij}^{(k)} x_k + d_{ij}$, for $i < j$ in I.

if and only if

(1) For every i in I, σ_i is a ring endomorphism of R and δ_i is σ_i-derivation.

(2) $h(x_jx_ir) = h(p(x_jx_i)r)$, for $i < j$ in I and $r \in R$.

(3) $h(x_kx_jx_i) = h(p(x_kx_j)x_i)$, for $i < j < k$ in I.

Proof. ($\Longrightarrow$) Numeral (1) is the content of Proposition 1.2. Conditions (2) and (3) follow from (a) and (b) and the associativity $x_j(x_ir) = (x_jx_i)r$ and $x_k(x_jx_i) = (x_kx_j)x_i$.

($\Longleftarrow$) Define $t : F_R(Mon) \to \mathbb{Z}\langle X \cup R\rangle$ as $t(\Sigma r_{\bar{x}}\bar{x}) := \Sigma r_{\bar{x}}\bar{x} \in \mathbb{Z}\langle X \cup R\rangle$, where $\Sigma r_{\bar{x}}\bar{x}$ is the unique expression of an element in $F_R(Mon)$ as a sum over a finite set, $\bar{x} \in Mon$ and $r_{\bar{x}} \neq 0$ is an element of R.

We define a product in $F_R(Mon)$ by

$$f \star g = h(t(f)t(g)), \qquad f, g \in F_R(Mon),$$

and we will prove in Lemma 2.8 below that $h(ab) = h(a) \star h(b)$, with $a, b \in \mathbb{Z}\langle X \cup R\rangle$. From this we get that $h : \mathbb{Z}\langle X \cup R\rangle \to F_R(Mon)$ is a surjection that preserves sums, products and $h(1) = 1$. This makes $F_R(Mon)$ a ring, which is a skew PBW extension of R by the definition of the product $\star$.

To complete the proof we proceed to prove Lemma 2.8, but for this, we have to show first some preliminary propositions under the hypothesis (1)-(3). □

Proposition 2.5. For $a, b \in W$ and $r, s \in R$ the following equalities hold:

(i) $h(a0b) = 0$.

(ii) $h(a(-r)b) = -h(arb)$.

(iii) $h(a(r + s)b) = h(arb + asb)$.

(iv) $h(a1b) = h(ab)$.

(v) $h(a(rs)b) = h(arsb)$.

Proof. (i) and (ii) follow from (iii) since $r \mapsto h(arb)$ is a group homomorphism from the additive group of R into $F_R(Mon)$.

(iii) is proven by induction on $c(a(r+s)b)$ and applying the definition of h. Here the conditions $\delta_i(a+b) = \delta_i(a)+\delta_i(b)$ and $\sigma_i(a+b) = \sigma_i(a)+\sigma_i(b)$ in the hipothesis (1) of Theorem 2.4 are used.

(iv) is proven by induction on $c(a1b)$ and making use of part (i). The relevant hypothesis are $\sigma_i(1) = 1$ and $\delta_i(1) = 0$ which are part of the hypothesis (1) in Theorem 2.4.

(v) This part is proven by induction on $c(a(rs)b)$ and making use of (iii). The relevant hypothesis are $\sigma_i(ab) = \sigma_i(a)\sigma_i(b)$ and $\delta_i(ab) = \sigma_i(a)\delta_i(b) + \delta_i(a)b$. □

Proposition 2.6. Let $y, z \in \mathbb{Z}\langle X \cup R\rangle$ and $a \in \mathbb{Z}T$. Then $h(yaz) = h(ytq(a)z)$.

Proof. This is because we can obtain $tq(a)$ from a with a finite number of operations described in Proposition 2.5. Indeed if $a \in \mathbb{Z}T$ then by definition of T we heave $a = \Sigma n_u u$ where the sum is over $u \in T$, $n_u \in \mathbb{Z}$ and $u = r_{1,u} \dots r_{m,u} x_{j_1} \dots x_{j_k}$ ($j_1, \dots j_k$ and m, k depend on u) here $r_s \in R$ and $1 \leqslant j_1 \leqslant \dots \leqslant j_k \leqslant n$. Then by definition of t, q we have $tq(a) = \Sigma_{x \in A} a(x)x$ where $A = \{x \in Mon(X) : a(x) \neq 0\}$, and $a(x) = \Sigma_{u \in B(x)} n_u \Pi_s r_{s,u} \in R$ where $B(x) = \{u \in T : x_{j_1} \dots x_{j_k} = x\}$. Using the Proposition 2.5 (i) we obtain that

$$h(ytq(a)z) = h(y\Sigma_{x \in Mon(X)} a(x)xz).$$

Using that h is linear we get

$$h(y\Sigma_{x \in Mon(X)} a(x)xz) = \Sigma_{x \in Mon(x)} h(ya(x)xz).$$

Using Proposition 2.5 (i),(ii),(iii) we get that

$$h(ya(x)xz) = \Sigma_{u \in B(x)} n_u h(y(\Pi_s r_{s,u})xz).$$

Further, using Proposition 2.5 (iv)(v) we get that

$$h(y(\Pi_s r_{s,u})xz) = h(yr_{1,u} \dots r_{m,u} xz) = h(yuz). \qquad \square$$

Proposition 2.7. If $x, y, z \in \mathbb{Z}\langle X \cup R\rangle$ then $h(xp(y)z) = h(xyz)$.

Proof. The identity is linear in x, y, z, so we may assume they are words. Next we proceed by induction on $c(xyz)$. First assume that the first inversion from right to left in xyz is in y, say $y = w_1 x_j s w_2$ with $s = x_i$ with $i < j$ or $s \in R$, and $sw_2 \in T$. Then

$$h(xyz) = h(xw_1 p(x_j s) w_2 z) = h(xp(w_1 p(x_j s) w_2)z) = h(xp(y)z)$$

by the definition of p and induction.

Now assume that the first inversion of xyz is not contained in yz, or $xyz \in T$, in this case $y \in T$ and $p(y) = y$.

Next, assume that the first inversion of xyz is contained in z say $z = w_1 x_j s w_2$ with $sw_2 \in T$ and $s = x_i$ with $i < j$ or $s \in R$. Then

$$h(xyz) = h(xyw_1 p(x_j s) w_2) = h(xp(y)w_1 p(x_j s) w_2) = h(xp(y)z)$$

by definition of h and induction.

Now assume that the first inversion of xyz has a part in y and a part in z, say $y = y'x_j$ and $z = sz'$ with $z \in T$ and $s = x_i$ with $i < j$ or $s \in R$. Assume further that the first inversion of y exists and is contained in y', say $y' = w_1x_ks'w_2$ with $s'w_2 \in T$ an $s' = x_i$ with $i < k$ or $s' \in R$. Then

$$\begin{aligned} h(xyz) &= h(xy'p(x_js)z') = h(xp(y')p(x_js)z') \\ &= h(xp(w_1p(x_ks')w_2)p(x_js)z') = h(xw_1p(x_ks')w_2p(x_js)z') \\ &= h(xw_1p(x_ks')w_2x_jsz') = h(xp(w_1p(x_ks')w_2x_j)sz') \\ &= h(xp(y)z) \end{aligned}$$

by definition of h and induction applied alternatively. So the last case is $y = y'x_kx_j$ with $k > j$ and $z = sz'$ with $s = x_i$ with $i < j$ or $s \in R$ and $z \in T$. In this case

$$h(xyz) = h(xy'x_kp(x_js)z') = h(xy'p(x_kp(x_js))z')$$

by definition of h and induction, also observe

$$h(xy'p(x_kp(x_js))z') = h(xy'p(p(x_kx_j)s)z')$$

because $qp(p(x_kx_j)s) = qp(x_kp(x_js)$ by hipothesis (2) and (3) in Theorem 2.4, and also by Proposition 2.6. Also

$$h(xy'p(p(x_kx_j)s)z') = h(xy'p(x_kx_j)sz') = h(xp(y'p(x_kx_j))z)$$

by induction applied twice, and $h(xp(y'p(x_kx_j))z) = h(xp(y)z)$ by definition of p, as required. □

Lemma 2.8. $h(ab) = h(a) \star h(b)$, for $a, b \in \mathbb{Z}\langle X \cup R\rangle$.

Proof. $h(a) \star h(b) = h(tqp(a)tqp(b)) = h(p(a)p(b)) = h(ab)$, the first equality is from the definition of $\star$, the second equality is from Proposition 2.6 twice and the third equality is Proposition 2.7 twice. □

3. The universal property

In this section we will prove the main theorem about the characterization of skew PBW extensions by a universal property in a similar way as this is done for skew polynomial rings. This problem was studied in [2] where skew PBW extensions were generalized to infinite sets of generators.

Theorem 3.1 (Main theorem: The universal property)**.**

Let $A = \sigma(R)\langle x_1, \dots, x_n\rangle$ be a skew PBW extension with parameters $\sigma_i, \delta_i, c_{ij}, d_{ij}, a_{ij}^{(k)}$, $1 \leqslant i, j \leqslant n$. Let B be a ring with homomorphism $\varphi : R \to B$ and elements $y_1, \dots, y_n \in B$ such that

(i) $y_i\varphi(r) = \varphi(\sigma_i(r))y_i + \varphi(\delta_i(r))$, for every $r \in R$.

(ii) $y_jy_i = \varphi(c_{ij})y_iy_j + \varphi(a_{ij}^{(1)})y_1 + \cdots + \varphi(a_{ij}^{(n)})y_n + d_{ij}$.

Then, there exists an unique ring homomorphism $\widetilde{\varphi} : A \to B$ such that $\widetilde{\varphi}\iota = \varphi$ and $\widetilde{\varphi}(x_i) = y_i$, where ι is the inclusion of R in A.

Proof. Since A is a free R-module with basis $Mon(A)$, we define the R-homomorphism

$$\widetilde{\varphi} : A \to B, \quad r_1x^{\alpha_1} + \cdots + a_tx^{\alpha_t} \mapsto \varphi(r_1)y^{\alpha_1} + \cdots + \varphi(a_t)y^{\alpha_t},$$

where $y^\theta := y_1^{\theta_1} \cdots y_n^{\theta_n}$, with $\theta := (\theta_1, \dots, \theta_n) \in \mathbb{N}^n$. Note that $\widetilde{\varphi}(1) = 1$.

$\widetilde{\varphi}$ is multiplicative: In fact, applying induction on the degree $|\alpha + \beta|$ we have

$$\begin{aligned}
\widetilde{\varphi}(ax^\alpha bx^\beta) &= \widetilde{\varphi}(a[\sigma^\alpha(b)x^\alpha x^\beta + p_{\alpha,b}x^\beta]) \\
&= \widetilde{\varphi}[a\sigma^\alpha(b)[c_{\alpha,\beta}x^{\alpha+\beta} + p_{\alpha,\beta}] + ap_{\alpha,b}x^\beta] \\
&= \varphi(a)\varphi(\sigma^\alpha(b))\varphi(c_{\alpha,\beta})y^{\alpha+\beta} + \varphi(a)\varphi(\sigma^\alpha(b))\varphi(p_{\alpha,\beta})(y) \\
&\quad + \varphi(a)\varphi(p_{\alpha,b})(y)y^\beta,
\end{aligned}$$

where $\varphi(p_{\alpha,\beta})(y)$ is the element in B obtained replacing each monomial x^θ in $p_{\alpha,\beta}$ by y^θ and every coefficient c by $\varphi(c)$. In a similar way we have for $\varphi(p_{\alpha,b})(y)$ (observe that the degree of each monomial of $p_{\alpha,b}x^\beta$ is $< |\alpha + \beta|$). On the other hand, applying (i) and (ii) we get

$$\begin{aligned}
\widetilde{\varphi}(ax^\alpha)\widetilde{\varphi}(bx^\beta) &= \varphi(a)y^\alpha\varphi(b)y^\beta \\
&= \varphi(a)[\varphi(\sigma^\alpha(b))y^\alpha + \varphi(p_{\alpha,b})(y)]y^\beta \\
&= \varphi(a)\varphi(\sigma^\alpha(b))y^\alpha y^\beta + \varphi(a)\varphi(p_{\alpha,b})(y)y^\beta \\
&= \varphi(a)\varphi(\sigma^\alpha(b))[\varphi(c_{\alpha,\beta})y^{\alpha+\beta} + \varphi(p_{\alpha,\beta})(y)] \\
&\quad + \varphi(a)\varphi(p_{\alpha,b})(y)y^\beta \\
&= \varphi(a)\varphi(\sigma^\alpha(b))\varphi(c_{\alpha,\beta})y^{\alpha+\beta} + \varphi(a)\varphi(\sigma^\alpha(b))\varphi(p_{\alpha,\beta})(y) \\
&\quad + \varphi(a)\varphi(p_{\alpha,b})(y)y^\beta.
\end{aligned}$$

It is clear that $\widetilde{\varphi}\iota = \varphi$ and $\widetilde{\varphi}(x_i) = y_i$. Moreover, note that $\widetilde{\varphi}$ is the only ring homomorphism that satisfy these two conditions. □

Corollary 3.2. Let R be a ring and $A = \sigma(R)\langle x_1, \ldots, x_n\rangle$ be a skew PBW extension of R with parameters $\sigma_i, \delta_i, c_{ij}, d_{ij}, a_{ij}^{(k)}$, $1 \leqslant i, j \leqslant n$. Let B be a ring with homomorphism $\varphi : R \to B$ and elements $y_1, \ldots, y_n \in B$ such that the conditions (i)-(ii) in Theorem 3.1 are satisfied with respect to the system of parameters $\sigma_i, \delta_i, c_{ij}, d_{ij}, a_{ij}^{(k)}$, $1 \leqslant i, j \leqslant n$, of the ring R. If B satisfies the universal property, then $B \cong A = \sigma(R)\langle x_1, \ldots, x_n\rangle$. Moreover, the monomials $y_1^{\alpha_1} \cdots y_n^{\alpha_n}$, $\alpha_i \geqslant 0$, $1 \leqslant i \leqslant n$ are a R-basis of B.

Proof. By the universal property of A there exists $\widetilde{\varphi}$ such that $\widetilde{\varphi}\iota = \varphi$; by the universal property of B there exists $\widetilde{\iota}$ such that $\widetilde{\iota}\varphi = \iota$. Note that $\widetilde{\iota}\widetilde{\varphi}\iota = \iota$ and $\widetilde{\varphi}\widetilde{\iota}\varphi = \varphi$. The uniqueness gives that $\widetilde{\iota}\widetilde{\varphi} = i_A$ and $\widetilde{\varphi}\widetilde{\iota} = i_B$. Moreover, in the proof of Theorem 3.1 we observed that $\widetilde{\varphi}$ is not only a ring homomorphism but also a R-homomorphism, whence

$$\widetilde{\varphi}(Mon(A)) = \{y_1^{\alpha_1} \cdots y_n^{\alpha_n} | \alpha_i \geqslant 0, 1 \leqslant i \leqslant n\}$$

is a R-basis of B. □

Corollary 3.3. Let R be a ring and $A = \sigma(R)\langle x_1, \ldots, x_n\rangle$ be a skew PBW extension of R with parameters $\sigma_i, \delta_i, c_{ij}, d_{ij}, a_{ij}^{(k)}$, $1 \leqslant i, j \leqslant n$. Let B be a ring that satisfies the following conditions with respect to the system of parameters $\sigma_i, \delta_i, c_{ij}, d_{ij}, a_{ij}^{(k)}$, $1 \leqslant i, j \leqslant n$, of the ring R.

(i) There exists a ring homomorphism $\varphi : R \to B$.

(ii) There exist elements $y_1, \ldots, y_n \in B$ such that B is a left free B-module with basis $Mon(y_1, \ldots, y_n)$, and the product is given by $r \cdot b := \varphi(r)b$, $r \in R, b \in B$.

(iii) The conditions (i) and (ii) in Theorem 3.1 hold.

Then $B \cong A = \sigma(R)\langle x_1, \ldots, x_n\rangle$.

Proof. According to the universal property of A, there exists a ring homomorphism $\widetilde{\varphi} : A \to B$ given by $r_1 x^{\alpha_1} + \cdots + a_t x^{\alpha_t} \mapsto \varphi(r_1)y^{\alpha_1} + \cdots + \varphi(a_t)y^{\alpha_t}$; from (ii) we get that $\widetilde{\varphi}$ is bijective. □

4. The Poincaré-Birkhoff-Witt theorem

Using the results of the previous sections, we will give now a new short proof of the Poincaré-Birkhoff-Witt theorem about the bases of enveloping algebras of finite-dimensional Lie algebras. Recall that if K is a field and $\mathcal{G}$

is a Lie algebra with K-basis $Y := \{y_1, \dots, y_n\}$, the enveloping algebra of $\mathcal{G}$ is the associative K-algebra $\mathcal{U}(\mathcal{G})$ defined by $\mathcal{U}(\mathcal{G}) = K\{y_1, \dots, y_n\}/I$, where $K\{y_1, \dots, y_n\}$ is the free K-algebra in the alphabet Y and I the two-sided ideal generated by all elements of the form $y_j y_i - y_i y_j - [y_j, y_i]$, $1 \leqslant i, j \leqslant n$, where $[\,,\,]$ is the Lie bracket of $\mathcal{G}$ (see [9]).

Theorem 4.1 (Poincaré-Birkhoff-Witt theorem)**.** The standard monomials $y_1^{\alpha_1} \cdots y_n^{\alpha_n}$, $\alpha_i \geqslant 0$, $1 \leqslant i \leqslant n$, conform a K-basis of $\mathcal{U}(\mathcal{G})$.

Proof. For the ring K we consider the following system of variables and parameters:

$$\begin{gathered} X := \{x_1, \dots, x_n\}, \quad \sigma_i := i_K, \quad \delta_i := 0, \quad c_{i,j} := 1, \quad d_{ij} := 0, \\ [x_i, x_j] = a_{ij}^{(1)} x_1 + \cdots + a_{ij}^{(n)} x_n, \quad 1 \leqslant i, j \leqslant n. \end{gathered} \tag{4.1}$$

We want to prove that conditions (1)–(3) in Theorem 2.4 hold. Condition (1) trivially holds. For (2) we have

$$\begin{aligned} h(x_j x_i r) &= h(x_j r x_i) = h(r x_j x_i) = r x_i x_j + r[x_j, x_i]; \\ h(p(x_j x_i) r) &= h(x_i x_j r) + h([x_j, x_i] r) = h(x_i r x_j) + r[x_j, x_i] \\ &= r x_i x_j + r[x_j, x_i]. \end{aligned}$$

Condition (3) of Theorem 2.4 also holds: In fact,

$$\begin{aligned} h(p(x_k x_j) x_i) &= h(x_j x_k x_i) + h([x_k, x_j] x_i) \\ &= h(x_j x_i x_k) + h(x_j [x_k, x_i]) + h([x_k, x_j] x_i) \\ &= x_i x_j x_k + h([x_j, x_i] x_k) + h(x_j [x_k, x_i]) + h([x_k, x_j] x_i) \\ &= x_i x_j x_k + (h(x_k [x_j, x_i]) + h([[x_j, x_i], x_k])) + (h([x_k, x_i] x_j) \\ &\quad + h([x_j, [x_k, x_i]])) + (h(x_i [x_k, x_j]) + h([[x_k, x_j], x_i])) \\ &= h(x_k x_j x_i) + h([[x_j, x_i], x_k] + [x_j, [x_k, x_i]] + [[x_k, x_j], x_i]) \\ &= h(x_k x_j x_i). \end{aligned}$$

The last equality holds by the Jacobi identity, the second to the last equality follows regrouping the terms and applying the definition of h to $h(x_k x_j x_i)$.

From Theorem 2.4 we conclude that there exists a skew PBW extension $A = \sigma(K)\langle x_1, \dots, x_n\rangle$ that satisfies (4.1), in particular, the monomials $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, $\alpha_i \geqslant 0$, $1 \leqslant i \leqslant n$, conform a K-basis of A. But note that $\mathcal{U}(\mathcal{G})$ satisfies the hypothesis in Corollary 3.2, so $\mathcal{U}(\mathcal{G}) \cong A$ and $\mathcal{U}(\mathcal{G})$ has K-basis $y_1^{\alpha_1} \cdots y_n^{\alpha_n}$, $\alpha_i \geqslant 0$, $1 \leqslant i \leqslant n$.

□

References

[1] Acosta, J.P., Chaparro, C., Lezama, O., Ojeda, I., and Venegas, C., *Ore and Goldie theorems for skew PBW extensions*, Asian-European Journal of Mathematics, 6 (4), 2013, 1350061-1; 1350061-20.

[2] Acosta, J. P., *Ideales Primos en Extensiones PBW torcidas*, Tesis de Maestría, Universidad Nacional de Colombia, Bogotá, 2014.

[3] Chaparro, C., *Valuations of skew quantum polynomials*, Asian-European Journal of Mathematics, 8(2), 2015.

[4] Dixmier, J., *Enveloping Algebras*, GSM 11, AMS, 1996.

[5] Bueso, J., Gómez-Torrecillas, J. and Verschoren, A., *Algorithmic Methods in noncommutative Algebra: Applications to Quantum Groups*, Kluwer, 2003.

[6] Humphreys, J. E., *Introduction to Lie Algebras and Representation Theory*, GTM 9, Springer, 1980.

[7] Lezama, O. and Gallego, C., *Gröbner bases for ideals of sigma-PBW extensions*, Communications in Algebra, 39 (1), 2011, 50-75.

[8] Lezama, O. & Reyes, M., *Some homological properties of skew PBW extensions*, Comm. in Algebra, 42, (2014), 1200-1230.

[9] McConnell, J. and Robson, J., *Non-commutative Noetherian Rings*, Graduate Studies in Mathematics, AMS, 2001.

[10] Venegas, C., *Automorphisms for skew PBW extensions and skew quantum polynomial rings*, Communications in Algebra, 42(5), 2015, 1877-1897.

CONTACT INFORMATION

Juan Pablo Acosta,
Oswaldo Lezama
Departamento de Matemáticas
Universidad Nacional de Colombia, Sede Bogotá
E-Mail(*s*): jolezamas@unal.edu.co

Received by the editors: 02.03.2015
and in final form 16.03.2015.

Algebra and Discrete Mathematics
Volume **20** (2015). Number 1, pp. 13–31

RESEARCH ARTICLE

Lie and Jordan structures of differentially semiprime rings

Orest D. Artemovych and Maria P. Lukashenko

Communicated by A. P. Petravchuk

Abstract. Properties of Lie and Jordan rings (denoted respectively by R^L and R^J) associated with an associative ring R are discussed. Results on connections between the differentially simplicity (respectively primeness, semiprimeness) of R, R^L and R^J are obtained.

1. Introduction

Throughout here, R is an associative ring (with respect to the addition "$+$" and the multiplication "$\cdot$") with an identity, $\operatorname{Der} R$ is the set of all derivations in R. On the set R we consider two operations: the Lie multiplication "$[-,-]$" and the Jordan multiplication "$(-,-)$" defined by the rules

$$[a,b] = a \cdot b - b \cdot a$$

and

$$(a,b) = a \cdot b + b \cdot a$$

for any $a, b \in R$. Then

$$R^L = (R, +, [-,-])$$

is a Lie ring and

$$R^J = (R, +, (-,-))$$

2010 MSC: Primary 16W25, 16N60; Secondary 17B60, 17C50.
Key words and phrases: Derivation, semiprime ring, Lie ring.

is a Jordan ring (see [13] and [14]) associated with the associative ring R. Recall that an additive subgroup A of R is called:

- *a Lie ideal* of R if
$$[a, r] \in A,$$
- *a Jordan ideal* of R if
$$(a, r) \in A$$

for all $a \in A$ and $r \in R$. Obviously, A is a Lie (respectively Jordan) ideal of R if and only if A^L (respectively A^J) is an ideal of R^L (respectively R^J).

In all that follows Δ will be any subset of $\operatorname{Der} R$ (in particular, $\Delta = \{0\}$) and $\delta \in \operatorname{Der} R$. A subset K of R is called Δ-*stable* if $d(a) \in K$ for all $d \in \Delta$ and $a \in K$. An ideal I of a (Lie, Jordan or associative) ring A is said to be a Δ-*ideal* if I is Δ-stable. A (Lie, Jordan or associative) ring A is said to be:

- *simple* (respectively Δ-*simple*) if there no two-sided ideals (respectively Δ-ideals) other 0 or A,
- *prime* (respectively Δ-*prime*) if, for all two-sided ideals (respectively Δ-ideals) K, S of A, the condition $KS = 0$ implies that $K = 0$ or $S = 0$ (if $\Delta = \{\delta\}$ and A is Δ-prime, then we say that A is δ-*prime*),
- *semiprime* (respectively Δ-*semiprime*) if, for any two-sided ideal (respectively Δ-ideal) K of A, the condition $K^2 = 0$ implies that $K = 0$,
- *primary* if, for any two-sided ideals K, S of A, the condition $KS = 0$ implies that $K = 0$ or S is nilpotent.

Every non-commutative Δ-simple ring is Δ-prime and every Δ-prime ring is Δ-semiprime. We say that R is $\mathbb{Z}$-*torsion-free* if, for any $r \in R$ and integers n, the condition $nr = 0$ holds if and only if $r = 0$. If the implication
$$2r = 0 \Rightarrow r = 0$$
is true for any $r \in R$, then R is said to be 2-*torsion-free*. Let
$$F_p(R) = \{a \in R \mid a \text{ has an additive order } p^k \text{ for some non-negative} k = k(a)\}$$
be the p-part of R, where p is a prime. Then $F_p(R)$ is a Δ-ideal of R. If R is Δ-semiprime, then
$$pF_p(R) = 0.$$

In particular, in a Δ-prime ring R it holds $F_p(R) = 0$ (and so the characteristic $\operatorname{char} R = 0$) or $F_p(R) = R$ (and therefore $\operatorname{char} R = p$). Obviously that the additive group R^+ of a Δ-prime ring R is torsion-free if and only if $\operatorname{char} R = 0$. Recall that a ring R is said to be *of bounded index* m, if m is the least positive integer such that $x^m = 0$ for all nilpotent elements $x \in R$. We say that a ring R satisfies *the condition* (X) if one of the following holds:

(1) R or $R/\mathbb{P}(R)$ is $\mathbb{Z}$-torsion-free, where $\mathbb{P}(R)$ is the prime radical of R,

(2) R is of bounded index m such that an additive order of every nonzero torsion element of R, if any, is strictly larger than m.

As noted in [16, p.283], a $\mathbb{Z}$-torsion-free δ-prime ring is semiprime. In this way we prove the following

Proposition 1. *For a ring R the following hold:*

(1) *if R is a Δ-semiprime ring with the condition (X), then it is semiprime,*

(2) *if R is both semiprime (respectively satisfies the condition (X)) and Δ-prime, then R is prime.*

Relations between properties of an associative ring R, a Lie ring R^L and a Jordan ring R^J was studied by I.N. Herstein and his students (see [7, 8, 11] and bibliography in [9] and [5]); he has obtained, for a ring R of characteristic different from 2, that the simplicity of R implies the simplicity of a Jordan ring R^J [7, Theorem 1], and also that every Lie ideal of a simple Lie ring R is contained in the center $Z(R)$ [7, Theorem 3]. K. McCrimmon [20, Theorem 4] has proved that R is a simple algebra if and only if R^J is a simple Jordan algebra. Our result is the following

Theorem 1. *For a 2-torsion-free ring R the following statements are true:*

(1) *R is a Δ-simple ring if and only if R^J is a Δ-simple Jordan ring,*

(2) *R is a Δ-prime ring if and only if R^J is a Δ-prime Jordan ring,*

(3) *R is a Δ-semiprime ring if and only if R^J is a Δ-semiprime Jordan ring.*

Let us $d \in \Delta$. Since $C(R)$ and $\operatorname{ann} C(R)$ are Δ-ideals, the rule

$$\overline{d} : R/\operatorname{ann} C(R) \ni r + \operatorname{ann} C(R) \mapsto d(r) + \operatorname{ann} C(R) \in R/\operatorname{ann} C(R)$$

determines a derivation $\overline{d}$ of the quotient ring $R/\operatorname{ann} C(R)$. Then

$$\overline{\Delta} = \{\overline{d} \mid d \in \Delta\} \subseteq \operatorname{Der}(R/\operatorname{ann} C(R)).$$

Inasmuch $d(Z(R)) \subseteq Z(R)$, the rule

$$\widehat{d} : R^L/Z(R) \ni r + Z(R) \mapsto d(r) + Z(R) \in R^L/Z(R)$$

determines a derivation $\widehat{d}$ of the Lie ring $R^L/Z(R)$. Then

$$\widehat{\Delta} = \{\widehat{d} \mid d \in \Delta\} \subseteq \operatorname{Der}(R^L/Z(R)).$$

Since the center $Z(R)$ is a nonzero Lie ideal of an associative ring R with an identity, a Lie ring R^L is not Δ-simple. Our next result is the following

Theorem 2. *Let R be a 2-torsion-free ring. Then the following are true:*

(1) *if the quotient ring $R^L/Z(R)$ is a $\widehat{\Delta}$-simple Lie ring, then R is non-commutative and $R/\operatorname{ann} C(R)$ is a $\overline{\Delta}$-simple ring,*

(2) *if R is a Δ-simple ring, then $R^L/Z(R)$ is a $\widehat{\Delta}$-simple Lie ring or R is commutative,*

(3) *if $R^L/Z(R)$ is a $\widehat{\Delta}$-semiprime Lie ring, then R is non-commutative and the quotient ring $R/\operatorname{ann} C(R)$ is a $\overline{\Delta}$-semiprime ring,*

(4) *if R is a Δ-semiprime ring, then $R^L/Z(R)$ is a $\widehat{\Delta}$-semiprime Lie ring or R is commutative,*

(5) *if $R^L/Z(R)$ is a $\widehat{\Delta}$-prime Lie ring, then R is non-commutative and $R/\operatorname{ann} C(R)$ is a $\overline{\Delta}$-prime ring,*

(6) *if R is a Δ-prime ring, then $R^L/Z(R)$ is a $\widehat{\Delta}$-prime Lie ring or R is commutative.*

Throughout, let $Z(R)$ denote the center of R, $[A, B]$ (respectively (A, B)) an additive subgroup of R generated by all commutators $[a, b]$ (respectively (a, b)), where $a \in A$ and $b \in B$, $C(R)$ the commutator ideal of R, $N(R)$ the set of nilpotent elements in R, $\operatorname{char} R$ the characteristic of R, $\operatorname{ann}_l I = \{a \in R \mid aI = 0\}$ the left annihilator of I in R, $\operatorname{ann}_r I = \{a \in R \mid Ia = 0\}$ the right annihilator of I in R, $\operatorname{ann} I = (\operatorname{ann}_r I) \cap (\operatorname{ann}_l I)$, $C_R(I) = \{a \in R \mid ai = ia \text{ for all } i \in I\}$ the centralizer of I in R and $\partial_a(x) = [a, x]$ for $a, x \in R$.

All other definitions and facts are standard and it can be found in [10], [17] and [19].

2. Differentially prime right Goldie rings

Let agree that

$$d^0 = \mathrm{id}_R$$

is the identity endomorphism for $d \in \Delta$.

Lemma 1. *The following conditions are equivalent:*

(1) *R is a Δ-semiprime ring,*

(2) *for any Δ-ideals A, B of R the implication*

$$AB = 0 \Rightarrow A \cap B = 0$$

is true,

(3) *if $a \in R$ is such that*

$$aR\delta_1^{m_1} \dots \delta_k^{m_k}(a) = 0$$

for any integers $k \geqslant 1$, $m_i \geqslant 0$ and derivations $\delta_i \in \Delta$ $(i = 1, \dots, k)$, then $a = 0$.

Proof. A simple modification of Proposition 2 from [17, §3.2]. □

Lemma 2. *The following conditions are equivalent:*

(1) *R is a Δ-prime ring,*

(2) *a left annihilator $\mathrm{ann}_l\, I$ of a left Δ-ideal I of R is zero,*

(3) *a right annihilator $\mathrm{ann}_r\, I$ of a right Δ-ideal I of R is zero,*

(4) *if $a, b \in R$ are such that*

$$aR\delta_1^{m_1} \dots \delta_k^{m_k}(b) = 0$$

for any integers $k \geqslant 1$, $m_j \geqslant 0$ and derivations $\delta_j \in \Delta$ $(j = 1, \dots, k)$, then $a = 0$ or $b = 0$.

Proof. A simple consequence of Lemma 2.1.1 from [10]. □

If I is an ideal of a ring R, then

$$\mathcal{C}_R(I) = \{x \in R \mid x + I \text{ is regular in the quotient ring } R/I\}$$

(see [19, Chapter 2, §1]). The next lemma extends Proposition 1 of [15].

Lemma 3. *Let R be a right Goldie ring and $\delta \in \operatorname{Der} R$. If R is δ-prime, then:*

(*a*) *the set $N = N(R)$ of nilpotent elements of R is its prime radical,*

(*b*) $\bigcap_{i=1}^{k} \delta^{-1}(N) = 0$ *for some integer k,*

(*c*) $\mathcal{C}_R(0) = \mathcal{C}_R(N)$.

Proof. From Theorem 2.2 of [16] (see the part $(ii) \Rightarrow (iii)$ of its proof), we obtain (a) and (b). By Proposition 4.1.3 of [19], $\mathcal{C}_R(0) \subseteq \mathcal{C}_R(N)$. By the same argument as in [16, p.284], we can obtain that $\mathcal{C}_R(0) = \mathcal{C}_R(N)$. □

Corollary 1. *If R is a commutative δ-prime Goldie ring and $\delta \in \operatorname{Der} R$, then $N(R)$ contains all zero-divisors of R.*

By Corollary 1.4 of [6], if I is a δ-prime ideal of a right Noetherian ring R and R/I has characteristic 0, then I is prime. The following lemma is an extension of Lemma 2.5 from [6].

Lemma 4. *Let R be a 2-torsion-free commutative Goldie ring and $\delta \in \operatorname{Der} R$. If R is δ-prime, then it is an integral domain.*

Proof. Assume that $a \in \operatorname{ann} N(R)$, $b \in N(R)$ and $r \in R$. Then

$$\begin{aligned} 0 = \delta^2(arb) &= \delta(\delta(a)rb + a\delta(r)b + ar\delta(b)) \\ &= \delta^2(a)rb + 2\delta(a)\delta(r)b + 2\delta(a)r\delta(b) + a\delta^2(r)b + 2a\delta(r)\delta(b) + ar\delta^2(b) \end{aligned}$$

and so

$$2\delta(a)R\delta(b) \subseteq N(R).$$

This means that $\delta(a) \in N(R)$ or $\delta(b) \in N(R)$. Hence $N(R)$ is δ-stable. By Lemma 3, $N(R)$ is a ideal and therefore $N(R) = 0$. By Lemma 1.2 of [4], R is prime and consequently it is an integral domain. □

Proof of Proposition 1.

(1) By Proposition 1.3 of [6] and Theorem 1 of [1], the prime radical $\mathbb{P}(R)$ is a Δ-ideal and so $\mathbb{P}(R) = 0$ is zero.

(2) Since $\mathbb{P}(R) = 0$, R is prime by Lemma 1.2 from [4]. □

By Theorem 4 of [22], a Δ-simple ring R of characteristic 0 is prime. Since every non-commutative Δ-simple ring is Δ-prime, in view of Proposition 1 we obtain the following

Corollary 2. *Let R be a semiprime ring (respectively a ring R satisfy the condition (X)). If R is Δ-simple, then it is prime.*

3. Differential analogues of Herstein's results

For the proof of Theorem 2 we need the next results. In the proofs below we use the same consideration, as in [12, Chapter 1, §1], and present them here in order to have the paper more self-contained. Let agree that everywhere in this section $k \geqslant 1$ and $m_i \geqslant 0$ are integers $(i = 1, \dots, k)$.

Lemma 5. *Let R be a Δ-semiprime ring, A and B its Δ-ideals. Then the following statements hold:*

(*i*) *if $AB = 0$, then $BA = 0$.*

(*ii*) $\operatorname{ann}_l A = \operatorname{ann}_r A$.

(*iii*) $A \cap \operatorname{ann}_r A = 0$.

Proof. (*i*) Indeed, BA is a Δ-ideal and $(BA)^2 = 0$ and so $BA = 0$.

(*ii*) We denote $(\operatorname{ann}_r A)A$ by X. Since X is a Δ-ideal and $X^2 = 0$, we deduce that $X = 0$. This means that

$$\operatorname{ann}_r A \subseteq \operatorname{ann}_l A.$$

The inverse inclusion we can prove similarly.

(*iii*) Since $A \cap \operatorname{ann}_r A$ is a nilpotent Δ-ideal, the assertion holds. □

Henceforth

$$X_a = \{[\delta_1^{m_1} \dots \delta_k^{m_k}(a), x] \mid x \in R,\ \delta_i \in \Delta,\ m_i \geqslant 0 \text{ and } k \geqslant 1 \text{ are integers } (i = 1, \dots, k)\}.$$

It is clear that $[a, x] \in X_a$.

Lemma 6. *Let R be a Δ-semiprime ring and $a \in R$. Then the following statements hold:*

(*i*) *if*

$$a[\delta_1^{m_1} \dots \delta_k^{m_k}(a), R] = 0$$

for any integers $k \geqslant 1$, $m_i \geqslant 0$ and derivations $\delta_i \in \Delta$ $(i = 1, \dots, k)$, then $a \in Z(R)$,

(*ii*) *if I is a right Δ-ideal of R, then $Z(I) \subseteq Z(R)$,*

(*iii*) *if I is a commutative right Δ-ideal of R and I is nonzero, then $I \subseteq Z(R)$. If, moreover, R is Δ-prime, then it is commutative.*

Proof. (i) Let $x, y \in R$ and $d, \delta \in \Delta$. Since

$$[b, xy] = [b, x]y + x[b, y] \tag{3.1}$$

for any $b \in X_a$ and $a[b, xy] = 0$, we conclude that $ax[b, y] = 0$. This gives that $ayx[b, y] = 0$ and $yax[b, y] = 0$ and consequently

$$(R[a, y]R)^2 = 0. \tag{3.2}$$

In addition,

$$0 = d(a[b, x]) = d(a)[b, x].$$

Multiplying (3.1) by $d(a)$ on left we get $d(a)x[b, y] = 0$. Moreover,

$$0 = \delta(ax[d(b), y]) = \delta(a)x[d(b), y]$$

and, by the similar argument, we obtain that

$$\delta_1^{m_1} \dots \delta_k^{m_k}(a)x[\delta_1^{m_1} \dots \delta_k^{m_k}(a), y] = 0$$

for any integers $k \geqslant 1$, $m_i \geqslant 0$ and derivations $\delta_i \in \Delta$ $(i = 1, \dots, k)$. As in the proof of the condition (3.2), we deduce that

$$(R[\delta_1^{m_1} \dots \delta_k^{m_k}(a), y]R)^2 = 0.$$

Then

$$I = \sum_{k=1}^{\infty} \sum_{\substack{\delta_1 \dots \delta_k \in \Delta \\ y \in R}} R[\delta_1^{m_1} \dots \delta_k^{m_k}(a), y]R$$

is a sum of nilpotent ideals and therefore it is a nil ideal. Since I is a Δ-ideal, we conclude that $I = 0$ and, as a consequence, $a \in Z(R)$.

(ii) Let $a \in Z(I)$ and $y \in R$. Then, for $\delta_1, \dots, \delta_k \in \Delta$, we have

$$\delta_1^{m_1} \dots \delta_k^{m_k}(a) \in Z(I)$$

and $ay \in I$. This gives that

$$a(\delta_1^{m_1} \dots \delta_k^{m_k}(a)y) = \delta_1^{m_1} \dots \delta_k^{m_k}(a)(ay) = a(y\delta_1^{m_1} \dots \delta_k^{m_k}(a)),$$

and thus

$$a[\delta_1^{m_1} \dots \delta_k^{m_k}(a), y] = 0.$$

By (i), $a \in Z(R)$ is central.

(*iii*) By (*ii*), $I \subseteq Z(R)$. Assume that R is Δ-prime, $u, v \in R$ and $a \in I$. Then $au \in I$ and so $au \in Z(R)$. Since

$$a(uv) = (au)v = v(au) = (va)u = a(vu),$$

we see that

$$[u, v] \in \operatorname{ann}_r I.$$

By Lemma 2(3), $[u, v] = 0$ and hence R is commutative. □

Lemma 7. *Let R be a Δ-prime ring and $a \in R$. If $a \in C_R(I)$ for some nonzero right Δ-ideal I of R, then $a \in Z(R)$.*

Proof. Let us $y \in R$ and $b \in I$. Then $by \in I$ and so $bay = a(by) = bya$. This yields that

$$I[a, y] = 0 = [a, y]I.$$

By Lemma 2(3), $[a, y] = 0$. Hence $a \in Z(R)$. □

Lemma 8. *The left annihilator $\operatorname{ann}_l(X_a)$ is a left Δ-ideal of R.*

Proof. Immediate from the definition. □

Lemma 9. *If R is a Δ-semiprime ring, then $C_R([R, R]) \subseteq Z(R)$.*

Proof. Let us $a \in C_R([R, R]), d, \delta \in \Delta$ and $x, y \in R$. Putting x for a and $xd(a)$ for xy in (3.1) we obtain

$$[x, xd(a)] = [x, x]d(a) + x[x, d(a)]$$

and, as a consequence, $[a, x[x, d(a)]] = 0$ and $[a, x][x, d(a)] = 0$. Then, by the same reasons as in the proof of Lemma 6(i), we obtain that $[a, x] \in \operatorname{ann}_l(X_a)$ and $A = \operatorname{ann}_l(X_a)$ is a Δ-ideal. Then

$$[\delta(a), x][d(a), x] = \delta([a, x][d(a), x]) = 0.$$

Since $A \cap \operatorname{ann}_l A = 0$, we deduce that is a nilpotent Δ-ideal and so $a \in Z(R)$. □

Lemma 10. *Let R be a 2-torsion-free Δ-semiprime ring. If $a \in R$ commutes with all elements of X_a, then $a \in Z(R)$.*

Proof. Let $r, x, y \in R$ and $d \in \Delta$. It is clear that $\partial_a^2(x) = 0$. From $\partial_a^2(xy) = 0$ it follows that

$$2\partial_a(x)\partial_a(y) = 0$$

and so $\partial_a(x)\partial_a(y) = 0$. Since

$$0 = \partial_a(x)\partial_a(rx) = \partial_a(x)\partial_a(r)x + \partial_a(x)r\partial_a(x) = \partial_a(x)r\partial_a(x),$$

we deduce that $\partial_a(x)R\partial_a(x) = 0$ and $(\partial_a(x)R)^2 = 0$. Moreover, $a[b, x] = [b, x]a$ for any $[b, x] \in X_a$ and therefore

$$d(a)[b, x] + a[d(b), x] + a[b, d(x)] = [b, x]d(a) + [d(b), x]a + [b, d(x)]a.$$

From this it holds that

$$d(a)[b, x] = [b, x]d(a).$$

This means that $C_R(X_a)$ is Δ-stable and $(\partial_{d(a)}(x)R)^2 = 0$. As a consequence,

$$I = \sum_{k=1}^{\infty} \sum_{\substack{x \in R \\ m_k \geqslant 0 \\ \delta_1, \ldots, \delta_k \in \Delta}} \partial_{\delta_1^{m_1} \ldots \delta_k^{m_k}(a)}(x)R$$

is a sum of nilpotent ideals and so I is a nil ideal. Since I is a Δ-ideal, we deduce that $I = 0$. Hence $a \in Z(R)$. □

The next lemma is an extension of Lemma 1 from [11] in the differential case.

Lemma 11. *Let R be a 2-torsion-free Δ-semiprime ring, T its Lie Δ-ideal. If $[T, T] \subseteq Z(R)$, then $T \subseteq Z(R)$.*

Proof. Let $x \in R$ and $t \in T$.

1) If $[T, T] = 0$, then $[t, x] \in T$ and so $[t, [t, x]] = 0$. By Lemma 10, $T \subseteq Z(R)$.

2) Now assume that $0 \neq [a, b] \in [T, T]$ for some $a, b \in T$. Then

$$\partial_a(b) \in Z(R) \text{ and } \partial_a^2(R) \subseteq Z(R).$$

Moreover, we have that

$$\begin{aligned} Z(R) \ni \partial_a^2(bx) &= \partial_a(\partial_a(b)x + b\partial_a(x)) \\ &= \partial_a^2(b)x + 2\partial_a(b)\partial_a(x) + b\partial_a^2(x) \\ &= 2\partial_a(b)\partial_a(x) + b\partial_a^2(x) \end{aligned}$$

and hence

$$[2\partial_a(b)\partial_a(x) + b\partial_a^2(x), b] = 0.$$

Then

$$\begin{aligned} 0 &= 2\partial_b(\partial_a(b))\partial_a(x) + 2\partial_a(b)\partial_b(\partial_a(x)) + \partial_b(b)\partial_a^2(x) + b\partial_b(\partial_a^2(x)) \\ &= 2\partial_a(b)\partial_b(\partial_a(x)) \end{aligned} \tag{3.3}$$

and

$$\partial_a(ba) = \partial_a(b)a + b\partial_a(a) = \partial_a(b)a.$$

Replacing ba for x in (3.3) we have

$$0 = 2\partial_a(b)\partial_b(\partial_a(b)a) = 2\partial_a(b)(\partial_b(\partial_a(b)) + \partial_a(b)\partial_b(a)) = -2\partial_a(b)^3$$

and thus $\partial_a(b)^3 = 0$. Then $R\partial_a(b)$ is a nilpotent ideal in R and, as a consequence,

$$\sum_{a,b\in T} R\partial_a(b)$$

is a nonzero nil Δ-ideal, a contradiction. □

Lemma 12. *If U is a Lie Δ-ideal of a ring R and $I(U) = \{u \in R \mid uR \subseteq U\}$, then $I(U)$ is the largest Δ-ideal of R such that $I(U) \subseteq U$.*

Proof. Let $u, v \in I(U)$, $x, y \in R$ and $\delta \in \Delta$. Clearly that $I(U)$ is an additive subgroup of $R, I(U) \subseteq U$ and $(ux)y = u(xy) \in (ux)R = u(xR) \subseteq uR \subseteq U$ that is $ux \in I(U)$. From

$$u(xy) - (yu)x = (ux)y - y(ux) = [ux, y] \in U$$

(and so $(yu)x \in U$) it holds that $yu \in I(U)$. Hence U is a two-sided ideal of R. Moreover,

$$\delta(u)x + u\delta(x) = \delta(ux) \in \delta(U) \subseteq U$$

and $u\delta(x) \in uR \subseteq U$. Therefore $\delta(u)x \in U$. This means that $I(U)$ is a Δ-ideal of R. If A is a Δ-ideal of R that is contained in U, then $AR \subseteq A \subseteq U$ and hence $A \subseteq I(U)$. □

Lemma 13. *Let U be a Lie Δ-ideal of R. If U is an associative subring of R, then $[U, U] = 0$ or U contains a nonzero Δ-ideal of R.*

Proof. Assume that $x \in R$ and $[U,U] \neq 0$. Then $[u,v] \neq 0$ for some $u,v \in U$ and

$$[u,vx] = u(vx) - (vx)u = (uv - vu)x + v(ux - xu).$$

Since $[u,x],[u,vx] \in U$ and $v[u,x] \in U$, we deduce that $[u,v]x \in U$. This means that $[u,v] \in I(U)$. In view of Lemma 12, $I(U)$ is a nonzero Δ-ideal of R that is contained in U. □

Proposition 2. *If U is a Lie Δ-ideal of R, then $[U,U] = 0$ or there exists a nonzero Δ-ideal I_U of R such that $[I_U,R] \subseteq U$.*

Proof. By Lemma 3 of [7],

$$T(U) = \{t \in R \mid [t,R] \subseteq U\}$$

is both a Lie ideal and an associative subring of R and $U \subseteq T(U)$. Moreover, for $\delta \in \Delta$, we have

$$[\delta(t),R] + [t,\delta(R)] = \delta([t,R]) \subseteq \delta(U) \subseteq U$$

and so $[\delta(t),R] \subseteq U$. Hence $T(U)$ is Δ-stable. If $[U,U] \neq 0$, then, by Lemmas 12 and 13,

$$I_U = I(T(U)) \subseteq T(U)$$

is a nonzero Δ-ideal of R such that $[I_U,R] \subseteq U$. □

Lemma 14. *Let U be a Lie Δ-ideal of a ring R. If $[U,U] = 0$, then the centralizer $C_R(U)$ is a Lie Δ-ideal and an associative subring of R.*

Proof. Is immediately. □

We extend Theorem 1.3 of [9] in the following

Proposition 3. *Let R be a Δ-simple ring of characteristic 2. If U is a Lie Δ-ideal of R, then one of the following holds:*

(1) $[R,R] \subseteq U$,

(2) $U \subseteq Z(R)$,

(3) *R contains a subfield P such that $U \subseteq P$ and $[P,R] \subseteq P$.*

Proof. If $[U,U] \neq 0$, then $[R,R] \subseteq U$ by Proposition 2. Therefore we assume that $[U,U] = 0$. By Lemma 14, $C_R(U)$ is a Lie Δ-ideal and an associative subring of R such that $U \subseteq C_R(U)$.

a) If $C_R(U)$ is non-commutative, then $C_R(U) = R$ by Lemma 13. Hence $U \subseteq Z(R)$.

b) Now assume that the centralizer $C_R(U)$ is commutative. If $c \in C_R(U)$ and $x \in R$, then

$$c^2 \in C_R(U) \text{ and } [c^2, x] = [[c,x],x] = 2c[c,x] = 0.$$

This gives that $c^2 \in Z(R)$. By Theorem 2 of [22], $Z(R)$ is a field. As a consequence, c^2 (and so c) is invertible in $C_R(U)$. Hence $C_R(U)$ is a field. □

Corollary 3. *Let R be a Δ-simple ring. If U is a Lie Δ-ideal of R, then one of the following holds:*

(1) $[R,R] \subseteq U$,

(2) $U \subseteq Z(R)$,

(3) $\operatorname{char} R = 2$ *and R contains a subfield P such that $U \subseteq P$ and $[P,R] \subseteq P$.*

4. Jordan properties

Lemma 15. *Let R be a Δ-simple ring of characteristic $\neq 2$, U its proper Jordan Δ-ideal and $a \in U$. If $[a,R] \subseteq U$, then $a = 0$.*

Proof. Let us $x, y \in R$. Since $[a,x] \in U$ and $(a,x) \in U$, we obtain that $2ax \in U$ and, as a consequence, $ax \in U$ and $(ax,y) \in U$. Moreover, from $axy \in U$ it follows that $yax \in U$. This means that $RaR \subseteq U$. Since $d(a) \in U$ for any $d \in \Delta$, in view of [21, Lemma 1.1] we obtain that

$$\sum_{k=1}^{\infty} \sum_{\substack{\delta_1,\dots,\delta_k \in \Delta \\ (m_1,\dots,m_k) \in \mathbb{N}^k}} R\delta_1^{m_1} \dots \delta_k^{m_k}(a)R$$

is a proper Δ-ideal of R that is contained in U. Hence $a = 0$. □

Remark 1. Let R be a 2-torsion-free ring, U its Jordan Δ-ideal. If Δ contains all inner derivations of R, then U is an ideal of R.

In fact, we have

$$2xa = [a, x] + (a, x) \in U$$

for any $a, b, x \in U$ and so $xa \in U$. By the same argument, we can conclude that $ax \in U$.

Proof of Theorem 1.

(1) ($\Leftarrow$) If A is a nonzero proper Δ-ideal of a ring R, then A^J is a nonzero proper Δ-ideal of R^J, a contradiction.

($\Rightarrow$) Let U be a proper Jordan Δ-ideal of R, $a, b \in U$ and $x \in R$. By Lemma 1 of [7], $[(a, b), x] \in U$, and, by Lemma 15, we see that

$$(a, b) = 0. \tag{4.4}$$

In particular, $2a^2 = 0$ and, as a consequence, $a^2 = 0$ and $2axa = (a, (a, x)) = 0$. It follows that $axa = 0$. Since

$$0 = (a + b)x(a + b) = axb + bxa$$

and

$$0 = (b, (a, x)) = b(ax + xa) + (ax + xa)b = bax + bxa + axb + xba,$$

we deduce that $bax + xab = 0$. But $ab = -ba$ and so $bax - xba = 0$. This means that $ba \in Z(R)$. Then $(RabR)^2 = 0$. Since

$$I = \sum_{k=1}^{\infty} \sum_{\substack{a,b \in U,\ \delta_1,\ldots,\delta_k \in \Delta \\ (m_1,\ldots,m_k) \in \mathbb{N}^k}} Ra\delta_1^{m_1} \ldots \delta_k^{m_k}(b)R$$

is a Δ-ideal of R that is a sum of nilpotent ideals, we obtain that $I = 0$. Therefore

$$0 = (b, x)a = (bx + xb)a = bxa + xba = 2bxa.$$

We conclude that $URU = 0$. From $(RUR)^2 = 0$ and $\delta(RUR) \subseteq RUR$ for any $\delta \in \Delta$ it holds that $U = 0$.

(2) ($\Leftarrow$) If A, B are Δ-ideals of R such that $AB = 0$, then $(BA)^2 = 0$ and so BA is a Jordan ideal of R satisfying the condition

$$(BA, BA) = 0.$$

Thus the condition (4.4) is true for $U = BA$. As in the proof of the part (1), we obtain that $BA = 0$. Then A^J, B^J are Δ-ideals of a Jordan ring R^J such that

$$(A^J, B^J) = 0.$$

Hence $A = 0$ or $B = 0$.

($\Rightarrow$) Let $a_1, a_2 \in A$ and $x, y \in R$. Suppose that R^J is not Δ-prime and therefore there exist nonzero Jordan Δ-ideals A, B of R such that

$$(A, B) = 0.$$

By the same reasons as above, we conclude that $A \cap B = 0$. Then, by Lemma 1 of [7], we have $[(a_1, a_2), x] \in A$ and hence

$$[(a_1, a_2), x] \pm ((a_1, a_2), x) \in A.$$

Therefore $x(a_1, a_2)y \in A$. Thus R contains Δ-ideals $R(A, A)R \subseteq A$ and $R(B, B)R \subseteq B$ such that

$$R(A, A)R(B, B)R \subseteq A \cap B = 0.$$

Hence $(A, A) = 0$ or $(B, B) = 0$ and this leads to a contradiction.

(3) ($\Leftarrow$) If A is a nonzero Δ-ideal of R such that $A^2 = 0$, then A^J is a nonzero Δ-ideal of the Jordan ring R^J such that

$$(A^J, A^J) = 0,$$

a contradiction.

($\Rightarrow$) Suppose that R has a nonzero Jordan Δ-ideal U such that

$$(U, U) = 0.$$

Then the condition (4.4) is true for any $a, b \in U$. As in the proof of the part (1), we obtain that $U = 0$.

□

If R is a ring, then on the set R we can to define a left Jordan multiplication "$\langle -, - \rangle$" by the rule

$$\langle a, b \rangle = 2ab$$

for any $a, b \in R$. Then the equalities

$$\langle \langle \langle a, a \rangle, b \rangle, a \rangle = \langle \langle a, a \rangle, \langle b, a \rangle \rangle \quad \text{and} \quad \langle \langle \langle a, b \rangle, a \rangle = \langle a, \langle b, a \rangle \rangle$$

are true and hence

$$R^{lJ} = (R, +, \langle -, - \rangle)$$

is a non-commutative Jordan ring (which is called *a left Jordan ring associated with an associative ring* R). It is clear that, for commutative ring R, we have

$$R^J = R^{lJ}.$$

If A is an additive subgroup of R that $\langle a, r\rangle, \langle r, a\rangle \in A$ for any $a \in A$ and $r \in R$, then A is called *an ideal* of R^{lJ}. If $\delta \in \Delta$ and $a, b \in R$, then

$$\delta(\langle a, b\rangle) = \delta(2ab) = 2\delta(a)b + 2a\delta(b) = \langle \delta(a), b\rangle + \langle a, \delta(b)\rangle$$

and therefore $\delta \in \mathrm{Der}(R^{lJ})$. By the other hand, if $\delta \in \mathrm{Der}(R^{lJ})$, then

$$2\delta(ab) = \delta(\langle a, b\rangle) = \langle \delta(a), b\rangle + \langle a, \delta(b)\rangle = 2(\delta(a)b + a\delta(b)).$$

If R is a 2-torsion-free ring, then $\delta \in \mathrm{Der}\, R$. Similarly, as in Theorem 1, we can prove the following

Proposition 4. *For a 2-torsion-free ring R the following conditions are true:*

(1) *R is a Δ-simple ring if and only if R^{lJ} is a Δ-simple Jordan ring,*

(2) *R is a Δ-prime ring if and only if R^{lJ} is a Δ-prime Jordan ring,*

(3) *R is a Δ-semiprime ring if and only if R^{lJ} is a Δ-semiprime Jordan ring.*

5. Proofs

The next lemma in the prime case is contained in [18, Lemma 7].

Lemma 16 ([2, Lemma 1.7])**.** *Let R be a ring. If $[[R, R], [R, R]] = 0$, then the commutator ideal $C(R)$ is nil.*

Corollary 4. *If R is a non-commutative Δ-semiprime ring, then $[R, R]$ is non-commutative.*

Proof of Theorem 2.

(1) It is clear that a ring R is non-commutative. If A is a nonzero proper Δ-ideal of R, then A^L is a nonzero proper Δ-ideal of R^L. Therefore $A \subseteq Z(R)$ and, as a consequence, $A \cdot C(R) = 0$.

(2) Suppose that a Δ-simple ring R is non-commutative and U is its nonzero proper Lie Δ-ideal. By Proposition 2, $[U, U] = 0$. Then, by Lemma 11, $U \subseteq Z(R)$. Hence the quotient ring $R^L/Z(R)$ is $\widehat{\Delta}$-simple.

(3) Let A be a nonzero Δ-ideal of R such that $A^2 = 0$. Then A^L is a nonzero Δ-ideal of a Lie ring R^L and, moreover,

$$[A^L, A^L] = 0.$$

By Lemma 11, $A \subseteq Z(R)$ and hence $A \cdot C(R) = 0$.

(4) Suppose that R is non-commutative. Let A be a nonzero Lie Δ-ideal of R such that $[A, A] = 0$. Then, by Lemma 11, $A \subseteq Z(R)$ and, as a consequence, the Lie ring $R^L/Z(R)$ is $\widehat{\Delta}$-semiprime.

(5) Let A, B be nonzero Δ-ideals of R such that $AB = 0$. Obviously, $[A, B] \subseteq Z(R)$. Then $A \subseteq Z(R)$ or $B \subseteq Z(R)$.

(6) Assume that R is non-commutative and A, B are nonzero Lie Δ-ideals of R such that

$$[A, B] = 0.$$

Then $A \cap B \subseteq Z(R)$. Since $A \cap B \subseteq \operatorname{ann} C(R)$ in a Δ-prime ring R, we have that the intersection $A \cap B = 0$ is zero. If $T(A) = R$ (see proof of Proposition 2), then $[R, R] \subseteq A$ and $B \subseteq C_R([R, R])$. By Lemma 9, $B \subseteq Z(R)$. So we assume that $T(A) \neq R$. If $[T(A), T(A)] = 0$, then $[A, A] = 0$ and, by Lemma 11, $A \subseteq Z(R)$. Suppose that $[T(A), T(A)] \neq 0$. By Lemma 13, $T(A)$ contains a nonzero Δ-ideal I of R. Since

$$[I, B] \subseteq A \cap B = 0,$$

we conclude that $B \subseteq Z(R)$ by Lemma 7. □

The map

$$\partial_a : R \ni x \mapsto [a, x] \in R$$

is called *an inner derivation* of a ring R induced by $a \in R$. The set $\operatorname{IDer} R$ of all inner derivations of R is a Lie ring. Every prime Lie ring is primary Lie.

Lemma 17. *There is the Lie ring isomorphism*

$$\operatorname{IDer} R \ni \partial_a \mapsto a + Z(R) \in R^L/Z(R).$$

Proof. Evident. □

Corollary 5. *Let R be a ring. Then the following statements hold:*

(1) IDer R *is a simple Lie ring if and only if* $R^L/Z(R)$ *is a simple Lie ring,*

(2) IDer R *is a prime Lie ring if and only if* $R^L/Z(R)$ *is a prime Lie ring,*

(3) IDer R *is a semiprime Lie ring if and only if* $R^L/Z(R)$ *is a semiprime Lie ring,*

(4) IDer R *is a primary Lie ring if and only if* $R^L/Z(R)$ *is a primary Lie ring.*

References

[1] K. I. Beidar, A. V. Mikhal′ev, *Ortogonal completeness and minimal prime ideals* (in Russian), Trudy Sem. Petrovski, **10**, 1984, pp.227-234.

[2] H. E. Bell, A. A. Klein, *Combinatorial commutativity and finiteness conditions for rings*, Comm. Algebra, **29**, 2001, pp.2935-2943.

[3] J. Bergen, I. N. Herstein, E. W. Kerr, *Lie ideals and derivations of prime rings*, J. Algebra, **71**, 1981, pp.259-267.

[4] J. Bergen, S. Montgomery, D. S. Passmann, *Radicals of crossed products of enveloping algebras*, Israel J. Math., **59**, 1987, pp.167-184.

[5] M. Brešar, M. A. Chebotar and W.S. Martindale, 3rd, *Functional identities*, Birkh′auser Verlag, Basel Boston Berlin, 2000.

[6] K. R. Goodearl, R. B. Warfield, Jr., *Primitivity in differential operator rings*, Math. Z., **180**, 1982, pp.503-523.

[7] I. N. Herstein, *On the Lie and Jordan rings of a simple associative ring*, Amer. J. Math., **77**, 1955, pp.279-285.

[8] I. N. Herstein, *The Lie ring of a simple associative ring*, Duke Math. J., **22**, 1955, pp.471-476.

[9] I. N. Herstein, *Topics in Ring Theory*, The University of Chicago Press, Chicago London, 1965.

[10] I. N. Herstein, *Noncommutative rings*, The Mathematical Association of America, J. Wiley and Sons, 1968.

[11] I. N. Herstein, *On the Lie structure of an associative ring*, J. Algebra, **14**, 1970, pp.561-571.

[12] I. N. Herstein, *Rings with involution*, The University of Chicago Press, Chicago London 1976.

[13] N. Jacobson, *Lie algebras*, Interscience, New York, 1962.

[14] N. Jacobson, *Structure and representations of Jordan algebras*, Amer. Math. Soc. Colloq. Publ., V. 39, Providence, R.I., 1968.

[15] C. R. Jordan, D. A. Jordan, *The Lie structure of a commutative ring with derivation*, J. London Math. Soc.(2), **18**, 1978, pp.39-49.

[16] D. A. Jordan, *Noetherian Ore extensions and Jacobson rings*, J. London Math. Soc. (2), **10**, 1975, pp.281-291.

[17] J. Lambek, *Lectures on rings and modules*, Blaisdell Publ. Co. A division of Ginn and Co. Waltham Mass. Toronto London, 1966.

[18] W. S. Martindale, 3rd, *Lie isomorphisms of prime rings*, Trans. Amer. Math. Soc., **142**, 1969, pp.437-455.

[19] J. C. McConnell, J. C. Robson, *Noncommutative Noetherian rings*. With the cooperation of L. W. Small. Revised edition. Grad. Stud. in Math., 30. Amer. Math. Soc., Providence, RI, 2001.

[20] K. McCrimmon, *On Herstein's theorem relating Jordan and associative algebras*, J. Algebra, **13**, 1969, pp.382-392.

[21] A. Nowicki, *The Lie structure of a commutative ring with a derivation*, Arch. Math., **45**, 1985, pp.328-335.

[22] E. C. Posner, *Differentiably simple rings*, Proc. Amer. Math. Soc., **11**, 1960, pp.337-343.

[23] I. I. Zuev, *Lie ideals of associative rings* (in Russian), Uspehi Mat. Nauk, **18**, 1963, pp.155–158.

Contact information

O. D. Artemovych Institute of Mathematics
Cracow University of Technology
ul. Warszawska 24
Cracow 31-155 POLAND
E-Mail(*s*): artemo@usk.pk.edu.pl

M. P. Lukashenko Faculty of Mathematics and Informatics
PreCarpathian National University of Vasyl Stefanyk
Shevchenko St 57
Ivano-Frankivsk 76025 UKRAINE
E-Mail(*s*): bilochka.90@mail.ru

Received by the editors: 22.01.2015
and in final form 22.03.2015.

Algebra and Discrete Mathematics
Volume **20** (2015). Number 1, pp. 32–39

RESEARCH ARTICLE

On characteristic properties of semigroups

Vitaliy M. Bondarenko, Yaroslav V. Zaciha

Communicated by V. V. Kirichenko

Abstract. Let $\mathcal{K}$ be a class of semigroups and $\mathcal{P}$ be a set of general properties of semigroups. We call a subset Q of $\mathcal{P}$ characteristic for a semigroup $S \in \mathcal{K}$ if, up to isomorphism and anti-isomorphism, S is the only semigroup in $\mathcal{K}$, which satisfies all the properties from Q. The set of properties $\mathcal{P}$ is called char-complete for $\mathcal{K}$ if for any $S \in \mathcal{K}$ the set of all properties $P \in \mathcal{P}$, which hold for the semigroup S, is characteristic for S. We indicate a 7-element set of properties of semigroups which is a minimal char-complete set for the class of semigroups of order 3.

Introduction

All properties of semigroups are assumed to be invariant with respect to isomorphism and anti-isomorphism.

Let $\mathcal{K}$ be a class of semigroups and $\mathcal{P}$ be some set of general (qualitative and quantitative) properties of semigroups. For $S \in \mathcal{K}$, we denote by $\mathcal{P}(S)$ the set of all properties $P \in \mathcal{P}$ which hold for the semigroup S.

We say that a subset Q of $\mathcal{P}$ *characteristic for a semigroup* $S \in \mathcal{K}$ if, up to isomorphism and anti-isomorphism, S is the only semigroup in $\mathcal{K}$, which satisfies all the properties from Q; if $Q = \{q_1, \ldots, q_s\}$, then the properties $q_1, \ldots, q_s$ are called *characteristic for* S. Obviously, if $Q \subset Q' \subseteq \mathcal{P}$ and Q is characteristic for S then so is Q'. The set of properties $\mathcal{P}$ is called *char-complete for* $\mathcal{K}$ if for any $S \in \mathcal{K}$ the subset $\mathcal{P}(S)$ of $\mathcal{P}$ is characteristic

2010 MSC: 20M.
Key words and phrases: semigroup, anti-isomorphism, idempotent, Cayley table, characteristic property, char-complete set.

for S. A char-complete set of properties is called *minimal* if it does not contain a proper char-complete subset of ones.

In this paper we indicate a 7 properties of semigroups which form a minimal char-complete set for the class of semigroups of order 3.

1. Formulation of the main result

We consider the following properties of a semigroup of order 3:

$P(C)$: commutativity;

$P(1)$: the existence of a unit element;

$P(0)$: the existence of a zero element;

$P^+(0)$: the existence of an added zero element;

$P_{id}(1)$: the number of idempotents is equal to 1;

$P_{id}(2)$: the number of idempotents is equal to 2;

$P_{gen}(2)$: the smallest number of generators is equal to 2.

The set of all these properties is denoted by $P_3(7)$.

Our aim is to prove the following theorem.

Theorem 1. *The set $P_3(7)$ is a minimal char-complete set of properties for the class of semigroups of order 3.*

2. Preliminaries

In this section we present results from the paper [1].

Let $S = \{\langle 1\rangle, \langle 2\rangle, \dots, \langle n\rangle\}$ be a finite semigroup which is given by the Cayley table T. One wants to find some its minimal system of generators and the complete set of defining relations for these generators.

In the first step one chooses an element $\langle s\rangle$ of S that is (according to the table) the product of two elements $\langle i\rangle \neq \langle s\rangle$ and $\langle j\rangle \neq \langle s\rangle$; then in the Cayley table T (including the header row and the header column) one substitutes $\langle i\rangle\langle j\rangle$ instead of $\langle s\rangle$. The new table is denoted by T_1.

In the second step one chooses an element $\langle s^{(1)}\rangle$ of the set $S^{(1)} = S\setminus\{\langle s\rangle\}$ that is the product of two elements $i^{(1)}$ and $j^{(1)}$, where $i^{(1)} = \langle i_1\rangle$, $j^{(1)} = \langle j_1\rangle$, or $i^{(1)} = \langle i_1\rangle\langle i_2\rangle$, $j^{(1)} = \langle j_1\rangle$, or $i^{(1)} = \langle i_1\rangle$, $j^{(1)} = \langle j_1\rangle\langle j_2\rangle$, or $i^{(1)} = \langle i_1\rangle\langle i_2\rangle$, $j^{(1)} = \langle j_1\rangle\langle j_2\rangle$, with $\langle i_1\rangle, \langle i_2\rangle, \langle j_1\rangle, \langle j_2\rangle \neq \langle s^1\rangle$; then in the table T_1 (including the header row and header column) one substitutes $\langle i^1\rangle\langle j^1\rangle$ instead of $\langle s^1\rangle$. The new table is denoted by T_2.

In the next step one chooses an element $s^{(2)}$ of the set $S^{(2)} = S \setminus \{\langle s\rangle, \langle s^{(1)}\rangle\}$, and so on. Upon completion of this process, say after m steps ($m \geqslant 0$), one has a minimal system of generators $S^{(m)}$ of the semigroup S (consisting of those elements of the header column of the last table that are of the form $\langle k\rangle$) and an appropriate set of defining relations in the form of the last table (which must be taken fully).

Note that the specified process is ambiguous and so one can get different final system of generators.

In the paper [1] this algorithm is applied to all semigroups of order 3 which are considered up to isomorphism and anti-isomorphism (if S be a semigroup, then a semigroup S' with multiplication $\circ$ is called anti-isomorphic to S if $S' = S$ as sets and $x \circ y = yx$). In each from 18 cases the algorithm has less than 3 steps. Under the transition from one table to another, the equality between the arrows specifies a replacement in the first table.

1)

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 1\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 2\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$

$\Rightarrow (\langle 0\rangle = \langle 2\rangle^2) \Rightarrow$

	$\boldsymbol{\langle 2\rangle^2}$	$\langle 1\rangle$	$\langle 2\rangle$
$\boldsymbol{\langle 2\rangle^2}$	$\boldsymbol{\langle 2\rangle^2}$	$\boldsymbol{\langle 2\rangle^2}$	$\boldsymbol{\langle 2\rangle^2}$
$\langle 1\rangle$	$\boldsymbol{\langle 2\rangle^2}$	$\boldsymbol{\langle 2\rangle^2}$	$\boldsymbol{\langle 2\rangle^2}$
$\langle 2\rangle$	$\boldsymbol{\langle 2\rangle^2}$	$\boldsymbol{\langle 2\rangle^2}$	$\boldsymbol{\langle 2\rangle^2}$

2)

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 1\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 2\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 1\rangle$

$\Rightarrow (\langle 1\rangle = \langle 2\rangle^2) \Rightarrow$

	$\langle 0\rangle$	$\boldsymbol{\langle 2\rangle^2}$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\boldsymbol{\langle 2\rangle^2}$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 2\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\boldsymbol{\langle 2\rangle^2}$

$\Rightarrow$

$\Rightarrow (\langle 0\rangle = \langle 2\rangle^2 \cdot \langle 2\rangle = \langle 2\rangle^3) \Rightarrow$

	$\boldsymbol{\langle 2\rangle^3}$	$\langle 2\rangle^2$	$\langle 2\rangle$
$\boldsymbol{\langle 2\rangle^3}$	$\boldsymbol{\langle 2\rangle^3}$	$\boldsymbol{\langle 2\rangle^3}$	$\boldsymbol{\langle 2\rangle^3}$
$\langle 2\rangle^2$	$\boldsymbol{\langle 2\rangle^3}$	$\boldsymbol{\langle 2\rangle^3}$	$\boldsymbol{\langle 2\rangle^3}$
$\langle 2\rangle$	$\boldsymbol{\langle 2\rangle^3}$	$\boldsymbol{\langle 2\rangle^3}$	$\langle 2\rangle^2$

3)

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 1\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 2\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 2\rangle$

$\Rightarrow (\langle 0\rangle = \langle 1\rangle \cdot \langle 2\rangle) \Rightarrow$

$\Rightarrow$

	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\langle 1\rangle$	$\langle 2\rangle$
$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$
$\langle 1\rangle$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$
$\langle 2\rangle$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\langle 2\rangle$

4)

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 1\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 2\rangle$	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$

$\Rightarrow (\langle 0\rangle = \langle 1\rangle \cdot \langle 2\rangle) \Rightarrow$

$\Rightarrow$

	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\langle 1\rangle$	$\langle 2\rangle$
$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$
$\langle 1\rangle$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$
$\langle 2\rangle$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\langle 1\rangle$	$\langle 2\rangle$

5)

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 1\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 2\rangle$	$\langle 2\rangle$	$\langle 2\rangle$	$\langle 2\rangle$

$\Rightarrow (\langle 0\rangle = \langle 1\rangle \cdot \langle 2\rangle) \Rightarrow$

$\Rightarrow$

	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\langle 1\rangle$	$\langle 2\rangle$
$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$
$\langle 1\rangle$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$
$\langle 2\rangle$	$\langle 2\rangle$	$\langle 2\rangle$	$\langle 2\rangle$

6)

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 1\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 1\rangle$
$\langle 2\rangle$	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$

$\Rightarrow (\langle 0\rangle = \langle 1\rangle^2) \Rightarrow$

	$\mathbf{\langle 1\rangle^2}$	$\langle 1\rangle$	$\langle 2\rangle$
$\mathbf{\langle 1\rangle^2}$	$\mathbf{\langle 1\rangle^2}$	$\mathbf{\langle 1\rangle^2}$	$\mathbf{\langle 1\rangle^2}$
$\langle 1\rangle$	$\mathbf{\langle 1\rangle^2}$	$\mathbf{\langle 1\rangle^2}$	$\langle 1\rangle$
$\langle 2\rangle$	$\mathbf{\langle 1\rangle^2}$	$\langle 1\rangle$	$\langle 2\rangle$

7)

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 1\rangle$	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 0\rangle$
$\langle 2\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 2\rangle$

$\Rightarrow (\langle 0\rangle = \langle 1\rangle \cdot \langle 2\rangle) \Rightarrow$

$\Rightarrow$

	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\langle 1\rangle$	$\langle 2\rangle$
$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$
$\langle 1\rangle$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\langle 1\rangle$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$
$\langle 2\rangle$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\langle 2\rangle$

8)

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 1\rangle$	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 0\rangle$
$\langle 2\rangle$	$\langle 2\rangle$	$\langle 2\rangle$	$\langle 2\rangle$

$\Rightarrow (\langle 0\rangle = \langle 1\rangle \cdot \langle 2\rangle) \Rightarrow$

$\Rightarrow$

	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\langle 1\rangle$	$\langle 2\rangle$
$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$
$\langle 1\rangle$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$	$\langle 1\rangle$	$\mathbf{\langle 1\rangle \cdot \langle 2\rangle}$
$\langle 2\rangle$	$\langle 2\rangle$	$\langle 2\rangle$	$\langle 2\rangle$

9)

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 1\rangle$	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 1\rangle$
$\langle 2\rangle$	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 1\rangle$

$\Rightarrow (\langle 1\rangle = \langle 2\rangle^2) \Rightarrow$

	$\langle 0\rangle$	$\mathbf{\langle 2\rangle^2}$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\mathbf{\langle 2\rangle^2}$	$\langle 0\rangle$	$\mathbf{\langle 2\rangle^2}$	$\mathbf{\langle 2\rangle^2}$
$\langle 2\rangle$	$\langle 0\rangle$	$\mathbf{\langle 2\rangle^2}$	$\mathbf{\langle 2\rangle^2}$

10)

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 1\rangle$	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 1\rangle$
$\langle 2\rangle$	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$

11)

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 1\rangle$	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 1\rangle$
$\langle 2\rangle$	$\langle 0\rangle$	$\langle 2\rangle$	$\langle 2\rangle$

12)

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 1\rangle$	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 2\rangle$	$\langle 0\rangle$	$\langle 2\rangle$	$\langle 1\rangle$

$\Rightarrow (\langle 1\rangle = \langle 2\rangle^2) \Rightarrow$

	$\langle 0\rangle$	$\boldsymbol{\langle 2\rangle^2}$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\boldsymbol{\langle 2\rangle^2}$	$\langle 0\rangle$	$\boldsymbol{\langle 2\rangle^2}$	$\langle 2\rangle$
$\langle 2\rangle$	$\langle 0\rangle$	$\langle 2\rangle$	$\boldsymbol{\langle 2\rangle^2}$

13)

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 1\rangle$	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 2\rangle$	$\langle 2\rangle$	$\langle 2\rangle$	$\langle 2\rangle$

14)

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 1\rangle$	$\langle 1\rangle$	$\langle 1\rangle$	$\langle 1\rangle$
$\langle 2\rangle$	$\langle 2\rangle$	$\langle 2\rangle$	$\langle 2\rangle$

15)

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 2\rangle$
$\langle 1\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 2\rangle$
$\langle 2\rangle$	$\langle 2\rangle$	$\langle 2\rangle$	$\langle 0\rangle$

$\Rightarrow (\langle 0\rangle = \langle 2\rangle^2) \Rightarrow$

	$\boldsymbol{\langle 2\rangle^2}$	$\langle 1\rangle$	$\langle 2\rangle$
$\boldsymbol{\langle 2\rangle^2}$	$\boldsymbol{\langle 2\rangle^2}$	$\boldsymbol{\langle 2\rangle^2}$	$\langle 2\rangle$
$\langle 1\rangle$	$\boldsymbol{\langle 2\rangle^2}$	$\boldsymbol{\langle 2\rangle^2}$	$\langle 2\rangle$
$\langle 2\rangle$	$\langle 2\rangle$	$\langle 2\rangle$	$\boldsymbol{\langle 2\rangle^2}$

16)

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 0\rangle$	$\langle 2\rangle$
$\langle 1\rangle$	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 2\rangle$	$\langle 2\rangle$	$\langle 2\rangle$	$\langle 0\rangle$

$\Rightarrow (\langle 0\rangle = \langle 2\rangle^2) \Rightarrow$

	$\boldsymbol{\langle 2\rangle^2}$	$\langle 1\rangle$	$\langle 2\rangle$
$\boldsymbol{\langle 2\rangle^2}$	$\boldsymbol{\langle 2\rangle^2}$	$\boldsymbol{\langle 2\rangle^2}$	$\langle 2\rangle$
$\langle 1\rangle$	$\boldsymbol{\langle 2\rangle^2}$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 2\rangle$	$\langle 2\rangle$	$\langle 2\rangle$	$\boldsymbol{\langle 2\rangle^2}$

17)

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 1\rangle$
$\langle 1\rangle$	$\langle 1\rangle$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 2\rangle$	$\langle 1\rangle$	$\langle 0\rangle$	$\langle 0\rangle$

$\Rightarrow (\langle 1\rangle = \langle 0\rangle \cdot \langle 2\rangle) \Rightarrow$

$\Rightarrow$

	$\langle 0\rangle$	$\boldsymbol{\langle 0\rangle \cdot \langle 2\rangle}$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\boldsymbol{\langle 0\rangle \cdot \langle 2\rangle}$	$\boldsymbol{\langle 0\rangle \cdot \langle 2\rangle}$
$\boldsymbol{\langle 0\rangle \cdot \langle 2\rangle}$	$\boldsymbol{\langle 0\rangle \cdot \langle 2\rangle}$	$\langle 0\rangle$	$\langle 0\rangle$
$\langle 2\rangle$	$\boldsymbol{\langle 0\rangle \cdot \langle 2\rangle}$	$\langle 0\rangle$	$\langle 0\rangle$

$\Rightarrow$

$$\Rightarrow (\langle 0\rangle = \langle 2\rangle^2) \Rightarrow$$

	$\mathbf{\langle 2\rangle^2}$	$\langle 2\rangle^3$	$\langle 2\rangle$
$\mathbf{\langle 2\rangle^2}$	$\mathbf{\langle 2\rangle^2}$	$\langle 2\rangle^3$	$\langle 2\rangle^3$
$\langle 2\rangle^3$	$\langle 2\rangle^3$	$\mathbf{\langle 2\rangle^2}$	$\mathbf{\langle 2\rangle^2}$
$\langle 2\rangle$	$\langle 2\rangle^3$	$\mathbf{\langle 2\rangle^2}$	$\mathbf{\langle 2\rangle^2}$

18)

	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
$\langle 1\rangle$	$\langle 1\rangle$	$\langle 2\rangle$	$\langle 0\rangle$
$\langle 2\rangle$	$\langle 2\rangle$	$\langle 0\rangle$	$\langle 1\rangle$

$\Rightarrow (\langle 2\rangle = \langle 1\rangle^2) \Rightarrow$

	$\langle 0\rangle$	$\langle 1\rangle$	$\mathbf{\langle 1\rangle^2}$
$\langle 0\rangle$	$\langle 0\rangle$	$\langle 1\rangle$	$\mathbf{\langle 1\rangle^2}$
$\langle 1\rangle$	$\langle 1\rangle$	$\mathbf{\langle 1\rangle^2}$	$\langle 0\rangle$
$\mathbf{\langle 1\rangle^2}$	$\mathbf{\langle 1\rangle^2}$	$\langle 0\rangle$	$\langle 1\rangle$

$\Rightarrow$

$$\Rightarrow (\langle 0\rangle = \langle 1\rangle \cdot \langle 1\rangle^2 = \langle 1\rangle^3) \Rightarrow$$

	$\mathbf{\langle 1\rangle^3}$	$\langle 1\rangle$	$\langle 1\rangle^2$
$\mathbf{\langle 1\rangle^3}$	$\mathbf{\langle 1\rangle^3}$	$\langle 1\rangle$	$\langle 1\rangle^2$
$\langle 1\rangle$	$\langle 1\rangle$	$\langle 1\rangle^2$	$\mathbf{\langle 1\rangle^3}$
$\langle 1\rangle^2$	$\langle 1\rangle^2$	$\mathbf{\langle 1\rangle^3}$	$\langle 1\rangle$

From the above, it easily follows the next statement (which was not formulated in [1]) .

Theorem 2. *Let S be a semigroup of order* 3. *Then any two its minimal systems of generators are of the same order, and coincide if S is not a group.*

Note that the group of order 3 is given by the case 18).

3. Proof of Theorem 1

The following table $\mathcal{T}$, which follows from the results of section 2, shows what the properties hold for the semigroups 1) – 18) ("+" means that the corresponding property holds, and its absence means that the corresponding property does not hold).

Since all rows of this table (without the header row and the header column) are mutually different, the set of properties $P_3(7)$ is char-complete.

To prove that the char-complete set $P_3(7)$ is minimal it suffices to verify that the table $\mathcal{T}$ without any fix column X (and, of course, without the header row and the header column) has two equal rows. It is easy to see that if X is equal to $C, P(1), P(0), P^+(0), P_{id}(1), P_{id}(2), P_{gen}(2)$, then, respectively, the following two rows are equal: 3 and 4, 13 and 14, 4 and 5, 3 and 9, 1 and 7, 5 and 8, 8 and 14.

Table 1. $\mathcal{T}$

	C	$P(1)$	$P(0)$	$P^+(0)$	$P_{id}(1)$	$P_{id}(2)$	$P_{gen}(2)$
1	+		+		+		+
2	+		+		+		
3	+		+			+	+
4			+			+	+
5						+	+
6	+	+	+			+	+
7	+		+				+
8							+
9	+		+	+		+	+
10	+	+	+	+			
11			+	+			
12	+	+	+	+		+	+
13		+					
14							
15	+				+		+
16	+	+				+	+
17	+				+		
18	+	+			+		

References

[1] V. M. Bondarenko, Y. V. Zaciha, *On defining relations for minimal generator systems of three-order semigroups*, Science Journal of National Pedagogical Dragomanov University, Series 1: Physics and Mathematics (2013), no. 14, 62-67 (in Ukrainian).

Contact information

V. M. Bondarenko, Y. V. Zaciha
Institute of Mathematics, Tereshchenkivska 3, 01601 Kyiv, Ukraine
E-Mail(s): vit-bond@imath.kiev.ua, zaciha@mail.ru

Received by the editors: 07.09.2015
and in final form 07.09.2015.

Algebra and Discrete Mathematics
Volume **20** (2015). Number 1, pp. 40–54

RESEARCH ARTICLE

New families of Jacobsthal and Jacobsthal-Lucas numbers

Paula Catarino, Paulo Vasco, Helena Campos, Ana Paula Aires and Anabela Borges

Communicated by V. Mazorchuk

Abstract. In this paper we present new families of sequences that generalize the Jacobsthal and the Jacobsthal-Lucas numbers and establish some identities. We also give a generating function for a particular case of the sequences presented.

Introduction

Several sequences of positive integers were and still are object of study for many researchers. Examples of these sequences are the well known Fibonacci sequence and the Lucas sequence, both related with the golden mean, with so many applications in diverse fields such as mathematics, engineering, biology, physics, architecture, stock market investing, among others (see [9] and [17]). About these and other sequences like Pell sequence, Pell-Lucas sequence, Modified Pell sequence, Jacobsthal sequence and the Jacobsthal-Lucas sequence, among others, there is a vast literature where several properties are studied and well known identities are derived, see for example, [13, 18–20].

In 1965, Horadam studied some properties of sequences of the type, $w_n(a, b; p, q)$, where $a,\ b$ are nonnegative integers and $p,\ q$ are arbitrary

2010 MSC: 11B37, 11B83, 05A15.
Key words and phrases: Jacobsthal numbers, Jacobsthal-Lucas numbers, Binet formula, Generating matrix, Generating function.

integers, see [11] and [12]. Such sequences are defined by the recurrence relations of second order

$$w_n = pw_{n-1} - qw_{n-2}, (n \geqslant 2)$$

with initial conditions $w_0 = a$, $w_1 = b$. For example, the Fibonacci and the Lucas sequences can be considered as special cases of sequences of this type, $w_n(1, 1; 1, -1)$ and $w_n(2, 1; 1, -1)$, respectively. Also, the Jacobsthal and the Jacobsthal-Lucas sequences can be considered as $w_n(0, 1; 1, -2)$ and $w_n(2, 1; 1, -2)$, respectively. Recall that the second-order recurrence relations and the initial conditions for the Jacobsthal numbers, $J_n,\ n \geqslant 0$, and for the Jacobsthal-Lucas numbers, $j_n,\ n \geqslant 0$, respectively, are given by

$$J_{n+2} = J_{n+1} + 2J_n,\ J_0 = 0,\ J_1 = 1$$

and

$$j_{n+2} = j_{n+1} + 2j_n,\ j_0 = 2,\ j_1 = 1.$$

The Binet formulae for these sequences are

$$J_n = \frac{2^n - (-1)^n}{3} \text{ and } j_n = 2^n + (-1)^n,$$

where 2 and -1 are the roots of the characteristic equation associated with the above recurrence relations.

More recently, some of these sequences were generalized for any positive real number k. The studies of k-Fibonacci sequence, k-Lucas sequence, k-Pell sequence, k-Pell-Lucas sequence, Modified k-Pell sequence, k-Jacobhstal and k-Jacobhstal-Lucas sequence, can be found in [1,3–7,14].

The aim of this work is to study some properties of two new sequences that generalize the Jacobhstal and the Jacobsthal-Lucas numbers. In this work we will follow closely the work of El-Mikkawy and Sogabe (see [10]) where the authors give a new family that generalizes the Fibonacci numbers, different from the one defined in [1], and establish relations with the ordinary Fibonacci numbers.

So, in this Section we start giving the new definition of generalized Jacobsthal and Jacobsthal-Lucas numbers, and we exhibit some elements of them. We also present relations of these sequences with ordinary Jacobsthal and Jacobsthal-Lucas. In Section 1 we deduce some properties of these new families, as well as in Section 2, but using different methods. In Section 3, we study a particular case, that is two sequences of the new defined families for $k = 2$. For these sequences we present some recurrence relations and generating functions.

Following our ideas, we give a new definition of generalized Jacobsthal and Jacobsthal-Lucas numbers.

Definition 1. Let n be a nonnegative integer and k be a natural number. By the division algorithm there exist unique numbers m and r such that $n = mk + r$ $(0 \leqslant r < k)$. Using these parameters we define the new generalized Jacobsthal and generalized Jacobsthal-Lucas numbers, $J_n^{(k)}$ and $j_n^{(k)}$ respectively by

$$J_n^{(k)} = \frac{1}{(r_1 - r_2)^k} \left(r_1^{m+1} - r_2^{m+1} \right)^r (r_1^m - r_2^m)^{k-r} \tag{1}$$

and

$$j_n^{(k)} = \left(r_1^{m+1} + r_2^{m+1} \right)^r (r_1^m + r_2^m)^{k-r}, \tag{2}$$

where $r_1 = 2$, $r_2 = -1$, respectively.

For $k = 1, 2, 3$ the first seven elements of these new sequences are:

$$\{J_n^{(1)}\}_{n=0}^5 = \{0, 1, 1, 3, 5, 11, 21\} \quad \{j_n^{(1)}\}_{n=0}^5 = \{2, 1, 5, 7, 17, 31, 65\}$$

$$\{J_n^{(2)}\}_{n=0}^5 = \{0, 0, 1, 1, 1, 3, 9\} \qquad \{j_n^{(2)}\}_{n=0}^5 = \{4, 2, 1, 5, 25, 35, 49\}$$

$$\{J_n^{(3)}\}_{n=0}^5 = \{0, 0, 0, 1, 1, 1, 1\} \qquad \{j_n^{(3)}\}_{n=0}^5 = \{8, 4, 2, 1, 5, 25, 125\}.$$

We also present more elements of some of these new sequences in the tables 1 and 2. We have found some interesting regularities. In the case of the generalized Jacobsthal sequences $\left\{J_n^{(k)}\right\}_n$ it is easy to prove that:

Proposition 1. *Let $J_i^{(k)}$ be the i^{th} term of the new family of Jacobsthal numbers. Then we have:*

a) $J_i^{(k)} = 0, \quad i \in \{0, \ldots, k-1\}$;

b) $J_i^{(k)} = 1, \quad i \in \{k, \ldots, k-1\}$;

c) $J_i^{(k)} = 3^{i-2k}, \quad i \in \{2k, \ldots, 3k\}$.

To the generalized Jacobsthal-Lucas sequences $\left\{j_n^{(k)}\right\}_n$ is easy to prove that:

Table 1. $J_n^{(k)}$, for $k = 1, 2, \ldots, 9$ and $n = 0, 1, \ldots, 27$.

n\k	1	2	3	4	5	6	7	8	9
0	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0
2	1	1	0	0	0	0	0	0	0
3	3	1	1	0	0	0	0	0	0
4	5	1	1	1	0	0	0	0	0
5	11	3	1	1	1	0	0	0	0
6	21	9	1	1	1	1	0	0	0
7	43	15	3	1	1	1	1	0	0
8	85	25	9	1	1	1	1	1	0
9	171	55	27	3	1	1	1	1	1
10	341	121	45	9	1	1	1	1	1
11	683	231	75	27	3	1	1	1	1
12	1365	441	125	81	9	1	1	1	1
13	2731	903	275	135	27	3	1	1	1
14	5461	1849	605	225	81	9	1	1	1
15	10923	3655	1331	375	243	27	3	1	1
16	21845	7225	2541	625	405	81	9	1	1
17	43691	14535	4851	1375	675	243	27	3	1
18	87381	29241	9261	3025	1125	729	81	9	1
19	174763	58311	18963	6655	1875	1215	243	27	3
20	349525	116281	38829	14641	3125	2025	729	81	9
21	699051	232903	79507	27951	6875	3375	2187	243	27
22	1398101	466489	157165	53361	15125	5625	3645	729	81
23	2796203	932295	310675	101871	33275	9375	6075	2187	243
24	5592405	1863225	614125	194481	73205	15625	10125	6561	729
25	11184811	3727815	1235475	398223	161051	34375	16875	10935	2187
26	22369621	7458361	2485485	815409	307461	75625	28125	18225	6561
27	44739243	14913991	5000211	1669647	586971	166375	46875	30375	19683

Table 2. $j_n^{(k)}$, for $k = 1, 2, \ldots, 9$ and $n = 0, 1, \ldots, 27$.

n\k	1	2	3	4	5	6	7	8	9
0	2	4	8	16	32	64	128	256	512
1	1	2	4	8	16	32	64	128	256
2	5	1	2	4	8	16	32	64	128
3	7	5	1	2	4	8	16	32	64
4	17	25	5	1	2	4	8	16	32
5	31	35	25	5	1	2	4	8	16
6	65	49	125	25	5	1	2	4	8
7	127	119	175	125	25	5	1	2	4
8	257	289	245	625	125	25	5	1	2
9	511	527	343	875	625	125	25	5	1
10	1025	961	833	1225	3125	625	125	25	5
11	2047	2015	2023	1715	4375	3125	625	125	25
12	4097	4225	4913	2401	6125	15625	3125	625	125
13	8191	8255	8959	5831	8575	21875	15625	3125	625
14	16385	16129	16337	14161	12005	30625	78125	15625	3125
15	32767	32639	29791	34391	16807	42875	109375	78125	15625
16	65537	66049	62465	83521	40817	60025	153125	39062	78125
17	131071	131327	130975	152303	99127	84035	214375	546875	390625
18	262145	261121	274625	277729	240737	117649	300125	765625	1953125
19	524287	523775	536575	506447	584647	285719	420175	1071875	2734375
20	1048577	1050625	1048385	923521	1419857	693889	588245	1500625	3828125
21	2097151	2098175	2048383	1936415	2589151	1685159	823543	2100875	5359375
22	4194305	4190209	4145153	4060225	4721393	4092529	2000033	2941225	7503125
23	8388607	8386559	8388223	8513375	8609599	9938999	4857223	4117715	10504375
24	16777217	16785409	16974593	17850625	15699857	24137569	11796113	5764801	14706125
25	33554431	33558527	33751039	34877375	28629151	44015567	28647703	14000231	20588575
26	67108865	67092481	67108097	68145025	60028865	80263681	69572993	34000561	28824005
27	134217727	134209535	133432831	133144895	125866975	146363183	168962983	82572791	40353607

Proposition 2. *Let $j_i^{(k)}$ be the i^{th} term of the new family of Jacobsthal-Lucas numbers. Then we have:*

a) $j_i^{(k)} = 2^{k-i}, \quad i \in \{0, \ldots, k-1\}$;

b) $j_i^{(k)} = 5^{i-k}, \quad i \in \{k, \ldots, 2k\}$.

The generalized Jacobsthal and Jacobsthal-Lucas numbers have the following relations with the ordinary Jacobsthal and Jacobsthal-Lucas numbers.

Lemma 1. *Given n a nonnegative integer and k a natural number*

$$J_{mk+r}^{(k)} = (J_m)^{k-r}(J_{m+1})^r$$

and

$$j_{mk+r}^{(k)} = (j_m)^{k-r}(j_{m+1})^r,$$

where m and r are nonnegative integers such that $n = mk+r$ $(0 \leqslant r < k)$.

Proof. We have

$$\begin{aligned}
(J_m)^{k-r}(J_{m+1})^r &= \left(\frac{2^m - (-1)^m}{3}\right)^{k-r} \left(\frac{2^{m+1} - (-1)^{m+1}}{3}\right)^r \\
&= \frac{1}{3^k}\left(2^m - (-1)^m\right)^{k-r} \left(2^{m+1} - (-1)^{m+1}\right)^r \\
&= \frac{1}{(r_1 - r_2)^k}\left(r_1^m - r_2^m\right)^{k-r} \left(r_1^{m+1} - r_2^{m+1}\right)^r \\
&= J_{mk+r}^{(k)}.
\end{aligned}$$

In a similar way we can show the second equality. □

Note that the use of the Lemma 1 allows us to conclude immediately that $J_n^{(1)}$ and $j_n^{(1)}$ are the Jacobsthal and the Jacobsthal-Lucas numbers, respectively.

1. Properties

Next we present some properties of these new families of integers.

Theorem 1. *Let k and m be fixed numbers where m is a nonnegative integer and k a natural number. The generalized Jacobsthal numbers and the ordinary Jacobsthal numbers satisfy:*

a) $\sum_{i=0}^{k-1} \binom{k-1}{i}(-1)^{-i} J_{mk+i}^{(k)} = (-2)^{k-1} J_m J_{(m-1)(k-1)}^{(k-1)}$;

b) $\sum_{i=0}^{k-1} \binom{k-1}{i} 2^{k-i-1} J_{mk+i}^{(k)} = J_m J_{(m+2)(k-1)}^{(k-1)}$;

c) $\sum_{i=0}^{k-1} J_{mk+i}^{(k)} = \frac{J_m}{2J_{m-1}} \left(J_{(m+1)k}^{(k)} - J_{mk}^{(k)} \right)$.

Proof. a) By Lemma 1 we have that $\sum_{i=0}^{k-1} \binom{k-1}{i}(-1)^{-i} J_{mk+i}^{(k)}$ is successively equal to

$$\begin{aligned}
&\sum_{i=0}^{k-1} \binom{k-1}{i} (-1)^{-i} (-1)^{k-1} (-1)^{1-k} (J_m)^{k-i} (J_{m+1})^i \\
&\quad = (-1)^{1-k} \sum_{i=0}^{k-1} \binom{k-1}{i} (-1)^{k-1-i} (J_m)^{k-1-i} J_m (J_{m+1})^i \\
&\quad = (-1)^{1-k} J_m \sum_{i=0}^{k-1} \binom{k-1}{i} (-J_m)^{k-1-i} (J_{m+1})^i,
\end{aligned}$$

that by the binomial theorem is equal to

$$(-1)^{1-k} J_m \left(J_{m+1} - J_m \right)^{k-1}.$$

Since, by the definition of the Jacobsthal sequence,

$$(-1)^{1-k} J_m \left(J_{m+1} - J_m \right)^{k-1} = (-1)^{1-k} J_m \left(2J_{m-1} \right)^{k-1}$$

and using Lemma 1 (considering $m-1$ instead of m, $k-1$ instead of k and $r=0$) we obtain

$$(-1)^{1-k} 2^{k-1} J_m J_{(m-1)(k-1)}^{(k-1)},$$

and the result follows.

b) By Lemma 1 we have

$$\begin{aligned}
\sum_{i=0}^{k-1} \binom{k-1}{i} 2^{k-i-1} J_{mk+i}^{(k)} &= \sum_{i=0}^{k-1} \binom{k-1}{i} (J_m)^{k-i} 2^{k-i-1} (J_{m+1})^i \\
&= \sum_{i=0}^{k-1} \binom{k-1}{i} (2J_m)^{k-i-1} J_m (J_{m+1})^i \\
&= J_m \sum_{i=0}^{k-1} \binom{k-1}{i} (2J_m)^{k-i-1} (J_{m+1})^i
\end{aligned}$$

and using again the binomial theorem we have

$$J_m(J_{m+1}+2J_m)^{k-1},$$

that is equal, by the definition of the Jacobsthal numbers, to

$$J_m(J_{m+2})^{k-1}$$

and by Lemma 1 (considering $m+2$ instead of m, $k-1$ instead of k and $r=0$), we get

$$J_m J^{(k-1)}_{(m+2)(k-1)}.$$

c) By Lemma 1 we can write

$$\begin{aligned}
\sum_{i=0}^{k-1} J^{(k)}_{mk+i} &= \sum_{i=0}^{k-1} (J_m)^{k-i}(J_{m+1})^i \\
&= (J_m)^k \sum_{i=0}^{k-1} \left(\frac{J_{m+1}}{J_m}\right)^i \\
&= (J_m)^k \left[\frac{\left(\frac{J_{m+1}}{J_m}\right)^k - 1}{\frac{J_{m+1}}{J_m} - 1}\right] \\
&= (J_m)^k \left[\frac{(J_{m+1})^k - (J_m)^k}{(J_m)^k} \times \frac{J_m}{J_{m+1}-J_m}\right] \\
&= \frac{J_m}{J_{m+1}-J_m}\left[(J_{m+1})^k - (J_m)^k\right] \\
&= \frac{J_m}{2J_{m-1}}\left[(J_{m+1})^k - (J_m)^k\right]
\end{aligned}$$

and, taking into account Lemma 1 (with $r=0$), the result follows. □

The following result for Jacobsthal-Lucas numbers can be deduced analogously:

Theorem 2. *Let k and m be fixed numbers where m is a nonnegative integer and k a natural number. The generalized Jacobsthal-Lucas numbers and the ordinary Jacobsthal-Lucas numbers satisfy:*

a) $\sum_{i=0}^{k-1} \binom{k-1}{i}(-1)^{-i} j^{(k)}_{mk+i} = (-2)^{k-1} j_m j^{(k-1)}_{(m-1)(k-1)}$;

b) $\sum_{i=0}^{k-1} \binom{k-1}{i} 2^{k-i-1} j^{(k)}_{mk+i} = j_m j^{(k-1)}_{(m+2)(k-1)}$;

c) $\sum_{i=0}^{k-1} j^{(k)}_{mk+i} = \frac{j_m}{2j_{m-1}}\left(j^{(k)}_{(m+1)k} - j^{(k)}_{mk}\right)$.

2. Generating matrices

In [15] the authors use a matrix method for generating the Jacobsthal numbers by defining the Jacobsthal A-matrix

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$$

and they proved that

$$A^n = \begin{bmatrix} J_{n+1} & 2J_n \\ J_n & 2J_{n-1} \end{bmatrix} = A^{(n-1)} \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix},$$

for any natural number n.

Thus, for any $n \geqslant 0$, $s \geqslant 0$ and $n + s \geqslant 2$, we have

$$\begin{bmatrix} J_{n+s} & 2J_{n+s-1} \\ J_{n+s-1} & 2J_{n+s-2} \end{bmatrix} = A^{(n+s-2)} \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}.$$

If we compute the determinant of both sides of the previous equality we obtain

$$2J_{n+s}J_{n+s-2} - 2\left(J_{n+s-1}\right)^2 = -2\left|A^{(n+s-2)}\right|$$

which is equivalent to

$$\left(J_{n+s-1}\right)^2 - J_{n+s}J_{n+s-2} = (-2)^{n+s-2}.$$

Since, by Lemma 1 (where $m = n + s - 1$, $k = 2$ and $r = 0$)

$$J^{(2)}_{2(n+s-1)} = \left(J_{n+s-1}\right)^2,$$

we have proved the following result:

Theorem 3. *If $n, s \geqslant 0$ and $n + s \geqslant 2$, then*

$$J^{(2)}_{2(n+s-1)} - J_{n+s}J_{n+s-2} = (-2)^{n+s-2}.$$

Also, by considering the generating Jacobsthal-Lucas B-matrix given in [16] and in [8]

$$B = \begin{bmatrix} 5 & 2 \\ 1 & 4 \end{bmatrix}$$

and proceeding in a similar way as we did for Jacobsthal numbers, we obtain for $n, s \geqslant 0$ and $n + s \geqslant 2$,

$$\begin{bmatrix} j_{n+s} & 2j_{n+s-1} \\ j_{n+s-1} & 2j_{n+s-2} \end{bmatrix} = B^{(n+s-2)} \begin{bmatrix} 5 & 2 \\ 1 & 4 \end{bmatrix}.$$

Computing the determinant of both sides of this equality we get

$$2j_{n+s}j_{n+s-2} - 2\left(j_{n+s-1}\right)^2 = \left(3^2 2\right)^{(n+s-2)} \times \left(3^2 2\right)$$

which is equivalent to

$$j_{n+s}j_{n+s-2} - \left(j_{n+s-1}\right)^2 = 3^{2(n+s-1)}2^{(n+s-2)}.$$

Using Lemma 1 again (with $m = n + s - 1$, $k = 2$ and $r = 0$) we obtain the following result:

Theorem 4. *If $n, s \geqslant 0$ and $n + s \geqslant 2$, then*

$$j_{n+s}j_{n+s-2} - j^{(2)}_{2(n+s-1)} = 3^{2(n+s-1)}2^{(n+s-2)}.$$

3. A particular case

In this section we study the particular case of the sequences $\left\{J_n^{(2)}\right\}_n$ and $\left\{j_n^{(2)}\right\}_n$ defined by (1) and (2), respectively, with $k = 2$.

3.1. Recurrence relations

First we present a recurrence relation for these sequences.

Theorem 5. *The sequences $\left\{J_n^{(2)}\right\}_n$ and $\left\{j_n^{(2)}\right\}_n$ satisfy, respectively, the following recurrence relations:*

$$J_n^{(2)} = J_{n-1}^{(2)} + 2J_{n-3}^{(2)} + 4J_{n-4}^{(2)},\ n = 4, 5, \ldots$$

and

$$j_n^{(2)} = j_{n-1}^{(2)} + 2j_{n-3}^{(2)} + 4j_{n-4}^{(2)},\ n = 4, 5, \ldots$$

Proof. First, we consider n even, that is $n = 2m$, for any natural number m. In this case, using Lemma 1, we have

$$\begin{aligned} J_{2m}^{(2)} &= (J_m)^2 = J_m J_m \\ &= J_m \left(J_{m-1} + 2J_{m-2}\right) \\ &= J_m J_{m-1} + 2J_m J_{m-2} \\ &= J_m J_{m-1} + 2\left(J_{m-1} + 2J_{m-2}\right) J_{m-2} \\ &= J_{2m-1}^{(2)} + 2J_{m-1}J_{m-2} + 4(J_{m-2})^2 \\ &= J_{2m-1}^{(2)} + 2J_{2m-3}^{(2)} + 4(J_{m-2})^2 \\ &= J_{2m-1}^{(2)} + 2J_{2m-3}^{(2)} + 4J_{2m-4}^{(2)} \end{aligned}$$

as required. Now, for n odd, that is $n = 2m + 1$, for any natural number m and, using again Lemma 1, we obtain:

$$\begin{aligned} J_{2m+1}^{(2)} &= J_m J_{m+1} \\ &= J_m \left(J_m + 2J_{m-1}\right) \\ &= (J_m)^2 + 2J_{m-1}J_m \\ &= J_{2m}^{(2)} + 2J_{m-1}(J_{m-1} + 2J_{m-2}) \\ &= J_{2m}^{(2)} + 2(J_{m-1})^2 + 4J_{m-2}J_{m-1} \\ &= J_{2m}^{(2)} + 2J_{2m-2}^{(2)} + 4J_{m-2}J_{m-1} \\ &= J_{2m}^{(2)} + 2J_{2m-2}^{(2)} + 4J_{2m-3}^{(2)}. \end{aligned}$$

So for every $n = 4, 5, \ldots$ the result is true. In a similar way we can prove the result for $j_n^{(2)}$. □

We also note that if we consider separately the even and the odd terms of the above defined sequences we can obtain shorter recurrence relations. In fact, for $n = 2m$, for any natural number m, by Theorem 3 (with $n = m$ and $s = 1$) we have

$$J_{2m}^{(2)} - J_{m+1}J_{m-1} = (-2)^{m-1}$$

and so

$$\begin{aligned} J_{2m}^{(2)} &= J_{m-1}J_{m+1} + (-2)^{m-1} \\ &= J_{m-1}(J_m + 2J_{m-1}) + (-2)^{m-1} \\ &= J_{m-1}J_m + 2(J_{m-1})^2 + (-2)^{m-1} \\ &= J_{2m-1}^{(2)} + 2J_{2m-2}^{(2)} + (-2)^{m-1}. \end{aligned}$$

In a similar way, if we consider $n = 2m + 1$, for any natural number m, we have $J^{(2)}_{2m+1} = J_m J_{m+1}$ that is equal to

$$J_m\left(J_m + 2J_{m-1}\right) = (J_m)^2 + 2J_{m-1}J_m = J^{(2)}_{2m} + 2J^{(2)}_{2m-1}.$$

Hence, in this case, we can conclude that

$$J^{(2)}_{2m+1} = J^{(2)}_{2m} + 2J^{(2)}_{2m-1}.$$

Therefore we can conclude the following:

Proposition 3. *A shorter recurrence relation for the sequence* $\left\{J^{(2)}_n\right\}_n$ *is given by*

$$\begin{cases} J^{(2)}_{2m} = J^{(2)}_{2m-1} + 2J^{(2)}_{2m-2} + (-2)^{m-1} \\ \\ J^{(2)}_{2m+1} = J^{(2)}_{2m} + 2J^{(2)}_{2m-1} \end{cases}$$

for the even and the odd terms.

In a similar way we obtain a shorter recurrence relation to $\left\{j^{(2)}_n\right\}_n$.

Proposition 4. *A shorter recurrence relation for the sequence* $\left\{j^{(2)}_n\right\}_n$ *is given by*

$$\begin{cases} j^{(2)}_{2m} = j^{(2)}_{2m-1} + 2j^{(2)}_{2m-2} - 3^{2m}2^{m-1} \\ \\ j^{(2)}_{2m+1} = j^{(2)}_{2m} + 2j^{(2)}_{2m-1} \end{cases}$$

for the even and the odd terms.

Proof. The proof of the second identity is similar to the one in the previous proposition. To the first identity, by Theorem 4 we have:

$$j_{m+1}j_{m-1} - j^{(2)}_{2m} = 3^{2m}2^{m-1}.$$

Hence

$$\begin{aligned} j^{(2)}_{2m} &= j_{m+1}j_{m-1} - 3^{2m}2^{m-1} \\ &= j_{m-1}(j_m + 2j_{m-1}) - 3^{2m}2^{m-1} \\ &= j_{m-1}j_m + 2(j_{m-1})^2 - 3^{2m}2^{m-1} \\ &= j^{(2)}_{2m-1} + 2j^{(2)}_{2m-2} - 3^{2m}2^{m-1}. \end{aligned}$$

□

3.2. Generating Functions

Next we find generating functions for these sequences. Let us suppose that the terms of the sequences $\{J_n^{(2)}\}_n$ and $\{j_n^{(2)}\}_n$ are the coefficients of a power series centred at the origin, that is convergent in $\left]-\frac{1}{r_1}, \frac{1}{r_1}\right[$, according the Proposition 2.5 in [14] and [2], respectively, for $k = 2$.

For $\{J_n^{(2)}\}_n$ we obtain the following result:

Theorem 6. *The generating function $f^{(2)}(x)$ for $J_n^{(2)}$ is given by*

$$f^{(2)}(x) = \frac{x^2 + 2x^3}{1 - x - 2x^3 - 4x^4}.$$

Proof. To the sum of this power series,

$$f^{(2)}(x) = \sum_{n=0}^{\infty} J_n^{(2)} x^n,$$

we call generating function of the generalized Jacobsthal sequence of numbers $\{J_n^{(2)}\}_n$.

Then

$$f^{(2)}(x) - xf^{(2)}(x) - 2x^3 f^{(2)}(x) - 4x^4 f^{(2)}(x)$$

is equal to

$$\left(J_0^{(2)} + J_1^{(2)}x + J_2^{(2)}x^2 + J_3^{(2)}x^3\right) - \left(J_0^{(2)}x - J_1^{(2)}x^2 - J_2^{(2)}x^3\right) - 2J_0^{(2)}x^3 + \sum_{n=4}^{\infty}\left(J_n^{(2)} - J_{n-1}^{(2)} - 2J_{n-3}^{(2)} - 4J_{n-4}^{(2)}\right)x^n.$$

Hence, taking into account the initial conditions of the sequence $\{J_n^{(2)}\}_n$, we have

$$\left(1 - x - 2x^3 - 4x^4\right) f^{(2)}(x) = \left(0 + 0x + x^2 + x^3\right) - \left(0x - 0x^2 - x^3\right) - 2 \times 0x^3 + \sum_{n=4}^{\infty}\left(J_n^{(2)} - \left(J_{n-1}^{(2)} + 2J_{n-3}^{(2)} + 4J_{n-4}^{(2)}\right)\right)x^n.$$

Now, by Theorem 5, this is equivalent to

$$\left(1 - x - 2x^3 - 4x^4\right) f^{(2)}(x) = x^2 + 2x^3 + \sum_{n=4}^{\infty}\left(J_n^{(2)} - J_n^{(2)}\right)$$

and therefore

$$f^{(2)}(x) = \frac{x^2 + 2x^3}{1 - x - 2x^3 - 4x^4}.$$

□

Theorem 7. *The generating function $g^{(2)}(x)$ for $j_n^{(2)}$ is given by*

$$g^{(2)}(x) = \frac{4 - 2x + 3x^2 - 2x^3}{1 - x - 2x^3 - 4x^4}.$$

Proof. To the sum of this power series,

$$g^{(2)}(x) = \sum_{n=0}^{\infty} j_n^{(2)} x^n$$

we call generating function of the generalized Jacobsthal-Lucas sequence of numbers $\left\{j_n^{(2)}\right\}_n$.

Then, in a similar way as in the proof of the previous theorem, we obtain

$$\begin{aligned}\left(1 - x - 2x^3 - 4x^4\right) g^{(2)}(x) &= \left(j_0^{(2)} + j_1^{(2)} x + j_2^{(2)} x^2 + j_3^{(2)} x^3\right) \\ &\quad - \left(j_0^{(2)} x - j_1^{(2)} x^2 - j_2^{(2)} x^3\right) - 2j_0^{(2)} x^3 \\ &\quad + \sum_{n=4}^{\infty} \left(j_n^{(2)} - j_{n-1}^{(2)} - 2j_{n-3}^{(2)} - 4j_{n-4}^{(2)}\right) x^n.\end{aligned}$$

Taking into account the initial conditions of the sequence $\left\{j_n^{(2)}\right\}_n$, we have

$$\begin{aligned}&\left(1 - x - 2x^3 - 4x^4\right) g^{(2)}(x) = \left(4 + 2x + x^2 + 5x^3\right) \\ &- \left(4x - 2x^2 - x^3\right) - 8x^3 + \sum_{n=4}^{\infty} \left(j_n^{(2)} - \left(j_{n-1}^{(2)} + 2j_{n-3}^{(2)} + 4j_{n-4}^{(2)}\right)\right) x^n.\end{aligned}$$

Now, by Theorem 5, this is equivalent to

$$\left(1 - x - 2x^3 - 4x^4\right) g^{(2)}(x) = 4 - 2x + 3x^2 - 2x^3 + \sum_{n=4}^{\infty} \left(j_n^{(2)} - j_n^{(2)}\right) x^n$$

and therefore

$$g^{(2)}(x) = \frac{4 - 2x + 3x^2 - 2x^3}{1 - x - 2x^3 - 4x^4}.$$

□

4. Conclusion

In this paper we have presented new families of sequences, $J_n^{(k)}$ and $j_n^{(k)}$, that generalize the Jacobsthal and the Jacobsthal-Lucas sequences and we have established some identities involving them.

We also gave generating functions for generalized Jacobsthal and Jacobsthal-Lucas sequences $\{J_n^{(2)}\}_n$ and $\{j_n^{(2)}\}_n$.

When we were looking for more elements of these new families we have found, first, that these families were not in the Encyclopedia of Integer Sequences [21]. Furthermore, we have found some interesting regularities, stated in Propositions 1 and 2.

Acknowledgements

This work has been supported by the Portuguese Government through the FCT - Fundação para a Ciência e a Tecnologia - under the project PEst-OE/MAT/ UI4080/2014.

References

[1] C. Bolat, H. Kőse, *On the Properties of k-Fibonacci Numbers*. Int. J. Contemp. Math. Sci. 5 (22) (2010) 1097-1105.

[2] H. Campos, P. Catarino, A. P. Aires, P. Vasco, A. Borges, *On some Identities of k-Jacobsthal-Lucas Numbers*. Int. J. Math. Anal. (Ruse) 8 (10) (2014) 489-494.

[3] P. Catarino, *On Some Identities and Generating Functions for k-Pell Numbers*. Int. J. Math. Anal. (Ruse) 7 (38) (2013) 1877 - 1884.

[4] P. Catarino, *On some identities for k-Fibonacci sequence*. Int. J. Contemp. Math. Sci. 9 (1) (2014) 37-42.

[5] P. Catarino, P. Vasco, *On some Identities and Generating Functions for k-Pell-Lucas sequence*. Appl. Math. Sci. 7 (98) (2013) 4867-4873.

[6] P. Catarino, P. Vasco, *Modified k-Pell Sequence: Some Identities and Ordinary Generating Function*. Appl. Math. Sci. 7 (121) (2013) 6031 - 6037.

[7] P. Catarino, P. Vasco, *Some basic properties and a two-by-two matrix involving the k-Pell Numbers*. Int. J. Math. Anal. (Ruse) 7 (45) (2013) 2209-2215.

[8] A. Dasdemir, *On the Jacobsthal numbers by matrix method*. Fen Derg. 7 (1) (2012) 69-76.

[9] R. A. Dunlap, *The Golden Ratio and Fibonacci numbers*. World Scientific Press, Singapore, 1997.

[10] M. El-Mikkawy, T. Sogabe, *A new family of k-Fibonacci numbers*. Appl. Math. Comput. 215 (12) (2010) 4456-4461.

[11] A. F. Horadam, *Basic properties of a certain generalized sequence of numbers*. Fibonacci Quart. 3 (3) (1965) 161-176.

[12] A. F. Horadam, *Generating functions for powers of a certain generalized sequences of numbers*. Duke Math. J. 32 (3) (1965) 437-446.

[13] A. F. Horadam, *Pell identities*. Fibonacci Quart. 9 (3) (1971) 245-263.

[14] D. Jhala, K. Sisodiya, G. P. S. Rathore, *On Some Identities for k-Jacobsthal Numbers*. Int. J. Math. Anal. (Ruse) 7 (12) (2013) 551-556.

[15] F. Koken, D. Bozkurt, *On the Jacobsthal numbers by matrix method*. Int. J. Contemp. Math. Sci. 3 (13) (2008) 605-614.

[16] F. Koken, D. Bozkurt, *On the Jacobsthal-Lucas numbers by matrix method*. Int. J. Contemp. Math. Sci. 3 (33) (2008) 1629-1633.

[17] T. Koshy, *Fibonacci and Lucas numbers with applications*. John Wiley and Sons. New York, 2001.

[18] T. Koshy, *Fibonacci, Lucas, and Pell numbers, and Pascal's triangle*. Applied Probability Trust. (2011) 125-132.

[19] F. Lu, Z. Jiang, *The sum and product of Fibonacci numbers and Lucas numbers, Pell numbers and Pell-Lucas numbers representation by matrix method*. Wseas Transactions on Mathematics. 12 (4) (2013) 449-458.

[20] D. Marques, *The Order of Appearance of the Product of Consecutive Lucas Numbers*. Fibonacci Quart. 51 (1) (2013) 38-43.

[21] OEIS Foundation Inc., *The On-line Encyclopedia of Integer Sequences*. 2011. http://oeis.org.

Contact information

P. Catarino,
P. Vasco,
H. Campos,
A. P. Aires,
A. Borges

Universidade de Trás-os-Montes e Alto Douro, UTAD, Quinta de Prados, 5001-801 Vila Real, Portugal

E-Mail(*s*): pcatarin@utad.pt, pvasco@utad.pt, hcampos@utad.pt, aaires@utad.pt, aborges@utad.pt

Web-page(*s*): http://www.utad.pt

Received by the editors: 18.11.2014
and in final form 31.12.2014.

Algebra and Discrete Mathematics
Volume **20** (2015). Number 1, pp. 55–68

RESEARCH ARTICLE

Quivers of 3×3-exponent matrices

M. Dokuchaev, V. Kirichenko and M. Plakhotnyk

Dedicated to 60-th anniversary of E. Zelmanov

Abstract. We show how to use generating exponent matrices to study the quivers of exponent matrices. We also describe the admissible quivers of 3×3 exponent matrices.

Introduction

Exponent matrices were introduced in the study of semi-maximal rings (see [10]), as important ingredients of tiled orders. Recall that a *semi-maximal ring* is a semiperfect semiprime right Noetherian ring A such that for any local idempotent $e \in A$ the endomorphism ring eAe is a (non-necessarily commutative) discrete valuation ring, i.e. all principal endomorphism rings of A are discrete valuation rings (see also [3, pp. 349-350]). A square $n \times n$ matrix $A = (\alpha_{ps})$ is called an *exponent matrix* if its diagonal entries are equal to zero and for all possible indices i, j, k, one has

$$\alpha_{ij} + \alpha_{jk} \geqslant \alpha_{ik}. \tag{1}$$

Throughout this paper, unless otherwise stated, n will denote the size of the matrix under consideration. We shall refer to (1) as *ring inequalities*, and this term is explained by the following fact.

Theorem 1 ([10], [3, Th. 14.5.2])**.** *An arbitrary semi-maximal ring is isomorphic to a direct product of rings of the form*

$$\Lambda = \sum_{i,j=1}^{n} e_{ij}(\pi^{\alpha_{ij}}\mathcal{O}) \subseteq M_n(\mathcal{O}), \tag{2}$$

2010 MSC: 16H99, 16Z05.
Key words and phrases: quiver, tiled Order, exponent matrix.

where $n \geqslant 1$, $\mathcal{O}$ is a discrete valuation ring with prime element π, (α_{ij}) is an exponent matrix, $e_{ij}(\pi^{\alpha_{ij}}\mathcal{O}) = \{e_{ij}(a), a \in \pi^{\alpha_{ij}}\mathcal{O}\}$ and $e_{ij}(a)$ is the $n \times n$-matrix whose unique non-zero entry a is placed in the (i, j)-position.

The ring $\mathcal{O}$ can be embedded into classical division ring $\mathcal{D}$ and (2) is the set of all matrices $(\alpha_{ij}) \in M_n(\mathcal{D})$ such that

$$\alpha_{ij} \in \pi^{\alpha_{ij}}\mathcal{O} = e_{ii}\Lambda e_{jj},$$

where $e_{11}, \ldots, e_{nn}$ are matrix units in $M_n(\mathcal{D})$.

Clearly, $Q = M_n(\mathcal{D})$ is a classical division ring of Λ. Obviously, Λ is left and right Noetherian.

We recall next some additional definitions and facts.

Definition 1. A module M is called distributive, if so is its lattice of submodules, i.e.

$$K \cup (L + N) = L \cup L + K \cup N$$

for all submodules K, L and N.

Clearly, that every submodule of a distributive lattice is also distributive.

A direct sum of distributive modules is called a semidistributive module. A ring A is called right (left) semidistributive, if it is semidistributive as a right (left) module over itself. We say that a ring is semidistributive if it is right and left semidistributive (see. [8]).

Theorem 2 ([7])**.** *The following conditions are equivalent for a semiprime right Noetherian ring A:*

1) *A is semidistributive;*
2) *A is a direct product of a semiprime Artinian ring and a semimaximal ring.*

A tiled order Λ over a discrete valuation ring $\mathcal{O}$ is a Noetherian prime semiperfect semidistributive ring with zero Jacobson radical. In this case, $\mathcal{O} = e\Lambda e$ where $e \in \Lambda$ is a primitive idempotent. We shall write

$$\Lambda = \{\mathcal{O}, \mathcal{E}(\Lambda)\},$$

where $\mathcal{E}(\Lambda) = (\alpha_{ij})$ is the exponent matrix of Λ, i.e. Λ is of the form (2).

A tiled order is called reduced, if $\Lambda/R(\Lambda)$ is a direct product of division rings. In this case $\alpha_{ij}+\alpha_{ji}>0$ for all $i\neq j$. Exponent matrices with this property are called reduced.

Denote by $\mathcal{M}(\Lambda)$ the partially ordered set (with respect to inclusion) of all projective right Λ-modules, which are contained in some fixed Q-module W. All simple Q-modules are isomorphic to each other, whence we may take any of them. Notice, that the partially ordered sets $\mathcal{M}_l(R)$ and $\mathcal{M}_r(R)$, which correspond to left and right modules are anti-isomorphic.

The set $\mathcal{M}(\Lambda)$ is completely determined by $\mathcal{E}(\Lambda)=(\alpha_{ij})$. More precisely, if Λ is a reduced, then

$$\mathcal{M}(\Lambda)=\{P_i^z : i=1,\dots,n,\ z\in\mathbb{Z}\},$$

where

$$P_i^z\leqslant P_j^{z'}\iff\begin{cases} z-z'\geqslant\alpha_{ij}, & \text{if}\quad \mathcal{M}(\Lambda)=\mathcal{M}_l(\Lambda)\\ z-z'\geqslant\alpha_{jj}, & \text{if}\quad \mathcal{M}(\Lambda)=\mathcal{M}_r(\Lambda).\end{cases}$$

Evidently, $\mathcal{M}(\Lambda)$ is an infinite periodical set.

Let Λ and Γ be tiled orders over discrete valuated rings $\mathcal{O}$ and Δ.

Definition 2 ([10])**.** An isomorphism $\varphi:\mathcal{M}(\Lambda)\simeq\mathcal{M}(\Gamma)$ is called coordinated, if

$$B\simeq C\iff\varphi(B)\simeq\varphi(C)$$

for all $B,C\in\mathcal{M}(\Lambda)$.

Theorem 3 ([10, Prop. 2.9])**.** *The tiled orders Λ and Γ are Morita equivalent if and only if the following hold:*

1) *The discrete valuated rings $\mathcal{O}$ and Δ are isomorphic;*
2) *There is coordinated isomorphism between the partially ordered sets $\mathcal{M}(\Lambda)$ and $\mathcal{M}(\Gamma)$.*

Let I be a two sided ideal of the tiled order Λ. Evidently,

$$I=\sum_{i,j}e_{ij}\pi^{\beta_{ij}}\mathcal{O},$$

where e_{ij} are matrix units. Denote by $\mathcal{E}(I)=(\beta_{ij})$ the exponent matrix of the ideal I.

For twosided ideals I and J with exponent matrices $\mathcal{E}(I)=(\beta_{ij})$ and $\mathcal{E}(J)=(\gamma_{ij})$ we have $\mathcal{E}(IJ)=(\delta_{ij})$, where $\delta_{ij}=\min\limits_k(\beta_{ik}+\gamma_{kj})$.

Assume that Λ is reduced and write $\mathcal{E}(\Lambda) = (\alpha_{ij})$. Then the exponent matrix $\mathcal{E}(R) = (\beta_{ij})$ of the Jacobson radical R of Λ can be found as follows: $\beta_{ij} = \alpha_{ij}$ for $i \neq j$ and $\beta_{ii} = 1$ for all i.

Let Q be the quiver of the reduced tiled order Λ and let $[Q(\Lambda)]$ be its adjacency matrix. By [3, Theor. 14.6.2], $[Q(\Lambda)]$ is a (0, 1)-matrix, more precisely, $[Q(\Lambda)] = \mathcal{E}(R^2) - \mathcal{E}(R)$.

For the $n \times n$-exponent matrix $\mathcal{E} = (\alpha_{ij})$ define the following matrices:

$$\mathcal{E}^{(1)} = (\beta_{ij}) = \mathcal{E} + E,$$

where E is the identity matrix.

$$\mathcal{E}^{(2)} = (\gamma_{ij}),\ \gamma_{ij} = \min_k(\beta_{ik} + \beta_{kj}). \tag{3}$$

Evidently, $[Q(\Lambda)] = \mathcal{E}^{(2)} - \mathcal{E}^{(1)}$.

Theorem 4 ([5]). *The matrix $[Q(\Lambda)]$ is the adjacency matrix of a strongly connected simply laced quiver.*

Definition 3. A quiver is called admissible, if it is the quiver of some exponent matrix.

Theorem 5 ([6]). *Let Q be a strongly connected quiver which has a loop at each vertex. Then Q is admissible.*

Theorem 6 ([1, Teor. 5.3]). *For every natural m, ($1 \leqslant m \leqslant n$, $m \neq n-1$), there exists an admissible quiver with n vertices and exactly m loops.*

Theorem 7 ([1]). *Let Q be a strongly connected quiver with n vertices which has exactly $n-1$ loop. Then Q is not admissible.*

Definition 4. Two exponent matrices $\mathcal{E} = (\alpha_{ij})$ and $\Theta = (\theta_{ij})$ are called equivalent if they can be obtained from each other by transformations of the following two types:

(1) subtraction of an integer from the i-th row with simultaneous addition of the same integer to the i-th column;

(2) simultaneous interchanging of two rows and of the equally numbered columns.

Proposition 1 ([1]). *Suppose, that $\mathcal{E}$ and Θ are exponent matrices and Θ can be obtained from $\mathcal{E}$ by transformations of type (1). Then $Q(\mathcal{E}) = Q(\Theta)$.*

For an $n \times n$-matrix A and a permutation σ of $\{1, \ldots, n\}$ denote by $\sigma \circ A$ the matrix, which is obtained from A by simultaneous permutation of rows and columns, defined by σ.

Proposition 2 ([1]). *Let τ be an arbitrary permutation of $\{1,\dots,n\}$. Suppose that $\mathcal{E}$ and Θ are exponent matrices such that Θ can be obtained applying τ to the rows and columns of $\mathcal{E}$. Then $[Q(\Theta)]=\tau\circ[Q(\mathcal{E})]$.*

Since any permutation is a product of transpositions, the above fact explains how does an adjacency matrix changes under transformations of the second type.

1. Generating exponent matrices in the study of quivers of exponent matrices

A non-negative exponent matrix is called **generating**, if it can not be represented as a sum of non-negative non-zero exponent matrices. Denote by $\mathcal{G}_n$ the set of all generating $n\times n$ exponent matrices. By [9] cardinality of $\mathcal{G}_n$ is finite.

For a quiver Q denote by Q^* the quiver, which is obtained from Q by deleating all loops.

Lemma 1. *Let $A_1,\dots,A_s$ be exponent matrices and Q be the quiver of $A=\sum\limits_{t=1}^{s}\alpha_tA_t$, where all α_s are positive integers, such that A is reduced.*

1) *Let $\widetilde{\alpha}_t=\min\{2,\alpha_s\}$ for all s. Then Q is also the quiver of $\widetilde{A}=\sum\limits_{t=1}^{s}\widetilde{\alpha}_tA_t$.*
2) *Let $\alpha_t^*=\min\{1,\alpha_s\}$ for all s. Then Q^* coincides with $(Q(A^*))^*$, where $A^*=\sum\limits_{t=1}^{s}\alpha_t^*A_t$.*

Proof. Write $A=(\alpha_{pq})$, $A+E=B=(\beta_{pq})$ and $C=(\gamma_{pq})$, where

$$\gamma_{ij}=\min_k\{\beta_{ik}+\beta_{kj}\}-\beta_{ij}.$$

Write also $\beta_{ijk}=\beta_{ik}+\beta_{kj}-\beta_{ij}$ and $\alpha_{ijk}=\alpha_{ik}+\alpha_{kj}-\alpha_{ij}$.

Notice, that if $k=i$, or $k=j$, then $\beta_{ijk}=1$. Indeed, if $i\neq j$, then $\beta_{iji}=(\alpha_{ii}+1)+\alpha_{ij}-\alpha_{ij}=1$ and $\beta_{ijj}=(\alpha_{ij}+1)+\alpha_{jj}-\alpha_{ij}=1$. Also if $i=j$, then $\beta_{iii}=(\alpha_{ii}+1)+(\alpha_{ii}+1)-(\alpha_{ii}+1)=1$.

We conclude that $\gamma_{ij}=\min\{1,\min\limits_{k\notin\{i,j\}}\beta_{ijk}\}$.

Notice, that C is the adjacency matrix of the quiver of A. We do next some transformations of formulas for entries of C, which will prove the Lemma.

For $i \neq j$ we can simplify γ_{ij} as follows

$$\begin{aligned}\gamma_{ij} &= \min\{1, \min_{k\notin\{i,j\}}(\beta_{ik}+\beta_{kj}-\beta_{ij})\}\\ &= \min\{1, \min_{k\notin\{i,j\}}(\alpha_{ik}+\alpha_{kj}-\alpha_{ij})\}\\ &= \min\{1, \min_{k\notin\{i,j\}}\alpha_{ijk}\}\\ &= \min\{1, \min_{k\notin\{i,j\}}\sum_{t=1}^{s}\alpha_t\alpha_{ijk}^t\}.\end{aligned}$$

Since $\alpha_{ijk}^t \geqslant 0$ for all i, j, k, t, the conditions $\sum\limits_{t=1}^{s}\alpha_t\alpha_{ijk}^t = 0$ and $\sum\limits_{t=1}^{s}\alpha_t^*\alpha_{ijk}^t = 0$ are equivalent. This proves the first part of Lemma.

For $i = j$ the formulas for γ_{ij} can be transformed as follows

$$\begin{aligned}\gamma_{ij} &= \min\{1, \min_{k\notin\{i,j\}}(\beta_{ik}+\beta_{ki}-1)\}\\ &= \min\{1, \min_{k\notin\{i,j\}}(\alpha_{ijk}-1)\}\\ &= \min\{1, \min_{k\notin\{i,j\}}\sum_{t=1}^{s}\alpha_t\alpha_{ijk}^t-1\}.\end{aligned}$$

Since A is reduced, then $\sum\limits_{t=1}^{s}\alpha_t\alpha_{ijk}^t \geqslant 1$. Nevertheless, the conditions $\sum\limits_{t=1}^{s}\alpha_t\alpha_{ijk}^t = 1$ and $\sum\limits_{t=1}^{s}\widetilde{\alpha}_t\alpha_{ijk}^t = 1$ are equivalent. This proves the second part of Lemma. □

The following two theorems follow from Lemma 1.

Theorem 8. *Let Q be an admissible quiver with n vertices and let $\mathcal{G}_n = \{A_1, \ldots, A_s\}$. Then there exist $\alpha_i \in \{0, 1, 2\}$, $1 \leqslant i \leqslant s$, such that Q is the quiver of $\sum\limits_{i=1}^{s}\alpha_i A_i$.*

Theorem 9. *Let Q be an admissible quiver with n vertices, which has no loops and let $\mathcal{G}_n = \{A_1, \ldots, A_s\}$. Then there exist $\alpha_i \in \{0, 1\}$, $1 \leqslant i \leqslant s$, such that Q is the quiver of $\sum\limits_{i=1}^{s}\alpha_i A_i$.*

2. Quivers of reduced exponent 3×3-matrices

The main result of this article is the following theorem.

Theorem 10. *The following* 10 *matrices are the adjacency matrices of the quivers of all* 3×3*-reduced exponent matrices, up to isomorphism of quivers:*

1) *The quivers with a loop at each vertex*

$$N_1 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, N_2 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}, N_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

$$N_4 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, N_5 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

2) *The quivers without loops:*

$$K_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, K_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

3) *The quivers with exactly one loop:*

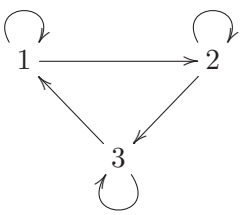
$$T_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}, T_2 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}, T_3 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

The Theorem will be proved in Section 3.

Remark 1. Notice, that the quivers with adjacency matrices $N_1, \ldots, N_5$ form the complete list of the strongly connected simply laced quivers on 3 vertices up to isomorphism.

Proof of the Remark. Consider an arbitrary strongly connected quiver Q with 3 vertices which has a loop at each vertex.

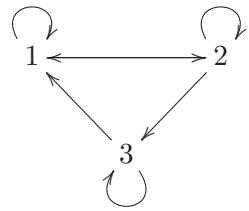
Assume, that Q has exactly 3 arrows, which are not loops. In this case Q is isomorphic to



(4)

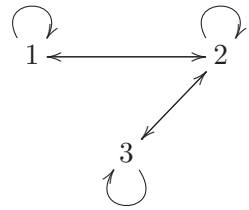
and $[Q] = N_5$.

If Q has more then 3 arrows, which are not loops, then there are vertices i and j, with arrows $i \to j$, $j \to i$. Without loss of generality, we may assume that $i = 1$ and $j = 2$. Assume, that Q has exactly 4 arrows, which are not loops. In this case Q is isomorphic to either



(5)

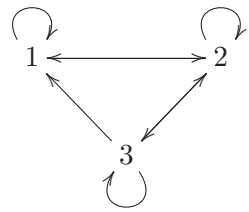
or



(6)

The quiver in (5) is isomorphic to the quiver with adjacency matrix N_2 and the quiver in (6) is isomorphic to one with adjacency matrix N_4.

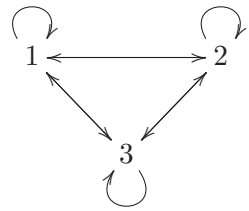
If Q has exactly 5 arrows, which are not loops, then there is a vertex, say 2, such that there is an arrow from 2 to all other vertices and there are arrows from all other vertices to 2. Without loss of generality we may assume that the last, 5-th arrow, goes from 3 to 1, whence, the quiver is as follows



(7)

This quiver is isomorphic to one with adjacency matrix N_3.

The last case is the complete simply laced quiver



(8)

Its adjacency matrix is N_1. □

Corollary 1. *There are* 5 *pairwise non-isomorphic strongly connected simply laced quivers with 3 vertices, which have no loops.*

Proof. There is a natural one to one correspondence between quivers with no loops and ones which have a loop at each the vertex. Now, the corollary follows from Remark 1. □

We shall say that two vertices i and j of a quiver Q are *similar* if the renumeration $i \to j$, $j \to i$ of the vertices of Q gives an isomorphic quiver.

Lemma 2. *There are exactly 10 pairwise non-isomorphic strongly connected simply laced quivers with 3 vertices, which have exactly* 1 *loop.*

Proof. Let Q be a quiver such as in the statement of Lemma. If we add two new loops, it will become isomorphic to one of (4),..., (8), mentioned in the proof of Remark 1.

If $\widetilde{Q}$ is of the form (4), then there is a unique possibility for Q (which we denote by Q_1), because all vertices of Q are pairwise similar.

If $\widetilde{Q}$ is of the form (5), then the three vertices are pairwise non-similar and there are three possibilities for Q (denote them Q_2, Q_3 and Q_4).

If $\widetilde{Q}$ is of the form (6), then vertices 1 and 3 are similar and 2 is not similar to them. Whence, there are two possibilities Q_5 and Q_6 for Q.

If $\widetilde{Q}$ is of the form (7), then all vertices of the quiver are pairwise non-similar and there are three possibilities Q_7, Q_8 and Q_9 for Q.

If $\widetilde{Q}$ is of the form (8), then all vertices of Q are pairwise similar, whence, there is a unique possibility, which we denote by Q_{10}. □

Corollary 2. *There are exactly 10 pairwise non-isomorphic strongly connected simply laced quivers with 3 vertices, which have exactly* 2 *loops.*

Corollary 3. *There are exactly 30 pairwise non-isomorphic strongly connected simply laced quivers with 3 vertices.*

3. Calculation of the admissible quivers with 3 vertices

From Proposition 1 it follows that for any admissible quiver there is a reduced exponent matrix, whose first row is zero. It immediately follows from the definition of the exponent matrix, that if one of its rows is zero, then all other entries are non-negative. The additive semigroup of the non-negative 3×3-exponent matrices was studied at [2]. This semigroup

is finitely generated and the unique set of its generators is as follows.

$$A_1=\begin{pmatrix}0&0&0\\1&0&0\\1&0&0\end{pmatrix};\quad A_2=\begin{pmatrix}0&1&0\\0&0&0\\0&1&0\end{pmatrix};\quad A_3=\begin{pmatrix}0&0&1\\0&0&1\\0&0&0\end{pmatrix};$$
$$A_4=\begin{pmatrix}0&1&1\\0&0&0\\0&0&0\end{pmatrix};\quad A_5=\begin{pmatrix}0&0&0\\1&0&1\\0&0&0\end{pmatrix};\quad A_6=\begin{pmatrix}0&0&0\\0&0&0\\1&1&0\end{pmatrix};$$
$$A_7=\begin{pmatrix}0&1&1\\0&0&0\\0&1&0\end{pmatrix};\quad A_8=\begin{pmatrix}0&1&1\\0&0&1\\0&0&0\end{pmatrix};\quad A_9=\begin{pmatrix}0&0&0\\1&0&1\\1&0&0\end{pmatrix};$$
$$A_{10}=\begin{pmatrix}0&0&1\\1&0&1\\0&0&0\end{pmatrix};\quad A_{11}=\begin{pmatrix}0&1&0\\0&0&0\\1&1&0\end{pmatrix};\quad A_{12}=\begin{pmatrix}0&0&0\\1&0&0\\1&1&0\end{pmatrix}.$$

Whence, the semigroup of the non-negative 3×3 exponent matrices, whose first line is zero, is generated by $A_1, A_5, A_6, A_9, A_{12}$.

Notice, that the matrices A_1, A_5 and A_6 are not reduced, but A_9 and A_{12} are. It follows that the semigroup of reduced exponent matrices with first zero line is not finitely generated. For example, for any $x_1>0$ the exponent matrix $x_1A_1+A_5$ is reduced, but it can not be represented as a sum of other non-negative reduced exponent matrices.

Nevertheless, notice, that if at least two of non-negative numbers from $\{x_1,\, x_5,\, x_6,\, x_9,\, x_{12}\}$ are greater then 0, then

$$\mathcal{E}=x_1A_1+x_5A_5+x_6A_6+x_9A_9+x_{12}A_{12} \tag{9}$$

is a reduced exponent matrix.

The matrix $\mathcal{E}$, defined by (9), can be written as

$$\mathcal{E}=\begin{pmatrix}0&0&0\\x_1+x_5+x_9+x_{12}&0&x_5+x_9\\x_1+x_6+x_9+x_{12}&x_6+x_{12}&0\end{pmatrix}.$$

We are going to find all possible quivers $Q(\mathcal{E})$ depending on the values of x_i. We will find the quivers up to their equivalence classes, because if some quiver is admissible, the all those, which are equivalent to it, are also admissible.

Notice, that the matrices A_9 and A_{12} are type (2) equivalent. By Proposition 2, without loss of generality we may assume, that $x_9\leqslant x_{12}$.

This means, that $x_{12} = x_9 + x_{13}$ for some non-negative x_{13}, whence the formulas for the entries of $\mathcal{E}^{(2)} = (\gamma_{ij})$ will be as follows:

$$\begin{aligned}
\gamma_{11} &= \min\{2,\ x_1 + x_5 + 2x_9 + x_{13},\ x_1 + x_6 + 2x_9 + x_{13}\};\\
\gamma_{12} &= \min\{1,\ x_6 + x_9 + x_{13}\};\\
\gamma_{13} &= \min\{1,\ x_5 + x_9\};\\
\gamma_{21} &= \alpha_{21} + \min\{1,\ x_6 + x_9\};\\
\gamma_{22} &= \min\{2,\ x_1 + x_5 + 2x_9 + x_{13},\ x_5 + x_6 + 2x_9 + x_{13}\};\\
\gamma_{23} &= \alpha_{23} + \min\{x_1 + x_9 + x_{13},\ 1\};\\
\gamma_{31} &= \alpha_{31} + \min\{1,\ x_5 + x_9 + x_{13}\};\\
\gamma_{32} &= \alpha_{32} + \min\{1,\ x_1 + x_9\};\\
\gamma_{33} &= \min\{x_1 + x_6 + 2x_9 + x_{13},\ x_5 + x_6 + 2x_9 + x_{13},\ 2\};
\end{aligned}$$

For each of the variables, which appear in the formulas for the entries of $\mathcal{E}^{(2)}$, consider cases, depending on whether it is zero, or is grater then zero. After we have made assumption about some variable, we will go on with assumptions about others. Also, if necessary, we will consider for a variable, which is earlier assumed to be positive, cases of its being equal to 1 or grater then 1.

We shall consider the following cases.

Case 1: $x_1 = 0$.

Case 1.1: $x_1 = 0$ and $x_5 = 0$.

Case 1.1.1: $x_1 = 0$, $x_5 = 0$ and $x_6 = 0$.

Case 1.1.1.1: $x_1 = 0$, $x_5 = 0$, $x_6 = 0$, and $x_9 = 0$. This leads to $x_{13} > 0$, otherwise A is the zero matrix.

Case 1.1.1.1.1: $x_1 = 0$, $x_5 = 0$, $x_6 = 0$, and $x_9 = 0$, $x_{13} = 1$. These assumptions lead to the quiver with adjacency matrix $[Q] = K_2$.

Case 1.1.1.1.2: $x_1 = 0$, $x_5 = 0$, $x_6 = 0$, and $x_9 = 0$, $x_{13} > 1$. In this case we have, that $[Q] = N_5$.

Similarly the rest of the cases are as follows.

Case 1.1.1.2: $x_1 = 0$, $x_5 = 0$, $x_6 = 0$, and $x_9 > 0$. $[Q] = N_1$.

Case 1.1.2: $x_1 = 0$, $x_5 = 0$ and $x_6 > 0$.

Case 1.1.2.1: $x_1 = 0$, $x_5 = 0$, $x_6 > 0$ and $x_9 = 0$. Notice, that in this case $x_{13} > 0$, otherwise, matrix A is not reduced.

Case 1.1.2.1.1: $x_1 = x_5x_9 = 0$, $x_6 > 0$ and $x_{13} = 1$. $[Q] = T_1$.

Case 1.1.2.1.2: $x_1 = x_5 = x_9 = 0$, $x_6 > 0$ and $x_{13} > 1$. $[Q] = N_2$.

Case 1.1.2.2: $x_1 = 0$, $x_5 = 0$, $x_6 > 0$ and $x_9 > 0$. $[Q] = N_1$.

Case 1.2: $x_1 = 0$ and $x_5 > 0$.

Case 1.2.1: $x_1 = 0$, $x_5 > 0$, and $x_6 = 0$.
Case 1.2.1.1: $x_1 = 0$, $x_5 > 0$, $x_6 = 0$, and $x_9 = 0$. In this case, $x_{13} > 0$, because otherwise A will not be reduced.
Case 1.2.1.1.1: $x_1 = x_6 = x_9 = 0$, $x_5 > 0$ and $x_{13} = 1$. $[Q] = (123) \circ T_1$.
Case 1.2.1.1.2: $x_1 = x_6 = x_9 = 0$, $x_5 > 0$ and $x_{13} > 1$. $[Q] = (123) \circ N_2$.
Case 1.2.1.2: $x_1 = 0$, $x_5 > 0$, $x_6 = 0$, and $x_9 > 0$. $[Q] = N_1$.
Case 1.2.2: $x_1 = 0$, $x_5 > 0$, and $x_6 > 0$.
Case 1.2.2.1: $x_1 = 0$, $x_5 > 0$, $x_6 > 0$ and $x_9 = 0$.
Case 1.2.2.1.1: $x_1 = 0$, $x_5 = 1$, $x_6 > 0$, $x_9 = 0$.
Case 1.2.2.1.1.1: $x_1 = 0$, $x_5 = 1$, $x_6 > 0$, $x_9 = 0$, and $x_{13} = 0$.
Case 1.2.2.1.1.1.1: $x_1 = x_9 = x_{13} = 0$ and $x_5 = x_6 = 1$.$[Q] = (12) \circ K_1$.
Case 1.2.2.1.1.1.2: $x_1 = x_9 = x_{13} = 0$, $x_5 = 1$ and $x_6 > 1$. $[Q] = T_2$.
Case 1.2.2.1.1.2: $x_1 = x_9 = 0$, $x_5 = 1$, $x_6 > 0$ and $x_{13} > 0$. $[Q] = N_3$.
Case 1.2.2.1.2: $x_1 = 0$, $x_5 > 1$, $x_6 > 0$ and $x_9 = 0$.
Case 1.2.2.1.2.1: $x_1 = 0$, $x_5 > 1$, $x_6 = 1$ and $x_9 = 0$. $[Q] = T_3$.
Case 1.2.2.1.2.2: $x_1 = 0$, $x_5 > 1$, $x_6 > 1$ and $x_9 = 0$. $[Q] = N_3$.
Case 1.2.2.2: $x_1 = 0$, $x_5 > 0$, $x_6 > 0$ and $x_9 > 0$. $[Q] = N_1$.
Case 2: $x_1 > 0$.
Case 2.1: $x_1 > 0$ and $x_5 = 0$.
Case 2.1.1: $x_1 > 0$, $x_5 = 0$, and $x_6 = 0$.
Case 2.1.1.1: $x_1 > 0$, $x_5 = 0$, $x_6 = 0$, and $x_9 = 0$. Notice, that in this case, $x_{13} > 0$, because otherwise A will not be reduced.
Case 2.1.1.1.1: $x_1 > 0$, $x_5 = x_6 = x_9 = 0$ and $x_{13} = 1$. $[Q] = (123) \circ T_3$.
Case 2.1.1.1.2: $x_1 > 0$, $x_5 = 0$, $x_6 = 0$, $x_9 = 0$ and $x_{13} > 1$.
$[Q] = (123) \circ N_3$.
Case 2.1.1.2: $x_1 > 0$, $x_5 = 0$, $x_6 = 0$, and $x_9 > 0$. $[Q] = N_1$.
Case 2.1.2: $x_1 > 0$, $x_5 = 0$, and $x_6 > 0$.
Case 2.1.2.1: $x_1 > 0$, $x_5 = 0$, $x_6 > 0$ and $x_9 = 0$.
Case 2.1.2.1.1: $x_1 = 1$, $x_5 = 0$, $x_6 > 0$ and $x_9 = 0$.
Case 2.1.2.1.1.1: $x_1 = 1$, $x_5 = 0$, $x_6 > 0$, $x_9 = 0$ and $x_{13} = 0$.
Case 2.1.2.1.1.1.1: $x_1 = x_6 = 1$ and $x_5 = x_9 = x_{13} = 0$. $[Q] = K_1$.
Case 2.1.2.1.1.1.2: $x_1 = 1$, $x_5 = x_9 = x_{13} = 0$ and $x_6 > 1$. $[Q] = (12) \circ T_2$.
Case 2.1.2.1.1.2: $x_1 = 1$, $x_5 = x_9 = 0$, $x_6 > 0$ and $x_{13} > 0$.
$[Q] = (132) \circ N_3$.
Case 2.1.2.1.2: $x_1 > 1$, $x_5 = 0$, $x_6 > 0$ and $x_9 = 0$.
Case 2.1.2.1.2.1: $x_1 > 1$, $x_5 = 0$, $x_6 > 0$, $x_9 = 0$ and $x_{13} = 0$.
Case 2.1.2.1.2.1.1: $x_1 > 1$, $x_5 = x_9 = x_{13} = 0$ and $x_6 = 1$.
$[Q] = (321) \circ T_2$.
Case 2.1.2.1.2.1.2: $x_1 > 1$, $x_5 = x_9 = x_{13} = 0$ and $x_6 > 1$. $[Q] = N_4$.

Case 2.1.2.1.2.2: $x_1 > 1$, $x_5 = x_9 = 0$, $x_6 > 0$, and $x_{13} > 0$.
$[Q] = (132) \circ N_3$.
Case 2.1.2.2: $x_1 > 0$, $x_5 = 0$, $x_6 > 0$ and $x_9 > 0$. $[Q] = N_1$.
Case 2.2: $x_1 > 0$ and $x_5 > 0$.
Case 2.2.1: $x_1 > 0$, $x_5 > 0$ and $x_6 = 0$.
Case 2.2.1.1: $x_1 > 0$, $x_5 > 0$, $x_6 = 0$ and $x_9 = 0$.
Case 2.2.1.1.1: $x_1 = 1$, $x_5 > 0$, $x_6 = 0$ and $x_9 = 0$.
Case 2.2.1.1.1.1: $x_1 = 1$, $x_5 > 0$, $x_6 = 0$, $x_9 = 0$ and $x_{13} = 0$.
Case 2.2.1.1.1.1.1: $x_1 = x_5 = 1$ and $x_6 = x_9 = x_{13} = 0$. $[Q] = (23) \circ K_1$.
Case 2.2.1.1.1.1.2: $x_1 = 1$, $x_5 > 1$, $x_6 = x_9 = x_{13} = 0$. $[Q] = (132) \circ T_2$.
Case 2.2.1.1.1.2: $x_1 = 1$, $x_5 > 0$, $x_6 = x_9 = 0$ and $x_{13} > 0$.
$[Q] = (132) \circ N_3$.
Case 2.2.1.1.2: $x_1 > 1$, $x_5 > 0$, $x_6 = 0$ and $x_9 = 0$.
Case 2.2.1.1.2.1.1: $x_1 > 1$, $x_5 = 1$, and $x_6 = x_9 = x_{13} = 0$.
$[Q] = (13) \circ T_2$.
Case 2.2.1.1.2.1.2: $x_1 > 1$, $x_5 > 1$ and $x_6 = x_9 = x_{13} = 0$. $[Q] = (23) \circ N_4$.
Case 2.2.1.1.2.2: $x_1 > 1$, $x_5 > 0$, $x_6 = x_9 = 0$ and $x_{13} > 0$.
$[Q] = (231) \circ N_3$.
Case 2.2.1.2: $x_1 > 0$, $x_5 > 0$, $x_6 = 0$ and $x_9 > 0$. $[Q] = N_1$.
Case 2.2.2: $x_1 > 0$, $x_5 > 0$ and $x_6 > 0$. $[Q] = N_1$.

Now for each adjacency matrix $[Q]$ from Theorem 10 point out a case, in which either $[Q]$, or $\sigma \circ [Q]$ (for some permutation σ) appears.

N_1 : 1.1.1.2;	N_2 : 1.1.2.1.2;	N_3 : 1.2.2.1.1.2;
N_4 : 2.1.2.1.2.1.2;	N_5 : 1.1.1.1.2;	
K_1 : 1.2.2.1.1.1.1;	K_2 : 1.1.1.1;	
T_1 : 1.1.2.1.1;	T_2 : 1.2.2.1.1.1.2;	T_3 : 1.2.2.1.2.1.

The above list shows, that all quivers from Theorem 10 are admissible.

We also see, that the matrix N_5 is obtained only in Case 1.1.1.1.2. In this case one of coefficients (precisely, x_{12} of A_{12}) is greater than 1. This gives the following example.

Example 1. For the admissible quiver Q with adjacency matrix

$$N = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix},$$

there is no $\alpha_1, \ldots, \alpha_{12}$ such that $\alpha_i \in \{0, 1\}$ for all i and Q is the quiver of $A = \sum_{i=1}^{12} \alpha_i A_i$, where $\{A_1, \ldots, A_{12}\} = \mathcal{G}_3$.

This example shows, that condition $\alpha_i \in \{0, 1, 2\}$ in Theorem 8 can not be changed to $\alpha_i \in \{0, 1\}$.

References

[1] Dokuchaev, M. A, Kirichenko, V. V., Zelensky, A. V., Zhuravlev V.N., Gorenstein Matrices, *Algebra Discrete Math.*, (2005), no. 1, pp. 8–29.

[2] Dokuchaev M., Kirichenko V., Plakhotnyk M., On one property of minimal exponent matrices, – Prerpint in *Trabalhos do Departamento de Matemática, Univ. de São Paulo, Instituto de Matemática e Ectatśctica,* São Paulo, Brasil. 2013, 9p.

[3] Hazewinkel, M., Gubareni, N., Kirichenko, V. V., *Algebras, rings and modules*, Vol. 1, Mathematics and its Applications, 575. Kluwer Academic Publishers, Dordrecht, 2004.

[4] Kirichenko, V.V., On quasi Frobenius rings and Gorenstein orders, *Trudy Math. Steklov Inst*, vol. **148**, (1978), pp. 168–174.

[5] Kirichenko V.V., Zelensky A.V. and Zhuravlev V.N., Exponent matrices and their quivers, Bul. Acad. de Stiinte a Rep. Moldova, Matematica, N1, (44), 2004, pp. 57-66.

[6] Kirichenko V.V., Zelensky A.V. and Zhuravlev V.N., Exponent matrices and tiled oreders over a discrete valuation rings, Intern. Journ. of Algebra and Computation, Vol. 15, No 5-6, 2005, pp. 997-1012.

[7] Kirichenko V.V., Khibina M.A., Semiperfect semidistributive rings, Infinite Groups and closed algebraic structures. Kyiv., Math. Inst., 1993, p. 457-480.

[8] Tuganbaev A. A., Semidistributive Modules and Rings. Dordrecht: Kluwer, 1998.

[9] Plakhotnyk, V.V., On minimal sets of integral non-negative solutions of a system of linear equations, *Visn., Ser. Fiz.-Mat. Nauky, Kyïv. Univ. Im. Tarasa Shevchenka*, No. 3, (2000), pp. 74–77.

[10] Zavadski, A.G., Kirichenko, V.V., Torsion free modules over prime rings, *Zap. Sci. Semin. LOMI USSR Akad. Sci.*, **57** (1976), pp. 100–116.

CONTACT INFORMATION

M. Dokuchaev Departamento de Matematica University de São Paulo, Caixa Postal 66281, São Paulo, SP 05314-970 – Brazil
E-Mail(*s*): dokucha@gmail.com
Web-page(*s*): www.ime.usp.br

V. Kirichenko, M. Plkakhotnyk Department of Mechanics and Mathematics, Taras Shevchenko National University of Kyiv, Volodymyrska str., 64, 01033 Kyiv, Ukraine.
E-Mail(*s*): vv.kirichenko@gmail.com, makar_plakhotnyk@ukr.net
Web-page(*s*): www.mechmat.univ.kiev.ua

Received by the editors: 18.10.2015
and in final form 18.10.2015.

Algebra and Discrete Mathematics
Volume **20** (2015). Number 1, pp. 69–88

RESEARCH ARTICLE

Finitely presented quadratic algebras of intermediate growth

Dilber Koçak*

Communicated by R. I. Grigorchuk

Abstract. In this article, we give two examples of finitely presented quadratic algebras (algebras presented by quadratic relations) of intermediate growth.

1. Introduction

Let A be a finitely generated algebra over a field k with generating set $S = \{x_1, \ldots, x_m\}$. We denote by A_n the subspace of elements of degree at most n, then $A = \bigcup_{n=0}^{\infty} A_n$. The growth function γ_A^S of A with respect to S is defined as the dimension of the vector space A_n over k,

$$\gamma_A^S(n) = \dim_k(A_n)$$

The function γ_A^S depends on the generating set S. This dependence can be removed by introducing an equivalence relation: Let f and g be eventually monotone increasing and positive valued functions on $\mathbb{N}$. Set $f \preceq g$ if and only if there exist $N > 0$, $C > 0$, such that $f(n) \leqslant g(Cn)$, for $n \geqslant N$, and $f \sim g$ if and only if $f \preceq g$ and $g \preceq f$. The equivalence class of f is called the *growth rate* of f. Simple verification shows that growth functions of an algebra with respect to different generating sets are equivalent. The growth rate is a useful invariant for finitely generated algebraic structures

*The author was partially supported by NSF grant DMS-1207699.

2010 MSC: 16P90, 16S37, 16S30, 17B70.

Key words and phrases: Finitely presented algebras, growth of algebras, quadratic relations.

such as groups, semigroups and algebras. The notion of growth function for groups was introduced by Schwarz [Šva55] and independently by Milnor [Mil68]. The description of groups of polynomial growth was obtained by Gromov in his celebrated work [Gro81]. He proved that every finitely generated group of polynomial growth contains a nilpotent subgroup of finite index.

The study of growth of algebras dates back to the papers by Gelfand and Kirillov, [GK66a, GK66b]. In this paper we are mainly interested in finitely presented algebras whose growth functions behave in intermediate way i.e., they grow faster than any polynomial function but slower than any exponential function. Govorov gave the first examples of finitely generated semigroups and associative algebras of intermediate growth in [Gov72]. Examples of algebras of intermediate growth can also be found in [Ste75, Smi76, She80, Ufn80, KKM83]. The first examples of finitely generated groups of intermediate growth were constructed by Grigorchuk [Gri83, Gri84]. It is still an open problem whether there exists a finitely presented group of intermediate growth. In contrast, there are examples of finitely presented algebras of intermediate growth. The first example is the universal enveloping algebra of a Lie algebra W with basis $\{w_{-1}, w_0, w_1, w_2, \dots\}$ and brackets defined by $[w_i, w_j] = (i-j)w_{i+j}$. W is a subalgebra of the generalized Witt algebra $W_{\mathbb{Z}}$ (see [AS74, p.206] for definitions). It was proven in [Ste75] that W has a finite presentation with two generators and six relations. It is also a graded algebra with generators of degree -1 and 2. Since W has linear growth, its universal enveloping algebra is an example of finitely presented associative algebra of intermediate growth.

The main goal of this paper is to present examples of finitely presented quadratic algebras (algebras defined by quadratic relations) of intermediate growth. The class of quadratic algebras contains a class of finitely presented algebras, called *Koszul algebras*. They play an important role in many studies. In [PP05], it is conjectured that the Hilbert series of a Koszul algebra A is a rational function and in particular, the growth of A is either polynomial or exponential.

In order to construct our first example of a finitely presented quadratic algebra of intermediate growth, we consider the Kac-Moody algebra for the generalized Cartan matrix $A = \left(\begin{smallmatrix} 2 & -2 \\ -2 & 2 \end{smallmatrix}\right)$. This is a graded Lie algebra of polynomial growth whose generators are of degree 1. Next, we consider a suitable subalgebra and its universal enveloping algebra.

Theorem 1. Let U be the associative algebra with generators x, y and relations $x^3y - 3x^2yx + 3xyx^2 - yx^3 = 0$, $y^3x - 3y^2xy + 3yxy^2 - xy^3 = 0$. Then

(i) It is the universal enveloping algebra of a subalgebra of the the Kac-Moody algebra for the generalized Cartan matrix $A = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$.

(ii) U is a graded algebra with generators of degree 1.

(iii) It has intermediate growth of type $e^{\sqrt{n}}$.

(iv) The Veronese subalgebra $V_4(U)$ of U is a quadratic algebra given by 14 generators and 96 quadratic relations and it has the same growth type with U.

The Kac-Moody algebra for the generalized Cartan matrix $A = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$ is the affine Lie algebra $A_1^{(1)}$. (For the definition of Kac-Moody algebras and classification of affine Lie algebras see [Kac85]). It has a subalgebra which is isomorphic to the Lie subalgebra L of $sl_2(\mathbb{C}[t])$ which consists of all matrices with entries on and under the diagonal divisible by t. That is,

$$L = \{a = (a_{ij})_{2\times 2} \mid a_{ij} \in \mathbb{C}[t],\ \mathrm{tr}(a) = 0 \\ \text{and for } (i,j) \neq (1,2),\ t \text{ divides } a_{ij}\}$$

with the usual Lie bracket $[a, b] = ab - ba$. It follows from [Kac85, Theorem 9.11] that L is finitely presented. In this paper we will prove this by using the axioms of Lie bracket without mentioning the theory of Kac-Moody algebras. In Section 2 we show that L is a finitely presented graded Lie algebra whose generators are all of degree 1 and L has linear growth. In Section 3 we explain the relation between the growth of a Lie algebra and its universal enveloping algebra. In Section 4 we consider the Veronese subalgebra of U to obtain a finitely presented quadratic algebra of intermediate growth and in Section 5 we complete the proof of Theorem 1. In Section 6 we give another example of finitely presented associative algebra A of intermediate growth related to the example of the monoid in [Kob95]. A has the following presentation:

$$A = \langle a, b, c \mid b^2a = ab^2,\ b^2c = aca,\ acc = 0, \\ aba = 0,\ abc = 0,\ cba = 0,\ cbc = 0\rangle$$

We show that A has intermediate growth of type $e^{\sqrt{n}}$ and its Veronese subalgebra $V_3(A)$ is an example of finitely presented quadratic algebra of

intermediate growth. In Section 7 , we give an explicit presentation of the Veronese subalgebra $V_4(U)$ of the first construction U as an example of a finitely presented quadratic algebra of intermediate growth.

2. An example of a finitely presented Lie Algebra of linear growth

The following example is a subalgebra of the Kac-Moody Algebra for the generalized Cartan matrix $A = \left(\begin{smallmatrix} 2 & -2 \\ -2 & 2 \end{smallmatrix} \right)$ [Kac85].

Consider the subalgebra L of $Sl_2(\mathbb{C}[t])$ over $\mathbb{C}$ (i.e., matrices of trace 0 with entries in $\mathbb{C}[t]$)) which consists of matrices whose entries on and under the diagonal are divisible by t. That is,

$$L = \{a = (a_{ij})_{2x2} |\ \ a_{ij} \in \mathbb{C}[t],\ \mathrm{tr}(a) = 0 \\ \text{and for } (i,j) \neq (1,2),\ t \text{ divides } a_{ij}\}$$

with the usual Lie bracket $[a, b] = ab - ba$.

Proposition 1. Let L be the Lie algebra described above. Then it has the following properties.

(i) L is finitely presented with generators

$$x := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad y := \begin{pmatrix} 0 & 0 \\ t & 0 \end{pmatrix}$$

and the defining relations $[x, [x, [x, y]]] = 0$ and $[y, [y, [y, x]]] = 0$.

(ii) $L = \bigoplus_{k \geqslant 1} L_k$ is graded and generated by L_1.

(iii) L has linear growth.

Proof. Take

$$x_1 := x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad y_1 := y = \begin{pmatrix} 0 & 0 \\ t & 0 \end{pmatrix}, \quad \text{and let} \quad z_1 := \begin{pmatrix} t & 0 \\ 0 & -t \end{pmatrix}.$$

In fact, define

$$x_i := \begin{pmatrix} 0 & t^{i-1} \\ 0 & 0 \end{pmatrix}, \ y_i := \begin{pmatrix} 0 & 0 \\ t^i & 0 \end{pmatrix}, \text{ and let } z_i := \begin{pmatrix} t^i & 0 \\ 0 & -t^i \end{pmatrix} \quad \text{for } i \geqslant 1.$$

An arbitrary element $w \in L$ is of the form:

$$w = \begin{pmatrix} \sum_{i=1}^{n} m_i t^i & \sum_{i=1}^{n} k_i t^{i-1} \\ \sum_{i=1}^{n} l_i t^i & \sum_{i=1}^{n} -m_i t^i \end{pmatrix} = \sum_{i=1}^{n} k_i x_i + \sum_{i=1}^{n} l_i y_i + \sum_{i=1}^{n} m_i z_i.$$

So, any element of L can be written as a linear combination of $x_i,\ y_i,\ z_i$ for $i \geqslant 1$ and $\{x_i,\ y_i,\ z_i\}_{i=1}^{\infty}$ forms a linearly independent set over $\mathbb{C}$.

Algebra L has the following relations

$$[x_i, y_j] = z_{i+j-1}, \tag{1}$$
$$[x_i, z_j] = -2x_{i+j}, \tag{2}$$
$$[y_i, z_j] = 2y_{i+j}, \tag{3}$$
$$[x_i, x_j] = 0, \tag{4}$$
$$[y_i, y_j] = 0, \tag{5}$$
$$[z_i, z_j] = 0. \tag{6}$$

for $i, j \geqslant 1$. In particular,

$$x_{i+1} = -\frac{1}{2}[x_i, z_1], \qquad y_{i+1} = \frac{1}{2}[y_i, z_1], \qquad z_i = [x_i, y_1].$$

It follows that L is generated by x_1 and y_1. In order to show that all the relations (1)–(6) can be derived from the relations $[x_1, [x_1, [x_1, y_1]]] = 0$ and $[y_1, [y_1, [y_1, x_1]]] = 0$, we apply induction on $i + j = n$. If $i + j = 2$, the relations (1)–(6) hold trivially. If $i + j = 3$,

$$\begin{aligned} [x_1, y_2] &= [x_1, \frac{[y_1, z_1]}{2}] \\ &= -\frac{1}{2}\left([z_1, [x_1, y_1]] + [y_1, [z_1, x_1]]\right) \\ &= [x_2, y_1] \\ &= z_2, \end{aligned}$$

$$\begin{aligned} [x_1, z_2] &= [x_1, [x_2, y_1]] \\ &= -[y_1, [x_1, x_2]] + [x_2, [y_1, x_1]] (\text{since } [x_1, x_2] = 0) \\ &= [x_2, [x_1, y_1]] \\ &= [x_2, z_1] \\ &= -2x_3, \end{aligned}$$

$$\begin{aligned} [y_1, z_2] &= [y_1, [x_1, y_2]] \\ &= -([y_2, [y_1, x_1]] + [x_1, [y_2, y_1]]) \text{ (since } [y_1, y_2] = 0) \\ &= [y_2, z_1] \\ &= 2y_3. \end{aligned}$$

The relations (4)-(5) for $n = 3$ correspond to relations of L_0. Observe the following three equations for $[z_2, z_1]$,

$$\begin{aligned}[z_2, z_1] &= [[x_2, y_1], z_1] \\ &= -([[z_1, x_2], y_1] + [[y_1, z_1], x_2]) \\ &= [[x_2, z_1], y_1] + [x_2, [y_1, z_1]] \\ &= -2[x_3, y_1] + 2[x_2, y_2] \\ &= k,\end{aligned}$$

$$\begin{aligned}[z_2, z_1] &= [[x_1, y_2], z_1] \\ &= -([[z_1, x_1], y_2] + [[y_2, z_1], x_1]) \\ &= [[x_1, z_1], y_2] + [x_1, [y_2, z_1]] \\ &= -2[x_2, y_2] + 2[x_1, y_3] \\ &= l,\end{aligned}$$

$$\begin{aligned}[z_2, z_1] &= [z_2, [x_1, y_1] \\ &= -([y_1, [z_2, x_1]] + [x_1, [y_1, z_2]]) \\ &= 2[x_3, y_1] - 2[x_1, y_3] \\ &= m.\end{aligned}$$

$3 \cdot [z_2, z_1] = k + l + m = 0$. So, (1)–(6) hold for $n = 3$. Now, suppose that (1)–(6) hold for $i + j \leqslant n$ for some $n \geqslant 3$. For $1 \leqslant i \leqslant n - 1$,

$$\begin{aligned}[x_i, y_{j+1}] &= \frac{1}{2}[x_i, [y_j, z_1]] \\ &= -\frac{1}{2}([z_1, [x_i, y_j]] + [y_j, [z_1, x_i]]) \\ &= [x_{i+1}, y_j],\end{aligned}$$

$$\begin{aligned}-2x_{n+1} &= [x_n, z_1] \\ &= -\frac{1}{2}[[x_1, z_{n-1}], z_1] \\ &= \frac{1}{2}([[z_1, x_1], z_{n-1}] + [[z_{n-1}, z_1], x_1]) \\ &= [x_2, z_{n-1}],\end{aligned}$$

and

$$\begin{aligned}[x_i, z_{j+1}] &= [x_i, [x_1, y_{j+1}]] \\ &= -([y_{j+1}, [x_i, x_1]] + [x_1, [y_{j+1}, x_i]]) \\ &= [x_1, z_{i+j}].\end{aligned}$$

Similarly, it can be shown that

$$2y_{n+1} = [y_i, z_{j+1}]$$

for any $i, j \geqslant 1$ such that $i + j = n$. So (1)–(3) hold for $i + j = n + 1$.

$$\begin{aligned}
[x_1, x_n] &= -\frac{1}{2}[x_1, [x_i, z_j]] \\
&= \frac{1}{2}([z_j, [x_1, x_i]] + [x_i, [z_j, x_1]]) \\
&= -\frac{1}{2}[x_i, [x_1, z_j]] \\
&= [x_i, x_j]
\end{aligned}$$

This equality implies $[x_i, x_j] = [x_j, x_i]$. Similarly, one checks that $[y_i, y_j] = [y_j, y_i]$. Hence, (4)–(5) hold for $i + j = n + 1$.

Finally, we need check that (6) holds for $i + j = n + 1$.

$$\begin{aligned}
[z_1, z_n] &= [z_1, [x_n, y_1]] = 2[x_{n+1}, y_1] - 2[x_n, y_2] \\
&= [z_1, [x_{n-1}, y_2]] = 2[x_n, y_2] - 2[x_{n-1}, y_3] \\
&\vdots \\
&= [z_1, [x_1, y_n]] = 2[x_2, y_n] - 2[x_1, y_{n+1}]
\end{aligned}$$

implies that $n \cdot [z_1, z_n] = 2[x_{n+1}, y_1] - 2[x_1, y_{n+1}]$ and,

$$\begin{aligned}
2[x_1, y_{n+1}] = [x_1, [y_1, z_n]] &= -[z_n, [x_1, y_1]] - [y_1, [z_n, x_1]] \\
&= [z_1, z_n] + 2[x_{n+1}, y_1].
\end{aligned}$$

So $[z_1, z_n] = 0$. Now, consider $[z_i, z_j]$ for $i \in \{1, \ldots, n-1\}$,

$$\begin{aligned}
[z_i, z_j] = [z_i, [x_j, y_1]] &= -([y_1, [z_i, x_j]] + [x_j, [y_1, z_i]]) \\
&= 2[x_{i+j}, y_1] - 2[x_j, y_{i+1}],
\end{aligned}$$

and

$$\begin{aligned}
[x_j, y_{i+1}] = \frac{1}{2}[x_j, [y_i, z_1] &= -\frac{1}{2}([z_1, [x_j, y_i]] + [y_i, [z_1, x_j]]) \\
&= -\frac{1}{2}([z_1, z_n] + [y_i, 2x_{j+1}]) \\
&= [x_{j+1}, y_i]
\end{aligned}$$

By applying this i times we get $[x_j, y_{i+1}] = [x_n, y_1]$, so that

$$[z_i, z_j] = 0 \text{ for } i + j = n + 1$$

i.e., (6) holds for $i+j=n+1$. By (1) - (3), the set $\{x_i, y_i, z_i\}_{i=1}^{\infty}$ forms a basis for L as a vector space. It can be observed that $L = \bigoplus_{k\geqslant 1} L_k$ where $L_{2k-1} = \langle x_k\rangle \oplus \langle y_k\rangle$ and $L_k = \langle z_k\rangle$ for $k \geqslant 1$. Since

$$[L_{2k-1}, L_{2m-1}] \subseteq L_{2(k+m-1)}, \quad [L_{2k}, L_{2m}] = 0,$$
$$[L_{2k-1}, L_{2m}] \subseteq L_{2(k+m)-1},$$

L admits an $\mathbb{N}$-gradation given by the sum of occurrences of x and y in each commutator i.e., $L = \bigoplus_{k\geqslant 1} L_k$ is a graded Lie algebra generated by two elements of degree 1 ($\deg(a) = min\{n | a \in \bigoplus_{k=1}^{n} L_k)\}$) and L has linear growth (dim $L_i \in \{1, 2\}$ for $i \geqslant 1$). □

Remark 1. We notice that L also admits a $\mathbb{Z}$-gradation. It is a 3-graded Lie algebra (in the sense of [dO03]) over $\mathbb{C}$ generated by elements x of degree 1 and y of degree -1 .

3. The relation between the growth of a Lie algebra and its universal enveloping algebra

Let L be any Lie algebra over a field k and $U(L)$ be its universal enveloping algebra. For an ordered basis $u_1, u_2, \ldots$ of L, monomials $u_{i_1} \ldots u_{i_r}$ with $i_1 \leqslant i_2 \leqslant \cdots \leqslant i_r$ form a basis for $U(L)$ (Poincaré-Birkhoff-Witt Theorem ([Ber78])). If $L = \bigoplus L_n$ is a graded Lie algebra such that all the components are finite dimensional, then

$$\sum_{n=0}^{\infty} b_n t^n = \prod_{n=1}^{\infty} (1-t^n)^{-a_n} \tag{7}$$

where $a_n := \dim(L_n)$ and b_n:=number of monomials of length n in $U(L)$ ([Smi76]). The proof of the following proposition can be found in various papers ([Ber83], [Pet93], [BG00]).

Proposition 2. If a_n and b_n are related by (7) and $a_n \sim n^d$, then $b_n \sim e^{n^{\frac{d+1}{d+2}}}$.

Corollary 1. If a Lie algebra L grows polynomially then its universal enveloping algebra $U(L)$ has intermediate growth. In particular, if L has linear growth, then $U(L)$ has growth of type $e^{\sqrt{n}}$.

4. Veronese subalgebra of an associative graded algebra

Let $A = k\langle x_1, \ldots, x_m\rangle$ be a free associative algebra over a field k with generating set $\{x_1, \ldots, x_m\}$. Each element u of A can be written uniquely as

$$u = u_0 + u_1 + \cdots + u_l,$$

where $A_0 = k$, $u_i \in A_i$ and A_i is the vector space over k spanned by m^i monomials of length i. Let $R = \{f_1, f_2, \ldots, f_s\}$ be a finite set of non-zero homogeneous polynomials and I be the ideal generated by R. Since I is generated by homogeneous polynomials, the factor algebra $\tilde{A} = A/I$ is graded:

$$\tilde{A} = \tilde{A}_0 \oplus \tilde{A}_1 \oplus \cdots \oplus \tilde{A}_n \oplus \ldots$$

where $\tilde{A}_i = (A_i + I)/I \cong A_i/(A_i \cap I)$. For $d \geqslant 1$, a *Veronese subalgebra* of $\tilde{A}$ is defined as

$$V_d(\tilde{A}) := k \oplus \tilde{A}_d \oplus \tilde{A}_{2d} \oplus \ldots$$

It is straightforward to see that,

$$\textit{growth of } \tilde{A} \sim \textit{growth of } V_d(\tilde{A})$$

Proposition 3. [BF85] For sufficiently large d, $V_d(\tilde{A})$ is quadratic.

Proof. Let $d_1, \ldots, d_s$ be the degrees of $f_1, f_2, \ldots, f_s$ respectively and $d \geqslant max\{d_i,\ 1 \leqslant i \leqslant s\}$. For any two words $v',\ v''$ such that

$$\deg(v') + d_i + \deg(v'') = d$$

consider the element $v' f_i v'' \in A_d$, and for any two words $w',\ w''$ such that

$$\deg(w') + d_i + \deg(w'') = 2d$$

consider the element $w' f_i w'' \in A_{2d}$. Let $R^* = \{v' f_i v'', w' f_i w''\}$ for $i \in \{1, \ldots, s\}$ and a be a homogeneous element from $A^{(n)} \cap I$. Say $a = \sum \alpha v f_i w$, where $\alpha \in k$, v and w are words. If we choose a summand and represent $v = v_1 v_2$, $\deg(v_1)$ is a multiple of d, $0 \leqslant \deg(v_2) < d$. Similarly, $w = w_2 w_1$, $\deg(w_1)$ is a multiple of d, $0 \leqslant \deg(w_2) < d$. Then we will get $\deg(v_2 f_i w_2) = d$ or $2d$. Hence $v_2 f_i w_2 \in R^*$. It shows that $V_d(A) \cap I$ is an ideal generated by the elements of R^* and an element $v' f_i v''$ is a linear combination of free generators of $A^{(n)}$ whereas $w' f_i w''$ is a quadratic element in these generators. So $V_d(\tilde{A}) = V_d(A)/(V_d(A) \cap I)$ is a quadratic algebra. □

5. Proof of Theorem 1

Let $L = \langle x_1, \ldots, x_m \mid f_1 = 0, \ldots, f_r = 0\rangle$ where each of f_i is a linear combination of the commutators (elements of the form $[x_{i_1}, \ldots, x_{i_k}]$ with an arbitrary distribution of parentheses inside). Then the universal enveloping algebra $U(L)$ of L is an associative algebra with the identical set of generators and relations, where the commutators are thought of as in the ordinary associative sense: $[x, y] = xy - yx$ [Bou89, Proposition 2, p.14]. The universal enveloping algebra $U(L)$ of $L = \langle x_1, y_1 \mid [x_1, [x_1, [x_1, y_1]]] = 0,\ [y_1, [y_1, [y_1, x_1]]] = 0\rangle$ has the following presentation:

$$U(L) = \langle x_1, y_1 \mid x_1^3 y_1 - 3x_1^2 y_1 x_1 + 3x_1 y_1 x_1^2 - y_1 x_1^3 = 0, \\ y_1^3 x_1 - 3y_1^2 x_1 y_1 + 3y_1 x_1 y_1^2 - x_1 y_1^3 = 0\rangle.$$

So, the associative algebra U in Theorem 1 is the universal enveloping algebra $U(L)$ of L. By Proposition 2, since L has linear growth, the growth rate of $U(L)$ is intermediate of type $e^{\sqrt{n}}$. In order to obtain a quadratic algebra of intermediate growth we consider a Veronese subalgebra of $V_4(U)$ as explained in the previous section and conclude that for a given finitely presented graded algebra with all generators of degree 1, one can construct a finitely presented graded algebra with all relations of degree 2. $V_4(U)$ is an example of a finitely presented graded algebra with intermediate growth. It has 14 generators and 96 relations. In the next section we compute all these relations.

6. A construction based on Kobayashi's example

In this section we construct another example of a finitely presented associative algebra with quadratic relations whose growth function is intermediate. For this, we consider the following example of a monoid with 0 that appears in the paper of Kobayashi [Kob95].

$$M = \langle a, b, c \mid ba = ab,\ bc = aca,\ acc = 0\rangle$$

where $w(a) = w(c) = 1,\ w(b) = 2$, w is a positive weight function on M. Kobayashi shows that M is a finitely presented monoid with solvable word problem which cannot be presented by a regular complete system. In order to prove that it cannot be presented by a regular complete system, he proves that M has intermediate growth. Now, we consider the semigroup

algebra $k[M]$ over a field k. $k[M]$ has the same presentation and growth function with M. So $k[M]$ is an example of finitely presented associative graded algebra of intermediate growth. But the generators of $k[M]$ have degrees $\deg(a) = \deg(c) = 1$ and $\deg(b) = 2$. To construct a quadratic algebra with these properties, we need to consider an algebra whose generators are all of degree 1. Thus we consider the following monoid:

$$\tilde{M} = \langle a, b, c \mid b^2a = ab^2,\ b^2c = aca,\ acc = 0,\\ aba = 0,\ abc = 0,\ cba = 0,\ cbc = 0\rangle$$

where $w(a) = w(b) = w(c) = 1$.

Now, we have the monoid algebra $A := k[\tilde{M}]$ over a field k:

$$A = \langle a, b, c \mid b^2a = ab^2,\ b^2c = aca,\ acc = 0,\\ aba = 0,\ abc = 0,\ cba = 0,\ cbc = 0\rangle$$

where $\deg(a) = \deg(b) = \deg(c) = 1$. To show that A has intermediate growth, we first find a complete rewriting system for A. Let $\prec$ be the shortlex order on $\langle X\rangle$ based on the order $a \prec b \prec c$ i.e.,

$$w_1 \prec w_2 \text{ implies } |w_1| < |w_2| \text{ or } |w_1| = |w_2| \ \&\ w_1 \prec_{lex} w_2.$$

Then A has the rewriting system R consisting of the following relations

$$b^2a \to ab^2$$
$$b^2c \to aca$$
$$acc \to 0$$
$$aba \to 0$$
$$abc \to 0$$
$$cba \to 0$$
$$cbc \to 0$$

It is easily seen that R is Noetherian. By applying the *Knuth-Bendix algorithm*, we obtain the following complete rewriting system R_∞ equivalent to R:

$$R_\infty = \{b^2a \to ab^2,\ b^2c \to aca,\ aba \to 0,\ abc \to 0,\ cba \to 0,\ cbc \to 0\}\\ \cup \bigcup_{n=1}^{\infty} \{a^nca^{n-1}c \to 0\}.$$

A monomial (word) m is called *irreducible* with respect to the rewriting system R if all the rewriting rules act trivially on m. The set of all irreducible words with respect to R is denoted by $Irr(R)$. Since R_∞ is a complete rewriting system, $Irr(R_\infty)$ is the set of words which do not contain u as a subword for any $u \to v \in R_\infty$. By *Bergman's Diamond Lemma* [Ber78], $Irr(R_\infty)$, forms a basis for A. Words in $Irr(R_\infty)$ are of the following form

$$b^s a^{m_1} c a^{m_2} c \dots a^{m_r} c a^l b^k$$

where $s \in \{0,1\}, l, k \in \mathbb{N}\cup\{0\}$ and $0 \leqslant m_1 \leqslant m_2 \leqslant \dots \leqslant m_r, m_i \in \mathbb{N}\cup\{0\}$ for $i \in \{1, \dots r\}$. So, the number of words in $Irr(R^\infty)$ of length n is equal to

$$\sum_{j=0}^{n}(2j+1)\cdot|\{(m_1,\dots,m_r) \mid 0\leqslant m_1\leqslant\dots\leqslant m_r, m_1+\dots+m_r=n-j-r\}|$$
$$= \sum_{j=0}^{n}(2j+1)\cdot p(n-j)$$

where $p(n)$ is the number of partitions of n. Hence

$$\gamma_A(n) \sim p(n) \sim e^{\sqrt{n}}.$$

A is an example of finitely presented graded algebra with generators of degree 1 and intermediate growth function and its *Veronese subalgebra* $V_3(A)$ can be presented by finitely many quadratic relations (to be precise with 21 generators and 280 relations).

7. Appendix: Presentation of the Veronese subalgebra $V_4(U)$ of U

As we noted in the Section 5, $U(L)$ is an associative algebra with generators x, y and the set of relations

$$R = \{x^3y-3x^2yx+3xyx^2-yx^3 = 0,\ y^3x-3y^2xy+3yxy^2-xy^3 = 0\}.$$

Since R is a set of two homogeneous polynomials, U is a graded algebra. Let $V_4(U)$ be the Veronese subalgebra of U. It was proven in Section 4 that $V_4(U)$ is a graded algebra generated by the set S of monomials of length 4 over $\{x, y\}$ and the set of relations $R^* = \{f_i = 0, vf_iw = 0\}$ where v, w are monomials such that $l(v) + l(w) = 4$ and, $f_1 = x^3y - 3x^2yx + 3xyx^2 - yx^3$, $f_2 = y^3x - 3y^2xy + 3yxy^2 - xy^3$. Basically, R^*

is the set of homogeneous polynomials of degree 4 or 8 generated by $R = \{f_1 = 0, f_2 = 0\}$ in $k[x, y]$. Since there are 48 different pairs (v, w) of monomials, R^* consists of 2 homogeneous polynomials of degree 4:

$$(i)\ yx^3 = x^3y - 3x^2yx + 3xyx^2,\ \ (ii)\ y^3x = xy^3 - 3yxy^2 + 3y^2xy$$

and 96 homogeneous polynomials of degree 8:

1) $xyx^2x^4 = x^4yx^3 - 3x^3yx^4 + 3x^2yxx^4$,
2) $x^3yx^4 = x^4x^2yx - 3x^4xyx^2 + 3x^4yx^3$,
3) $x^2y^2yx^3 = x^3yy^2x^2 - 3x^2yxy^2x^2 + 3x^2y^2xyx^2$,
4) $xyx^2x^3y = x^4yx^2y - 3x^3yx^3y + 3x^2yxx^3y$,
5) $x^3yx^3y = x^4x^2y^2 - 3x^4xyxy + 3x^4yx^2y$,
6) $x^2y^2yx^2y = x^3yy^2xy - 3x^2yxy^2xy + 3x^2y^2xyxy$,
7) $xyx^2x^2yx = x^4yxyx - 3x^3yx^2yx + 3x^2yxx^2yx$,
8) $x^2y^2x^4 = x^2yxx^2yx - 3x^2yxxyx^2 + 3x^2yxyx^3$,
9) $x^2y^2yxyx = x^3yy^3x - 3x^2yxy^3x + 3x^2y^2xy^2x$,
10) $xyx^2x^2y^2 = x^4yxy^2 - 3x^3yx^2y^2 + 3x^2yxx^2y^2$,
11) $x^2y^2x^3y = x^2yxx^2y^2 - 3x^2yxxyxy + 3x^2yxyx^2y$,
12) $x^2y^2yxy^2 = x^3yy^4 - 3x^2yxy^4 + 3x^2y^2xy^3$,
13) $xyx^2xyx^2 = x^4y^2x^2 - 3x^3yxyx^2 + 3x^2yxxyx^2$,
14) $xyxyx^4 = xyx^2x^2yx - 3xyx^2xyx^2 + 3xyx^2yx^3$,
15) $xy^3yx^3 = xyxyy^2x^2 - 3xy^2xy^2x^2 + 3xy^3xyx^2$,
16) $xyx^2xy^2x = x^4y^3x - 3x^3yxy^2x + 3x^2yxxy^2x$,
17) $xy^3x^4 = xy^2xx^2yx - 3xy^2xxyx^2 + 3xy^2xyx^3$,
18) $xy^3yxyx = xyxyy^3x - 3xy^2xy^3x + 3xy^3xy^2x$,
19) $xyx^2xyxy = x^4y^2xy - 3x^3yxyxy + 3x^2yxxyxy$,
20) $xyxyx^3y = xyx^2x^2y^2 - 3xyx^2xyxy + 3xyx^2yx^2y$,
21) $xy^3yx^2y = xyxyy^2xy - 3xy^2xy^2xy + 3xy^3xyxy$,
22) $xyx^2xy^3 = x^4y^4 - 3x^3yxy^3 + 3x^2yxxy^3$,
23) $xy^3x^3y = xy^2xx^2y^2 - 3xy^2xxyxy + 3xy^2xyx^2y$,
24) $xy^3yxy^2 = xyxyy^4 - 3xy^2xy^4 + 3xy^3xy^3$,
25) $y^2x^2x^4 = yx^3yx^3 - 3yx^2yx^4 + 3yxyxx^4$,
26) $yx^2yx^4 = yx^3x^2yx - 3yx^3xyx^2 + 3yx^3yx^3$,
27) $yxy^2yx^3 = yx^2yy^2x^2 - 3yxyxy^2x^2 + 3yxy^2xyx^2$,
28) $x^2y^2x^2yx = yx^3yxyx - 3yx^2yx^2yx + 3yxyxy^2xy$,
29) $yxy^2x^4 = yxyxx^2yx - 3yxyxxyx^2 + 3yxyxyx^3$,
30) $yxy^2yxyx = yx^2yy^3x - 3yxyxy^3x + 3yxy^2xy^2x$,
31) $y^2x^2x^2y^2 = yx^3yxy^2 - 3yx^2yx^2y^2 + 3yxyxx^2y^2$,

32) $yxy^2x^3y = yxyxx^2y^2 - 3yxyxxyxy + 3yxyxyx^2y$,
33) $yxy^2yxy^2 = yx^2yy^4 - 3yxyxy^4 + 3yxy^2xy^3$,
34) $y^2x^2x^3y = yx^3yx^2y - 3yx^2yx^3y + 3yxyxx^3y$,
35) $yx^2yx^3y = yx^3x^2y^2 - 3yx^3xyxy + 3yx^3yx^2y$,
36) $yxy^2yx^2y = yx^2yy^2xy - 3yxyxy^2xy + 3yxy^2xyxy$,
37) $y^2x^2xyx^2 = yx^3y^2x^2 - 3yx^2yxyx^2 + 3yxyxxyx^2$,
38) $y^2xyx^4 = y^2x^2x^2yx - 3y^2x^2xyx^2 + 3y^2x^2yx^3$,
39) $y^4yx^3 = y^2xyy^2x^2 - 3y^3xy^2x^2 + 3y^4xyx^2$,
40) $y^2x^2xyxy = yx^3y^2xy - 3yx^2yxyxy + 3yxyxxyxy$,
41) $y^2xyx^3y = y^2x^2x^2y^2 - 3y^2x^2xyxy + 3y^2x^2yx^2y$,
42) $y^4yx^2y = y^2xyy^2xy - 3y^3xy^2xy + 3y^4xyxy$,
43) $y^2x^2xy^2x = yx^3y^3x - 3yx^2yxy^2x + 3yxyxxy^2x$,
44) $y^4x^4 = y^3xx^2yx - 3y^3xxyx^2 + 3y^3xyx^3$,
45) $y^4yxyx = y^2xyy^3x - 3y^3xy^3x + 3y^4xy^2x$,
46) $y^2x^2xy^3 = yx^3y^4 - 3yx^2yxy^3 + 3yxyxxy^3$,
47) $y^4x^3y = y^3xx^2y^2 - 3y^3xxyxy + 3y^3xyx^2y$,
48) $y^4yxy^2 = y^2xyy^4 - 3y^3xy^4 + 3y^4xy^3$,
49) $x^2yxx^4 = x^4xyx^2 - 3x^4yx^3 + 3x^3yx^4$,
50) $xy^3x^4 = x^2y^2yx^3 - 3xyxyyx^3 + 3xy^2xyx^3$,
51) $x^3yy^2x^2 = x^4y^3x - 3x^3yxy^2x + 3x^3yyxyx$,
52) $x^2yxx^3y = x^4xyxy - 3x^4yx^2y + 3x^3yx^3y$,
53) $xy^3x^3y = x^2y^2yx^2y - 3xyxyyx^2y + 3xy^2xyx^2y$,
54) $x^3yy^2xy = x^4y^4 - 3x^3yxy^3 + 3x^3yyxy^2$,
55) $x^2yxx^2yx = x^4xy^2x - 3x^4yxyx + 3x^3yx^2yx$,
56) $xy^3x^2yx = x^2y^2yxyx - 3xyxyyxyx + 3xy^2xyxyx$,
57) $x^2y^2y^2x^2 = x^2yxy^3x - 3x^2y^2xy^2x + 3x^2y^2yxyx$,
58) $x^2yxx^2y^2 = x^4xy^3 - 3x^4yxy^2 + 3x^3yx^2y^2$,
59) $xy^3x^2y^2 = x^2y^2yxy^2 - 3xyxyyxy^2 + 3xy^2xyxy^2$,
60) $x^2y^2y^2xy = x^2yxy^4 - 3x^2y^2xy^3 + 3x^2y^2yxy^2$,
61) $xy^2xx^4 = xyx^2xyx^2 - 3xyx^2yx^3 + 3xyxyx^4$,
62) $xy^3xyx^2 = x^2y^2y^2x^2 - 3xyxyy^2x^2 + 3xy^2xy^2x^2$,
63) $xyxyy^2x^2 = xyx^2y^3x - 3xyxyxy^2x + 3xyxyyyxyx$,
64) $xy^2xx^2yx = xyx^2xy^2x - 3xyx^2yxyx + 3xyxyx^2yx$,
65) $xy^3xy^2x = x^2y^2y^3x - 3xyxyy^3x + 3xy^2xy^3x$,
66) $xy^3y^2x^2 = xy^2xy^3x - 3xy^3xy^2x + 3xy^3yxyx$,
67) $xy^2xx^3y = xyx^2xyxy - 3xyx^2yx^2y + 3xyxyx^3y$,
68) $xy^3xyxy = x^2y^2y^2xy - 3xyxyy^2xy + 3xy^2xy^2xy$,

69) $xyxyy^2xy = xyx^2y^4 - 3xyxyxy^3 + 3xyxyyxy^2$,
70) $xy^2xx^2y^2 = xyx^2xy^3 - 3xyx^2yxy^2 + 3xyxyx^2y^2$,
71) $xy^3xy^3 = x^2y^2y^4 - 3xyxyy^4 + 3xy^2xy^4$,
72) $xy^3y^2xy = xy^2xy^4 - 3xy^3xy^3 + 3xy^3yxy^2$,
73) $yxyxx^4 = yx^3xyx^2 - 3yx^3yx^3 + 3yx^2yx^4$,
74) $y^4x^4 = yxy^2yx^3 - 3y^2xyyx^3 + 3y^3xyx^3$,
75) $yx^2yy^2x^2 = yx^3y^3x - 3yx^2yxy^2x + 3yx^2yyxyx$,
76) $yxyxx^2yx = yx^3xy^2x - 3yx^3yxyx + 3yx^2yx^2yx$,
77) $y^4x^2yx = yxy^2yxyx - 3y^2xyyxyx + 3y^3xyxyx$,
78) $yxy^2y^2x^2 = yxyxy^3x - 3yxy^2xy^2x + 3yxy^2yxyx$,
79) $yxyxx^2y^2 = yx^3xy^3 - 3yx^3yxy^2 + 3yx^2yx^2y^2$,
80) $y^4x^2y^2 = yxy^2yxy^2 - 3y^2xyyxy^2 + 3y^3xyxy^2$,
81) $yxy^2y^2xy = yxyxy^4 - 3yxy^2xy^3 + 3yxy^2yxy^2$,
82) $yxyxx^3y = yx^3xyxy - 3yx^3yx^2y + 3yx^2yx^3y$,
83) $y^4x^3y = yxy^2yx^2y - 3y^2xyyx^2y + 3y^3xyx^2y$,
84) $yx^2yy^2xy = yx^3y^4 - 3yx^2yxy^3 + 3yx^2yyxy^2$,
85) $y^3xx^4 = y^2x^2xyx^2 - 3y^2x^2yx^3 - 3y^2xyx^4$,
86) $y^4xyx^2 = yxy^2y^2x^2 - 3y^2xyy^2x^2 + 3y^3xy^2x^2$,
87) $y^2xyy^2x^2 = y^2x^2y^3x - 3y^2xyxy^2x + 3y^2xyyxyx$,
88) $y^3xx^3y = y^2x^2xyxy - 3y^2x^2yx^2y + 3y^2xyx^3y$,
89) $y^4xyxy = yxy^2y^2xy - 3y^2xyy^2xy + 3y^3xy^2xy$,
90) $y^2xyy^2xy = y^2x^2y^4 - 3y^2xyxy^3 + 3y^2xyyxy^2$,
91) $y^3xx^2yx = y^2x^2yx^2x - 3y^2x^2yxyx + 3y^2xyx^2yx$,
92) $y^4xy^2x = yxy^2y^3x - 3y^2xyy^3x + 3y^3xy^3x$,
93) $y^4y^2x^2 = y^3xy^3x - 3y^4xy^2x + 3y^4yxyx$,
94) $y^3xx^2y^2 = y^2x^2xy^3 - 3y^2x^2yxy^2 + 3y^2xyx^2y^2$,
95) $y^4xy^3 = yxy^2y^4 - 3y^2xyy^4 - 3y^2xyy^4 + 3y^3xy^4$,
96) $y^4y^2xy = y^3xy^4 - 3y^4xy^3 + 3y^4yxy^2$.

We can rename the generators as follows:

$$\begin{array}{llll} y^4 = Y_1, & y^3x = Y_2, & y^2xy = Y_3, & y^2x^2 = Y_4, \\ yxy^2 = Y_5, & yxyx = Y_6, & yx^2y = Y_7, & yx^3 = Y_8, \\ xy^3 = X_1, & xy^2x = X_2, & xyxy = X_3, & xyx^2 = X_4, \\ x^2y^2 = X_5, & x^2yx = X_6, & x^3y = X_7, & x^4 = X_8. \end{array}$$

So the relations will be

$$(i)\ Y_8 = X_7 - 3X_6 + 3X_4, \quad (ii)\ Y_2 = X_1 - 3Y_5 + 3Y_3$$

1) $X_4X_8 = X_8Y_8 - 3X_7X_8 + 3X_6X_8$,
2) $X_7X_8 = X_8X_6 - 3X_8X_4 + 3X_8Y_8$,
3) $X_5Y_8 = X_7Y_4 - 3X_6Y_4 + 3X_5X_4$,
4) $X_4X_7 = X_8Y_7 - 3X_7X_7 + 3X_6X_7$,
5) $X_7X_7 = X_8X_5 - 3X_8X_3 + 3X_8Y_7$,
6) $X_5Y_7 = X_7Y_3 - 3X_6Y_3 + 3X_5X_3$,
7) $X_4X_6 = X_8Y_6 - 3X_7X_6 + 3X_6X_6$,
8) $X_5X_8 = X_6X_6 - 3X_6X_4 + 3X_6Y_8$,
9) $X_5Y_6 = X_7Y_2 - 3X_6Y_2 + 3X_5X_2$,
10) $X_4X_5 = X_8Y_5 - 3X_7X_5 + 3X_6X_5$,
11) $X_5X_7 = X_6X_5 - 3X_6X_3 + 3X_6Y_7$,
12) $X_5Y_5 = X_7Y_1 - 3X_6Y_1 + 3X_5X_1$,
13) $X_4X_4 = X_8Y_4 - 3X_7X_4 + 3X_6X_4$,
14) $X_3X_8 = X_4X_6 - 3X_4X_4 + 3X_4Y_8$,
15) $X_1Y_8 = X_3Y_4 - 3X_2Y_4 + 3X_1X_4$,
16) $X_4X_2 = X_8Y_2 - 3X_7X_2 + 3X_6X_2$,
17) $X_1X_8 = X_2X_6 - 3X_2X_4 + 3X_2Y_8$,
18) $X_1Y_6 = X_3Y_2 - 3X_2Y_2 + 3X_1X_2$,
19) $X_4X_3 = X_8Y_3 - 3X_7X_3 + 3X_6X_3$,
20) $X_3X_7 = X_4X_5 - 3X_4X_3 + 3X_4Y_7$,
21) $X_1Y_7 = X_3Y_3 - 3X_2Y_3 + 3X_1X_3$,
22) $X_4X_1 = X_8Y_1 - 3X_7X_1 + 3X_6X_1$,
23) $X_1X_7 = X_2X_5 - 3X_2X_3 + 3X_2Y_7$,
24) $X_1Y_5 = X_3Y_1 - 3X_2Y_1 + 3X_1X_1$,
25) $Y_4X_8 = Y_8Y_8 - 3Y_7X_8 + 3Y_6X_8$,
26) $Y_7X_8 = Y_8X_6 - 3Y_8X_4 + 3Y_8Y_8$,
27) $Y_5Y_8 = Y_7Y_4 - 3Y_6Y_4 + 3Y_5X_4$,
28) $Y_4X_6 = Y_8Y_6 - 3Y_7X_6 + 3Y_6X_6$,
29) $Y_5X_8 = Y_6X_6 - 3Y_6X_4 + 3Y_6Y_8$,
30) $Y_5Y_6 = Y_7Y_2 - 3Y_6Y_2 + 3Y_5X_2$,
31) $Y_4X_5 = Y_8Y_5 - 3Y_7X_5 + 3Y_6X_5$,
32) $Y_5X_7 = Y_6X_5 - 3Y_6X_3 + 3Y_6Y_7$,
33) $Y_5Y_5 = Y_7Y_1 - 3Y_6Y_1 + 3Y_5X_1$,
34) $Y_4X_7 = Y_8Y_7 - 3Y_7X_7 + 3Y_6X_7$,
35) $Y_7X_7 = Y_8X_5 - 3Y_8X_3 + 3Y_8Y_7$,
36) $Y_5Y_7 = Y_7Y_3 - 3Y_6Y_3 + 3Y_5X_3$,
37) $Y_4X_4 = Y_8Y_4 - 3Y_7X_4 + 3Y_6X_4$,

38) $Y_3X_8 = Y_4X_6 - 3Y_4X_4 + 3Y_4Y_8$,
39) $Y_1Y_8 = Y_3Y_4 - 3Y_2Y_4 + 3Y_1X_4$,
40) $Y_4X_3 = Y_8Y_3 - 3Y_7X_3 + 3Y_6X_3$,
41) $Y_3X_7 = Y_4X_5 - 3Y_4X_3 + 3Y_4Y_7$,
42) $Y_1Y_7 = Y_3Y_3 - 3Y_2Y_3 + 3Y_1X_3$,
43) $Y_4X_2 = Y_8Y_2 - 3Y_7X_2 + 3Y_6X_2$,
44) $Y_1X_8 = Y_2X_6 - 3Y_2X_4 + 3Y_2Y_8$,
45) $Y_1Y_6 = Y_3Y_2 - 3Y_2Y_2 + 3Y_1X_2$,
46) $Y_4X_1 = Y_8Y_1 - 3Y_7X_1 + 3Y_6X_1$,
47) $Y_1X_7 = Y_2X_5 - 3Y_2X_3 + 3Y_2Y_7$,
48) $Y_1Y_5 = Y_3Y_1 - 3Y_2Y_1 + 3Y_1X_1$,
49) $X_6X_8 = X_8X_4 - 3X_8Y_8 + 3X_7X_8$,
50) $X_1X_8 = X_5Y_8 - 3X_3Y_8 + 3X_2Y_8$,
51) $X_7Y_4 = X_8Y_2 - 3X_7X_2 + 3X_7Y_6$,
52) $X_6X_7 = X_8X_3 - 3X_8Y_7 + 3X_7X_7$,
53) $X_1X_7 = X_5Y_7 - 3X_3Y_7 + 3X_2Y_7$,
54) $X_7Y_3 = X_8Y_1 - 3X_7X_1 + 3X_7Y_5$,
55) $X_6X_6 = X_8X_2 - 3X_8Y_6 + 3X_7X_6$,
56) $X_1X_6 = X_5Y_6 - 3X_3Y_6 + 3X_2Y_6$,
57) $X_5Y_4 = X_6Y_2 - 3X_5X_2 + 3X_5Y_6$,
58) $X_6X_5 = X_8X_1 - 3X_8Y_5 + 3X_7X_5$,
59) $X_1X_5 = X_5Y_5 - 3X_3Y_5 + 3X_2Y_5$,
60) $X_5Y_3 = X_6Y_1 - 3X_5X_1 + 3X_5Y_5$,
61) $X_2X_8 = X_4X_4 - 3X_4Y_8 + 3X_3X_8$,
62) $X_1X_4 = X_5Y_4 - 3X_3Y_4 + 3X_2Y_4$,
63) $X_3Y_4 = X_4Y_2 - 3X_3X_2 + 3X_3Y_6$,
64) $X_2X_6 = X_4X_2 - 3X_4Y_6 + 3X_3X_6$,
65) $X_1X_2 = X_5Y_2 - 3X_3Y_2 + 3X_2Y_2$,
66) $X_1Y_4 = X_2Y_2 - 3X_1X_2 + 3X_1Y_6$,
67) $X_2X_7 = X_4X_3 - 3X_4Y_7 + 3X_3X_7$,
68) $X_1X_3 = X_5Y_3 - 3X_3Y_3 + 3X_2Y_2$,
69) $X_3Y_3 = X_4Y_1 - 3X_3X_1 + 3X_3Y_5$,
70) $X_2X_5 = X_4X_1 - 3X_4Y_5 + 3X_3X_5$,
71) $X_1X_1 = X_5Y_1 - 3X_3Y_1 + 3X_2Y_1$,
72) $X_1Y_3 = X_2Y_1 - 3X_1X_1 + 3X_1Y_5$,
73) $Y_6X_8 = Y_8X_4 - 3Y_8Y_8 + 3Y_7X_8$,
74) $Y_1X_8 = Y_5Y_8 - 3Y_3Y_8 + 3Y_2Y_8$,

75) $Y_7Y_4 = Y_8Y_2 - 3Y_7X_2 + 3Y_7Y_6$,
76) $Y_6X_6 = Y_8X_2 - 3Y_8Y_6 + 3Y_7X_6$,
77) $Y_1X_6 = Y_5Y_6 - 3Y_3Y_6 + 3Y_2Y_6$,
78) $Y_5Y_4 = Y_6Y_2 - 3Y_5X_2 + 3Y_5Y_6$,
79) $Y_6X_5 = Y_8X_1 - 3Y_8Y_5 + 3Y_7X_5$,
80) $Y_1X_5 = Y_5Y_5 - 3Y_3Y_5 + 3Y_2Y_5$,
81) $Y_5Y_3 = Y_6Y_1 - 3Y_5X_1 + 3Y_5Y_5$,
82) $Y_6X_7 = Y_8X_3 - 3Y_8Y_7 + 3Y_7X_7$,
83) $Y_1X_7 = Y_5Y_7 - 3Y_3Y_7 + 3Y_2Y_7$,
84) $Y_7Y_3 = Y_8Y_1 - 3Y_7X_1 + 3Y_7Y_5$,
85) $Y_2X_8 = Y_4X_4 - 3Y_4Y_8 + 3Y_3X_8$,
86) $Y_1X_4 = Y_5Y_4 - 3Y_3Y_4 + 3Y_2Y_4$,
87) $Y_3Y_4 = Y_4Y_2 - 3Y_3X_2 + 3Y_3Y_6$,
88) $Y_2X_7 = Y_4X_3 - 3Y_4Y_7 + 3Y_3X_7$,
89) $Y_1X_3 = Y_5Y_3 - 3Y_3Y_3 + 3Y_2Y_3$,
90) $Y_3Y_3 = Y_4Y_1 - 3Y_3X_1 + 3Y_3Y_5$,
91) $Y_2X_6 = Y_4X_2 - 3Y_4Y_6 + 3Y_3X_6$,
92) $Y_1X_2 = X_5Y_2 - 3Y_3Y_2 + 3Y_2Y_2$,
93) $Y_1Y_4 = Y_2Y_2 - 3Y_1X_2 + 3Y_1Y_6$,
94) $Y_2X_5 = Y_4X_1 - 3Y_4Y_5 + 3Y_3X_5$,
95) $Y_1X_1 = Y_5Y_1 - 3Y_3Y_1 + 3Y_2Y_1$,
96) $Y_1Y_3 = Y_2Y_1 - 3Y_1X_1 + 3Y_1Y_5$.

We see that $V_4(U)$ is a quadratic algebra with generators $X_1, \dots, X_8$, $Y_1, \dots Y_8$ and relations $(i), (ii)$, 1)–96). This may not be the simplest presentation of $V_4(U)$. Observe that the generators Y_8 and Y_2 are linear combinations of other generators by (i) and (ii), so they can be removed from the generating set.

Acknowledgements

I wish to thank Inna Capdeboscq for calling my attention to subalgebras of Kac-Moody algebras and Efim Zelmanov for the idea of considering Veronese subalgebras to get quadratic algebras. I also thank my advisor Rostislav Grigorchuk for his assistance in writing this paper.

References

[AS74] Ralph K. Amayo and Ian Stewart. *Infinite-dimensional Lie algebras.* Noordhoff International Publishing, Leyden, 1974.

[Ber78] George M. Bergman. The diamond lemma for ring theory. *Adv. in Math.*, 29(2), 178–218, 1978.

[Ber83] A. E. Bereznyĭ. Discrete subexponential groups. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 123, 155–166, 1983. Differential geometry, Lie groups and mechanics, V.

[BF85] Jörgen Backelin and Ralf Fröberg. Koszul algebras, Veronese subrings and rings with linear resolutions. *Rev. Roumaine Math. Pures Appl.*, 30(2), 85–97, 1985.

[BG00] Laurent Bartholdi and Rostislav I. Grigorchuk. Lie methods in growth of groups and groups of finite width. In *Computational and geometric aspects of modern algebra (Edinburgh, 1998)*, volume 275 of *London Math. Soc. Lecture Note Ser.*, pages 1–27. Cambridge Univ. Press, Cambridge, 2000.

[Bou89] Nicolas Bourbaki. *Lie groups and Lie algebras. Chapters 1–3.* Elements of Mathematics (Berlin). Springer-Verlag, Berlin, 1989. Translated from the French, Reprint of the 1975 edition.

[dO03] M. P. de Oliveira. On 3-graded Lie algebras in a pair of generators: a classification. *J. Pure Appl. Algebra*, 178(1), 73–85, 2003.

[GK66a] I. M. Gel′fand and A. A. Kirillov. On fields connected with the enveloping algebras of Lie algebras. *Dokl. Akad. Nauk SSSR*, 167, 503–505, 1966.

[GK66b] I. M. Gelfand and A. A. Kirillov. Sur les corps liés aux algèbres enveloppantes des algèbres de Lie. *Inst. Hautes Études Sci. Publ. Math.*, (31), 5–19, 1966.

[Gov72] V. E. Govorov. Graded algebras. *Mat. Zametki*, 12, 197–204, 1972.

[Gri83] R. I. Grigorchuk. On the Milnor problem of group growth. *Dokl. Akad. Nauk SSSR*, 271(1), 30–33, 1983.

[Gri84] R. I. Grigorchuk. Construction of p-groups of intermediate growth that have a continuum of factor-groups. *Algebra i Logika*, 23(4), 383–394, 478, 1984.

[Gro81] Mikhael Gromov. Groups of polynomial growth and expanding maps. *Inst. Hautes Études Sci. Publ. Math.*, (53), 53–73, 1981.

[Kac85] V.G. Kac. *Infinite Dimensional Lie Algebras.* Cambridge University Press, 1985.

[KKM83] A. A. Kirillov, M. L. Kontsevich, and A. I. Molev. Algebras of intermediate growth. *Akad. Nauk SSSR Inst. Prikl. Mat. Preprint*, (39), 19, 1983. Translated in Selecta Math. Soviet. **9** (1990), no. 2, 137–153.

[Kob95] Yuji Kobayashi. A finitely presented monoid which has solvable word problem but has no regular complete presentation. *Theoret. Comput. Sci.*, 146(1-2), 321–329, 1995.

[Mil68] J. Milnor. A note on curvature and fundamental group. *J. Differential Geometry*, 2, 1–7, 1968.

[Pet93] V. M. Petrogradskiĭ. Some type of intermediate growth in Lie algebras. *Uspekhi Mat. Nauk*, 48(5(293)), 181–182, 1993.

[PP05] Alexander Polishchuk and Leonid Positselski. *Quadratic algebras*, volume 37 of *University Lecture Series.* American Mathematical Society, Providence, RI, 2005.

[She80] James B. Shearer. A graded algebra with a nonrational Hilbert series. *J. Algebra*, 62(1), 228–231, 1980.

[Smi76] M. K. Smith. Universal enveloping algebras with subexponential but not polynomially bounded growth. *Proc. Amer. Math. Soc.*, 60(1), 22–24, 1976.

[Ste75] Ian Stewart. Finitely presented infinite-dimensional simple Lie algebras. *Arch. Math. (Basel)*, 26(5), 504–507, 1975.

[Šva55] A. S. Švarc. A volume invariant of coverings. *Dokl. Akad. Nauk SSSR (N.S.)*, 105, 32–34, 1955.

[Ufn80] V. A. Ufnarovskiĭ. Poincaré series of graded algebras. *Mat. Zametki*, 27(1), 21–32, 157, 1980.

CONTACT INFORMATION

D. Koçak Department of Mathematics, Texas A&M University, College Station, Texas 77840
E-Mail(*s*): dkocak@math.tamu.edu

Received by the editors: 06.03.2015
and in final form 09.07.2015.

Algebra and Discrete Mathematics
Volume **20** (2015). Number 1, pp. 89–114

RESEARCH ARTICLE

A tabu search approach to the jump number problem

Przemysław Krysztowiak and Maciej M. Sysło

Communicated by D. Simson

Abstract. We consider algorithmics for the jump number problem, which is to generate a linear extension of a given poset, minimizing the number of incomparable adjacent pairs. Since this problem is NP-hard on interval orders and open on two-dimensional posets, approximation algorithms or fast exact algorithms are in demand.

In this paper, succeeding from the work of the second named author on semi-strongly greedy linear extensions, we develop a metaheuristic algorithm to approximate the jump number with the tabu search paradigm. To benchmark the proposed procedure, we infer from the previous work of Mitas [Order 8 (1991), 115–132] a new fast exact algorithm for the case of interval orders, and from the results of Ceroi [Order 20 (2003), 1–11] a lower bound for the jump number of two-dimensional posets. Moreover, by other techniques we prove an approximation ratio of $n/\log\log n$ for 2D orders.

1. Introduction

The jump number problem is to find a linear extension of a given poset minimizing the number of jumps, that is, incomparable adjacent pairs. It is best motivated by the following scheduling problem. Suppose a set of jobs is to be performed by a single machine, one at a time, with respect to

2010 MSC: 90C27, 90C59.
Key words and phrases: graph theory, poset, jump number, combinatorial optimization, tabu search.

some technological precedence constraints. Every job processed after one which is not constrained to precede it requires a warm-up (here, called a jump), which leads to a fixed unit of additional cost. The objective is to find a schedule which minimizes the number of warm-ups (jumps).

A purely theoretical interest in this problem is associated with a fact proved by Habib [8] that posets with isomorphic undirected comparability graphs have equal jump number. Consequently, the jump number fits the framework of comparability invariants, together with order dimension, the number of all linear extensions, the path partition number, and other properties. This leads to characterisation questions of comparability graphs satisfying given property and of possible interpretations of these invariants.

Another related topic is a classification of jump-critical posets. We recall from [26] that P is jump-critical if for any $p \in P$, $P\backslash\{p\}$ has less jumps than P. Some results have been established by El-Zahar et al. [26–28].

The main results of this work are as follows.

1) We design and benchmark a tabu search algorithm to approximate the jump number, see Section 3 and 4. It is built upon semi-strongly greedy linear extensions, defined by the second named author in terms of arc diagram representations of posets.
2) We give in Section 4.1 a new exact algorithm for the jump number of interval orders, based on the previous work of Mitas [16].
3) We show in Section 4.3 that the jump number has an $(n/\log\log n)$ approximation ratio on two-dimensional posets.

We now outline the paper structure.

The proposed tabu search algorithm explores semi-strongly greedy linear extensions, defined by the second named author (see Section 2.2). After reviewing the respective exact algorithm in Section 2, we verify how many solutions are generated when it is applied to various posets.

Our adaptation of the tabu search paradigm is proposed in Section 3 and tested in Section 4. In benchmarks we focus on two non-trivial classes of posets: interval orders and two-dimensional orders. A previous work of Mitas [16] on interval orders contains a characterization of optimal solutions in terms of subgraph packings.

In Section 4.1 we explore this idea to obtain a new exact algorithm for these orders. This allows us to calculate the jump number in reasonable time for posets having up to several hundred elements. In effect, we obtain

a testbed to verify the quality of solutions generated with the proposed tabu search algorithm.

Perhaps even more interesting is the case of two-dimensional orders, since the complexity status of the jump number problem has remained open in this class for several years now.

In Section 4.2 we exploit an interpretation given by Ceroi [4]. Chains to form a linear extension are seen as rectangles in the plane, and the bump number (see Section 1.1) corresponds to the maximum weight of an independent set (MWIS in short) of rectangles. Thus, a linear programming relaxation of the MWIS integer formulation yields a bound on $s(P)$. Even though the approximation ratio of our tabu search algorithm is unknown and only verified experimentally, we prove in Section 4.3 using other techniques that the jump number admits an $(n/\log\log n)$ approximation ratio on two-dimensional posets.

1.1. Preliminaries

We denote by $(P, <_P)$, or simply by P, a finite strict partially ordered set, in short a *poset* of cardinality$|P| = n$. That is, $<_P$ is a transitive and irreflexive relation on P. For any $p \in P$, $Succ_P(p) = \{q \in P : p <_P q\}$ is the *set of successors* of p and $Pred_P(p) = \{q \in P : q <_P p\}$ is the *set of predecessors* of p. $\mathcal{S}_P = \{Succ_P(p) : p \in P\}$ is the *family of distinct successor sets* and $\mathcal{P}_P = \{Pred_P(p) : p \in P\}$ is the *family of distinct predecessor sets* of a poset P. We say that p is covered by q if $p <_P q$ and for no r, $p <_P r <_P q$.

A *linear extension* $L = p_1, p_2, \ldots, p_n$ is a total ordering of P preserving the relation, that is, $p_i <_P p_j$ implies $i < j$. Two adjacent elements p_i, p_{i+1} in L form a *jump* if $p_i \not<_P p_{i+1}$ and otherwise they form a *bump*. Since jumps split L into chains of P, we can write $L = C_0 \oplus C_1 \oplus \ldots \oplus C_m$.

Problem 1.1.1. Let $s_L(P)$ denote the number of jumps in a linear extension L of a poset P. The *jump number problem* is to find

$$s(P) = \min\{s_L(P) : L \text{ is a linear extension of } P\}.$$

Problem 1.1.1 is equivalent to maximizing the number of bumps $b_L(P)$ amongst linear extensions of P, as for any L we have $s_L(P)+b_L(P) = n-1$. We write $b(P)$ for the maximum number of bumps in a linear extension of P. If $s_L(P) = s(P) = n - 1 - b_L(P)$ then L is called an *optimal linear extension* of P.

Definition 1.1.2. A poset $(P, <_P)$ is an *interval order* (or an *interval poset*) if there is a bijection between its elements and closed intervals on the real line, $P \longleftrightarrow \{I_p = [l(p), r(p)], l(p) \leqslant r(p)\}_{p \in P}$, such that $p <_P q$ if and only if $r(p) < l(q)$. $(P, <_P)$ is *two-dimensional* (or 2D) if $<_P$ is an intersection of two linear orders $\{L_1, L_2\}$, called a *realizer* of P.

Interval orders are characterized as posets excluding a subposet consisting of two independent 2-chains [6]. The recognition of two-dimensional posets can be accomplished in $\mathcal{O}(n^2)$ time, see [20] and also [10].

Definition 1.1.3. A chain C in P is *greedy* if $Pred(p) \cup \{p\} = C$, where $p = \sup C$, and for no element q covering p, the chain $C \cup \{q\}$ has this property. A linear extension $L = C_0 \oplus C_1 \oplus \ldots \oplus C_m$ is *greedy* if C_i is a greedy chain in $P \setminus \cup_{j<i} C_j$.

It is easy to prove that every poset has an optimal linear extension which is greedy.

1.2. Previous work

It has been proved by Pulleyblank [18] that the jump number problem is NP-hard on bipartite orders, that is, on posets having only minimal and maximal elements. Another NP-hardness proof has been given by Bouchitté and Habib [3]. Moreover, by Mitas [16] the problem remains NP-hard on interval orders. There are polynomial-time algorithms for some restricted classes of posets. These include semi-orders (i.e., interval orders formed by intervals of the same length) [2], and N-free orders (that is, with the N subposet forbidden) [19, 22]. The problem remains open on two-dimensional orders. However, Ceroi [4] proved NP-hardness of a generalized variant in which non-negative weights are associated with comparabilities and the objective is to maximize their sum on bumps of a linear extension. Due to high complexity of the problem, approximate algorithms are in demand. Those have been found only for interval orders (Sysło [25], Felsner [5], Mitas [16]). An exact algorithm has been designed by Sysło [24] (it is shortly reviewed in subsequent sections). As far as we know, the only metaheuristic approach published so far is that of Ngom [17], who adapted the genetic algorithm to the jump number problem.

2. Semi-strongly greedy linear extensions

In this paper a tabu search procedure is presented to search for valuable solutions amongst very particular greedy linear extensions, defined by the

second named author. Therefore, we now quickly recall what semi-strongly greedy linear extensions are and how they are generated. These linear extensions are formed from special greedy chains, defined by means of a digraph representation of $(P, <_P)$ explained below.

2.1. Arc diagrams

A *digraph* is denoted by $D = (V, A, t, h)$, where V is the vertex set, A is the arc set, and $t, h : A \to V$ are incidence mappings (t for tail, h for head). Then each arc $a \in A$ is of the form $a = (t(a), h(a))$. A sequence of arcs $\pi = (a_1, a_2, \ldots, a_l), l \geqslant 1$ is a *path* in D if $h(a_i) = t(a_{i+1})$ for $i = 1, 2, \ldots, l-1$. By $\mathrm{tc}(D) = (V, \mathrm{tc}(A), t^\star, h^\star)$ we denote a *transitive closure* of D, where $(a_1, \ldots, a_l), l \geqslant 1$ is a path in D if and only if $\mathrm{tc}(A)$ contains an arc b such that $t^\star(a_1) = t^\star(b)$ and $h^\star(a_l) = h^\star(b)$.

Definition 2.1.1. An *arc diagram* for a poset $(P, <_P)$ is an acyclic digraph $D(P) = (V, R, t, h)$ for which there is a mapping $\phi : P \to R$ such that for every $p, q \in P, p \neq q$, we have $p <_P q$ iff $(h^\star(\phi(p)), t^\star(\phi(q))) \in R^\star$, where $t^\star, h^\star$ are the incidence mappings of $\mathrm{tc}(D)$ and $R^\star = \mathrm{tc}(R) \cup \{(v, v) : v \in V\}$. An arc $a \in \phi(P)$ is a *poset arc* and otherwise a is a *dummy arc*.

Informally, certain arcs represent the elements of P, and the purpose of the remaining ones is to preserve the comparabilities along the paths of $D(P)$.

An example is shown in Figure 1. Elements $2, 3, 12, 8$ form one of the chains in this poset, so the four corresponding arcs are aligned in one of the paths leading from the source to the sink of the diagram. In any linear extension, these (and other) order constraints have to be respected, e.g., $(2, 3, 10) \oplus (7) \oplus (14, 5, 1) \oplus (4, 13) \oplus (6, 12) \oplus (11, 9, 8)$.

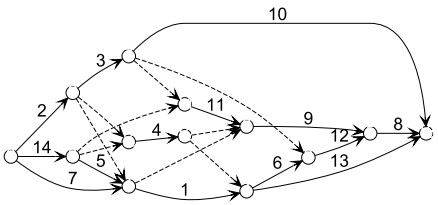


Figure 1. Arc diagram representation of a two-dimensional poset

Figure 2 is another example, whose one of linear extensions is $(1, 3) \oplus (2, 7) \oplus (4, 9) \oplus (5, 10) \oplus (6, 11) \oplus (8)$.

An algorithm to construct an adequate arc diagram for any finite poset P was given by Sysło [21]. For the sake of completeness, it is repeated as Algorithm 1.

Algorithm 1 Arc diagram for a poset P (see [21])

Input: A finite poset $(P, <_P)$.
Output: $D(P)$, an arc diagram for P.
Step **1. Let** $\mathcal{P}_P = \{Pred_1, \ldots, Pred_k\}$, $\mathcal{S}_P = \{Succ_1, \ldots, Succ_l\}$.
For each $Pred_i$ **let** $U_i = \bigcap_{p \in Pred_i} Succ_P(p)$.
Step **2.** {The vertices:}
Let $x_1, x_2, \ldots, x_h$ correspond to those $Pred_i$, for which there exists $Succ_j = U_i$.
Let $y_{h+1}, y_{h+2}, \ldots, y_k$ correspond to remaining $Pred_i$.
Let $z_{h+1}, z_{h+2}, \ldots, z_l$ correspond to remaining $Succ_j$.
Let $y_i = z_i = x_i$ for $i = 1, 2, \ldots, h$.
Step **3.** {The arcs:}
For each $p \in P$ add a poset arc (y_i, z_j), where
$P_i = Pred_P(p)$, $Succ_j = Succ_P(p)$.
For every $p, q \in P$, $p \neq q$,
if p is covered by q, and $z_j \neq y_i$, where
$Succ_j = Succ_P(p)$ and $Pred_i = Pred_P(q)$,
then add a dummy arc (z_j, y_i).
Finally, remove transitive arcs provided that they are not poset arcs.

2.2. Greedy paths

Definition 2.2.1. In an arc diagram $D(P)$, a natural counterpart of a greedy chain C (see Section 1.1) is a *greedy path* $\pi(C) = (a_1, a_2, \ldots, a_l)$, $l \geqslant 1$, i.e., a path satisfying the following conditions:

- No vertex of $\pi(C)$ except $h(a_l)$ is a head of any arc other than a_j, $j = 0, 1, \ldots, l-1$.
- a_j is a poset arc, $j = 1 \ldots l$.
- $\pi(C)$ cannot be extended to a longer path satisfying the above two conditions.

In the reverse direction, any greedy path π induces a greedy chain C_π in $D(P)$. Thus, an algorithm to generate a greedy linear extension can be formulated in terms of an arc diagram for P, see Algorithm 2.

Algorithm 2 Greedy linear extension

Input: $D(P)$, an arc diagram for P.
Output: $L = C_0 \oplus C_1 \oplus \ldots \oplus C_m$, a linear extension of P.
Step **1**. { Initialization }
$D := D(P)$
$L := \varnothing$
Step **2**. **while** $D \neq \varnothing$
($\star$) find a greedy path π in D
$L := L \oplus C_\pi$
$D := D(P \backslash L)$, an arc diagram for the remaining poset
Step **3**. **return** L

It is easy to design an analogous procedure to enumerate all greedy linear extensions, but initial experiments reveal quickly that it is a very time-consuming and hence inefficient process.

However, it was proved in a series of papers [21–25], that the search space can be significantly reduced, since for every poset the class of greedy linear extensions can be further restricted to the class of very particular greedy linear extensions which contains an optimal solution. These linear extensions are composed from two special types of greedy chains, as described below.

Definition 2.2.2. A *strongly greedy path* π is a greedy path satisfying:

- either $h(\pi)$ is the sink of $D(P)$, or
- $h(\pi)$ is the head of a poset arc $b \neq a_l$ such that no path terminating with b has a vertex incident with a dummy arc.

If $D(P)$ contains at least one strongly-greedy path π then there always exists an optimal linear extension beginning with C_π. If there are no strongly-greedy paths in $D(P)$ then there is at least one *semi-strongly greedy path*, i.e., a greedy path π such that

- π has a vertex which is a tail of a dummy arc but not a head of a dummy arc.

In such case, when searching for an optimal linear extension, we have to consider all semi-strongly greedy paths. It should be noted however that not every greedy path is semi-strongly greedy, so all in all, the search space is greatly reduced in comparison with an enumeration of all greedy linear extensions.

The arc diagram in Figure 2 has three semi-strongly greedy paths $(1, 3)$, $(1, 4)$ and $(1, 5)$. The path (2) is greedy, but it is neither strongly

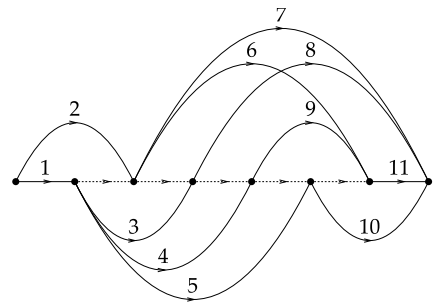


FIGURE 2. Arc diagram representation of an interval order

greedy, nor semi-strongly greedy. But the diagram in Figure 1 has two strongly greedy paths. The path $(2, 3, 10)$ passes through a tail of a dummy arc, so it would classify as semi-strongly greedy, but it also terminates in the sink, so it is strongly greedy. In addition, the vertex terminating $(14, 5)$ terminates also (7), which has no vertex incident with a dummy arc. So $(14, 5)$ is strongly greedy.

In conclusion, we have the following Theorem 2.2.3.

Theorem 2.2.3 (Sysło [23])**.** *Every poset has an optimal linear extension* $L = C_0 \oplus C_1 \oplus \ldots \oplus C_m$*, called semi-strongly greedy, such that each chain* C_i *is strongly greedy in* $P_i = P \setminus \bigcup_{j<i} C_j$ *or semi-strongly greedy in* P_i *if* P_i *has no strongly gredy chains.*

To design an algorithm generating one semi-strongly greedy linear extension, we simply replace Step **2**. $(\star)$ of Algorithm 2 with

$(\star)$ find a strongly greedy path π in D; **if** no such path has been found **then** set π to any semi-strongly greedy path in D.

It is now also easy to devise an exact algorithm for the jump number problem, which searches for optimal solution amongst all semi-strongly greedy linear extensions via backtracking. For this purpose, in Step **2**. $(\star)$ of Algorithm 2, if there are no strongly greedy paths, then instead of choosing an arbitrary semi-strongly greedy path we verify every one of them, and apply the procedure recursively on every respective subposet. We refer to this algorithm as `OptLinExt` [24] in subsequent sections, where a new tabu search algorithm is proposed, based on these special linear extensions. For a more in-depth treatment of the topic we refer the reader to the articles of Sysło [21–25].

2.3. The running time of `OptLinExt`

Let k denote the number of dummy arcs in $D(P)$. It was concluded in [24] that the pessimistic time complexity of `OptLinExt` is of order $\mathcal{O}(k! \cdot \text{poly}(n, k))$, since there are always at most $k!$ semi-strongly greedy linear extensions. That is, k is an important factor contributing to the complexity of the problem.

We have performed an experiment to learn how many solutions are generated in reality on posets for varying number of dummy arcs. For a fixed poset size ($n = 120$ in the case of interval orders and $n = 30$ in the case of two-dimensional orders), and for each number of dummy arcs $k \in \{5, 10, \ldots\}$, a hundred of posets were randomized, having these requested properties. To obtain posets with a given number of dummy arcs in their arc diagrams, we use a genetic algorithm, with distance from k in question being the optimality factor. The maximum number of generated solutions amongst a group of posets with k dummy arcs was recorded. This is plotted in Figure 3.

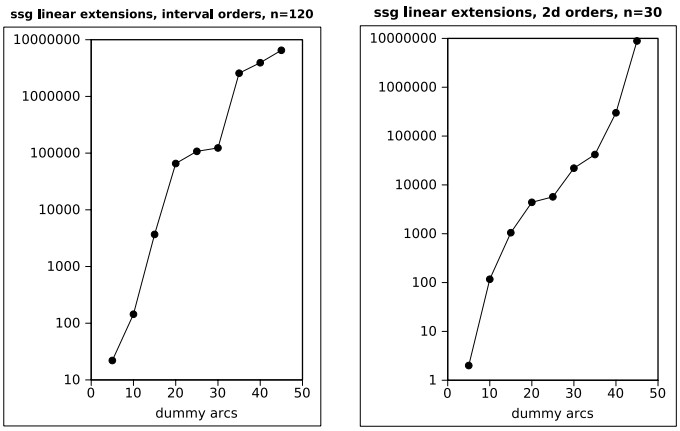


Figure 3. Total number of semi-strongly greedy linear extensions in sample posets containing increasing number of dummy arcs

This series of trials helps to asses the magnitude of the search space, browsed through by the tabu search algorithm described in Section 3. Interestingly, it shows that the search space is typically greater in the case of two-dimensional posets.

In the group of interval orders on 120 elements, containing 45 dummy arcs, a poset was spotted having 6 510 338 semi-strongly greedy linear extensions, and the exact algorithm took over 5 hours to list them. Typically,

the total number of solutions is lower. On the other hand, a 30-element two-dimensional poset was found, with 45 dummy arcs in its diagram, for which the search space contains 8 857 068 semi-strongly greedy linear extensions. In this case, the execution of `OptLinExt` took 3 445 seconds, benefitting from shorter time required to rebuild arc diagrams for no more than 30 elements.

We note that the number of dummy arcs is lesser than n on interval orders. This is not the case on two-dimensional posets. Thus, it is common for two-dimensional orders to have a significantly greater search space of semi-strongly greedy linear extensions than interval orders.

3. A tabu search algorithm to approximate the jump number

Tabu search is a famous algorithmic technique of moving stepwise towards an optimal solution of a computational problem. Its characteristic feature is maintaining a list of moves not allowed at given iteration, called a tabu list. The purpose of this list is the avoidance of repeatedly visiting the same solutions. In recent years, tabu search has become a major metaheuristic paradigm to approximate hard optimization problems. The rationale behind this method can be found in the monograph of Glover and Laguna [7].

In a tabu search algorithm, every solution is treated as a point in the search space. Initially, there is some solution *sol*, and in each step we move to another solution *sol*′ selected from a neighbourhood of *sol*. That is, a fixed number of solutions is generated from *sol*, and the algorithm follows to the best of them, provided it is not a tabu move. We now turn to a description of our adaptation for the jump number problem.

3.1. An adaptation for the jump number problem

In our approach, every solution is some semi-strongly greedy linear extension L of P. A neighbour solution is generated from L by splitting it between some consecutive chains, and completing it according to the semi-strongly greedy algorithm (see Section 2.2). A linear extension L is represented as a list of chains, which in turn are lists of poset elements. So if we decide to split L after kc chains, then $L[0] \oplus \ldots \oplus L[kc-1]$ becomes the initial part of a neighbour L'. Our adaptation is shown as Algorithms 3 and 4.

Algorithm 3 `TS-CompleteLinExt`

Input: A poset P, initial chains of some greedy linear extension $partialL = C_0 \oplus \ldots \oplus C_{c-1}$, and the minimum number of jumps $s_\star$ found so far.

Output: $L = C_0 \oplus \ldots \oplus C_{c-1} \oplus C_c \oplus \ldots \oplus C_m$, a linear extension of P.

Step **1**. { Initialization }
  $D := D(P\backslash partialL)$, an arc diagram for $P\backslash partialL$
  $c :=$ the number of chains in $partialL$
  $e :=$ the number of poset elements in $partialL$
  $s_{LB} := s(partialL) + 1 + s_{LB}(D)$
  { $s_{LB}(D)$ is the lower bound on $s(P\backslash partialL)$, Thm. 3.1.1 }

Step **2**. **if** $s_{LB} \geqslant s_\star$ **then** add (c, e) to $TabuPositions$, **return** $\varnothing$
  else go to 3

Step **3**. **if** D has no more dummies than $MaxDummies$ **then**
    $remainingL :=$ `OptLinExt`$(P\backslash partialL)$
    add (c, e) to $TabuPositions$
    if $s(partialL) + 1 + s(remainingL) \geqslant s_\star$ **then return** $\varnothing$
    else return $partialL \oplus remainingL$
  else go to 4

Step **4**. $remainingL := \varnothing$
  { complete the linear extension with s.-s.-greedy chains }
  while $D \neq \varnothing$
    $S :=$ strongly greedy paths in D
    $W :=$ semi-strongly greedy paths in D
    if $|S| > 0$ **then**
      $\pi :=$ any path from S
      $remainingL := remainingL \oplus C_\pi$
      if C_π is the first chain in $remainingL$ **then**
        add (c, e) to $TabuPositions$
    else { no strongly greedy paths }
      if $|W| = 1$ **then**
        $\pi :=$ the path from W, $remainingL := remainingL \oplus C_\pi$
        if C_π is the first chain in $remainingL$ **then**
          add (c, e) to $TabuPositions$
      else { several semi-strongly greedy paths }
        if $remainingL = \varnothing$ {i.e, first choice after split} **then**
          $\pi :=$ a random path from $W\backslash TabuPaths$;
          if not found **then** add (c, e) to $TabuPositions$, **return** $\varnothing$
          { π added to $TabuPaths$ in Alg. 4}
        else { not first chain after split } $\pi :=$ a random path from W
        $remainingL := remainingL \oplus C_\pi$
      $D := D(P\backslash(partialL \oplus remainingL))$

Step **5**. **return** $partialL \oplus remainingL$

Algorithm 4 `TS-OptimizeJumpNumber`

Input: A poset P, the number of iterations T.
Output: $L = C_0 \oplus \ldots \oplus C_m$, a semi-strongly greedy linear extension of P.

Step **1**. { Initialization }
$TabuPositions := \varnothing$
$TabuPaths := \varnothing$
$currentL :=$ `TS-CompleteLinExt`$(P, \varnothing, |P|)$
$bestL := currentL$
$t := 0$ {current iteration}

Step **2**. **while** $t \leqslant T$
$bestNeighbour = \varnothing$
{ the split position for the best neighbour }
$bestKC = 0$, $bestKE = 0$
$t := t + 1$

Step **3**. { select the best neighbour solution }
for $n = 1$ **to** $CheckedNeighbours$
$kc :=$ a random number less than $|currentL|$
$ke := |L[0]| + \ldots + |L[kc-1]|$
{ such that $(kc, ke) \notin TabuPositions$ }
$partialL :=$ first kc chains of $currentL$
$neighbourL :=$ `TS-CompleteLinExt`$(P, partialL,$
$s(bestLinExt))$
if $s(neighbourL) < s(bestNeighbour)$ **then**
$bestNeighbour := neighbourL$
$bestKC := kc$, $bestKE := |neighbourL[0..bestKC-1]|$ { $= ke$ }
$n := n + 1$

Step **4**. { move to the neighbour, update the result }
if $bestNeighbour \neq \varnothing$ **then**
$currentL := bestNeighbour$
update $TabuPaths$ with $(bestKC, bestKE,$
$currentL[bestKC-1], currentL[bestKC])$
if $s(currentL) < s(bestL)$ **then** $bestL := currentL$

Step **5**. **return** $bestL$

We keep two tabu lists. $TabuPositions$ is a cyclic list containing $TabuSize$ recent split positions. More precisely, if L is split after kc chains, then the pair (kc, ke) may be added to $TabuPositions$, where $ke = |L[0]| + \ldots + |L[kc-1]|$ is the total number of poset elements in kept chains. The second tabu list, $TabuPaths$, contains the greedy paths before and after split positions. That is, whenever L is split after kc chains and completed, we add to $TabuPaths$ the quadruple $(kc, ke, L[kc-1], L[kc])$. This is motivated by fact that there are two major decision points when generating a neighbour: firstly, L is split at some position kc; secondly, one of available greedy paths in $D(P\backslash(L[0] \oplus \ldots \oplus L[kc-1]))$ is selected. $TabuPositions$ is a cyclic list, so after adding a new entry, the oldest one is removed. On the other hand, $TabuPaths$ is a static list, from which no entry is removed in the process of the algorithm.

When generating or completing a linear extension, an obvious change is made with respect to the original semi-strongly greedy algorithm: a greedy path is not allowed to be selected, if it is contained in the tabu list $TabuPaths$ associated with current split position (Step **4** of Algorithm 3). Further, while `OptLinExt` proceeds with greedy paths in systematic manner, here we always select one at random. Obviously, the implementation is also augmented to update both tabu lists in each iteration.

A split position (kc, ke) is added to $TabuPositions$ when the remaining subposet has either a strongly greedy path, or only one semi-strongly greedy path, or when its jump number is assessed as non-promising in Step **2**, or has been completed exactly in Step **3** (Algorithm 3). In other words, (kc, ke) is not tabu, if there are still some unexplored choices of greedy paths in the remaining subposet. Each choice of a greedy path is recorded in $TabuPaths$.

The most time-consuming subprocedure is the construction of an arc diagram for the remaining subposet whenever a greedy path is selected and added to L. Therefore, it is important to quickly reject those solutions whose number of jumps will not improve over the best one found in the preceding course of the algorithm. Thus, when a split point is selected and the diagram is reconstructed, we asses this choice by calculating the lower bound for the jump number of the remaining poset. It may immediately turn out that another split position should be randomized. We use a lower bound given by the following theorem.

Theorem 3.1.1 (Sysło [24]). *If $D(P)$ is an arc diagram of a poset P then $\sum_{v\in V}\max\{0,\mathrm{indeg}_P(v)-1\}\leqslant s(P)$, where $\mathrm{indeg}_P(v)$ is the number of poset arcs coming into v.*

If the number of dummy arcs in the diagram is less than some fixed number $MaxDummies$, we apply `OptLinExt` to the remaining poset.

4. Benchmarks

In this section we benchmark the proposed tabu search algorithm on two non-trivial classes of posets.

4.1. Interval orders

In the case of interval orders, we first run `TS-OptimizeJumpNumber` and compare the quality of its solutions with optimal ones, obtained via a reduction to the subgraph packing problem. We now explain how these optimal values are computed.

The jump number of interval orders

Interval orders have a well-known characterization (see Fishburn [6]) which includes their canonical representation, that is, Algorithm 5.

Algorithm 5 Canonical representation of an interval order (see [16])

Input: $(P,<_P)$, an interval order.
Output: $\{I_p=[l(p),r(p)]\}_{p\in P}$, a compact family of intervals representing $(P,<_P)$.
Step **1**. Sort $(\mathcal{S}_P,\subseteq)$: $Succ_1\supseteq Succ_2\supseteq\ldots\supseteq Succ_e=\varnothing$.
Step **2**. Sort $(\mathcal{P}_P,\subseteq)$: $\varnothing=Pred_1\subseteq Pred_2\subseteq\ldots\subseteq Pred_e$.
Step **3**. Assign to each $p\in P$ its left endpoint $l(p)=i-1$ such that $Pred_i=Pred_P(p)$ and its right endpoint $r(p)=j-1$ such that $Succ_j=Succ_P(p)$.

The obtained *canonical intervals* are then written into a *table* of size $e\times e$, where $e=|\mathcal{P}_P|=|\mathcal{S}_P|$ (see Figure 4). For an interval $[l(p),r(p)]$ its corresponding element p is put in the cell in row $l(p)$, column $r(p)$. Then, successive bumps of a linear extension may be read from the table along a sequence of ordered pairs of the form $T=\{t_i=(t_{col},t_{row})\}_{i=1,\ldots,b}$, called a *bump sequence*, where $b\leqslant e-1$ is its length. Every bump t in T satisfies $t_{col}<t_{row}$ and for every two consecutive bumps (s,t), $s_{row}\leqslant t_{col}$.

We say that the columns and rows outside T are *omitted*. The problem is to generate a bump sequence of maximum cardinality, i.e., to decide, which rows and columns should be omitted so as to obtain a *realizable* bump sequence (so some L can be read along it). Given a realizable bump sequence of length b_T one applies a quick procedure (see [16]) to obtain a linear extension with at least b_T bumps.

Definition 4.1.1. *Graph of intervals* $G_I(P)$ takes non-empty table cells as vertices. Edges are added for vertices positioned consecutively in a column or in a row (see Example 4.1.6 and Figure 4). A component C of this graph is *unsaturated* if none of its vertices is situated on the boundary of the table (i.e., in the lowest row or in the rightmost column, or on the diagonal), nor any cell of C contains a multiple element, nor C itself contains a cycle.

The number of unsaturated components is denoted by u and is typically much smaller than $e \leqslant n$.

In [16], Mitas characterizes realizable bump sequences as follows.

Theorem 4.1.2. *For a bump sequence T to be realizable it suffices that each unsaturated component C satisfies one of two properties:*

1) $(P1)$ *C contains a vertex in a column or in a row which is omitted by T.*
2) $(P2)$ *C contains an element $[j, j+q]$ such that the columns $j, \ldots, j+q-1$ and the rows $j+1, \ldots, j+q$ are omitted by T.*

Theorem 4.1.2 motivates the following definition.

Definition 4.1.3. In the *graph of unsaturated components* $G_U(P)$ vertices correspond to unsaturated components of $G_I(P)$. An edge joins v_C and v_D if component C contains a vertex in column i and component D contains a vertex in row i or row $i+1$.

Then, the properties $P1$ and $P2$ of bump sequences are mapped to edges and certain odd cycles (called *valid cycles*) of $G_U(P)$. With each edge and with each valid cycle there is an associated set of pairs (col_i, row_i) or (col_i, row_{i+1}), which when removed from a bump sequence result in $(P1)$ or $(P2)$ being satisfied by the corresponding unsaturated components. In this way the jump number problem is reduced to a subgraph packing problem which is to find an optimal packing of vertex-disjoint edges and

valid cycles. A packing involving v vertices and c cycles yields a bump sequence with $u - \frac{v+c}{2}$ lost bumps, so $v + c$ is to be maximized. We refer the reader to the article of Mitas [16] and to a recent work of the first named author [13, 14] for a detailed explanation of this reduction.

The following result is a corollary from the previous work of Mitas concerning approximation of the jump number.

Corollary 4.1.4. *The jump number of an interval order P can be computed in time $\mathcal{O}(2^n \cdot poly(n))$.*

Proof. By Mitas, for an optimal bump sequence there is an optimal packing with respect to $\frac{v+c}{2}$. It can be seen that the number of valid cycles is bounded by $e \leqslant n$. Hence, it suffices to enumerate all packings of vertex-disjoint valid cycles, and supplement each such packing H with a maximum matching M on $G_U(P)\backslash H$. From all candidate packings $H + M$ we choose the one maximizing $\frac{v+c}{2}$. □

Remark 4.1.5. Corollary 4.1.4 is an alternative proof of a more general fact, that the jump number can be computed in time dominated by 2^n for an arbitrary poset. Indeed, with any P one associates an instance of the Traveling Salesman Problem, in which vertices correspond to the points of P, and the distances (travel costs) are as follows:

- if $p <_P q$ then $c_{pq} = 0$,
- if $p \not<_P q$ then $c_{pq} = 1$,
- otherwise $c_{pq} = \infty$.

One more vertex d is added such that distances from d to the minimal elements are 0 and distances from the maximal elements to d are 0. Then, for every linear extension L there is a corresponding Hamiltonian cycle from d to d. Total cost of each such cycle is equal to $s_L(P)$ and other permutations contain a connection of cost ∞. Hence, a dynamic programming algorithm of Held and Karp [9] for TSP can be used to compute $s(P)$ in time complexity $\mathcal{O}(2^n \cdot \mathrm{poly}(n))$.

Nonetheless, in our experiments it is beneficial to apply the algorithm described in the proof of Corollary 4.1.4. In practice, due to typical structure of arising graphs, by incorporating simple heuristics, this algorithm is of satisfactory speed. There is often a limited number of valid cycles in $G_U(P)$. Moreover, they usually overlap, so that an exhaustive search algorithm may avoid many subsets of cycles, which have at least one common vertex. Thus, the jump number can be computed in reasonable

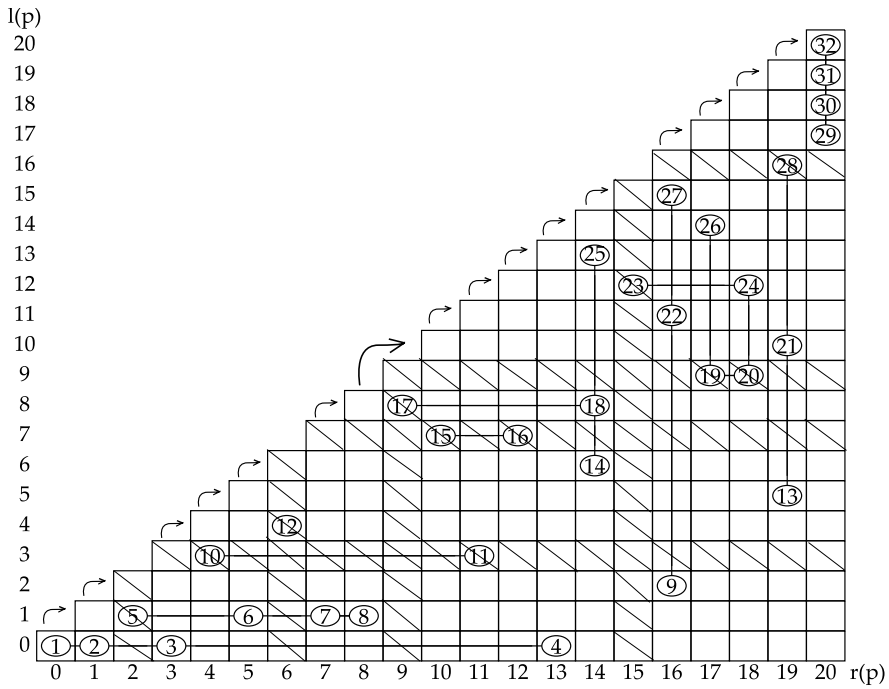


Figure 4. An interval order in canonical representation

time even for interval orders P having several hundred elements (i.e., within 10 seconds).

Example 4.1.6. An interval order in Figure 4 consists of 32 elements, with graph of intervals forming 10 components, of which 8 are unsaturated. Hence, its graph of unsaturated components (defined above) has 8 vertices, and there are three valid cycles: one pentagon, and two triangles. An optimal packing is composed of the triangle $\{\{9, 22, 27\}, \{13, 21, 28\}, \{19, 20, 23, 24, 26\}\}$ and edges $\{\{12\}, \{15, 16\}\}$, $\{\{10, 11\}, \{5, 6, 7, 8\}\}$. Corresponding bump sequence is visualised by arrows above the table, denoting successive bumps of the optimal linear extension.

Random interval orders

Our first experiment was to verify the quality of solutions generated by the proposed tabu search procedure on random interval orders. For each pair $(n, k), n \in \{100, 150, 200\}, k \in \{10, 20, \ldots, \frac{n}{2}\}$ a hundred of interval orders were randomized, consisting of n elements and containing k dummy

arcs in their arc diagrams. `TS-OptimizeJumpNumber` was started for every such instance, with a limit of iterations equal to n. The algorithm was aborted upon finding an optimal solution (known a priori by Corollary 4.1.4). Every time, an optimal linear extension has been found. Amongst 100-element posets it took at worst 45 iterations (10 seconds) for an instance containing 40 dummies. Amongst 200-element posets the worst found case required 42 iterations (47 seconds) and contained 60 dummies. On average, optimum was found after 20 iterations.

Hard interval orders

In our previous research concerning approximation of the jump number on interval orders [12], a vast amount of posets were recognized, which result in suboptimality of solutions generated by formerly known algorithms. We use them to benchmark approximation algorithms of Felsner [5], Sysło [25], and Mitas [16]. Consequently, we use them now to benchmark the tabu search procedure proposed in Section 3. For example, the poset in Figure 2 contains three semi-strongly greedy paths. It is unknown which of them should be chosen in order to reach an optimal linear extension. If many similar posets are joined by series compositions, then there are very many decision points in a process of a greedy algorithm.

In our second experiment those hard interval orders were taken on 100, 150 and 200 elements. Samples of 10 results for each cardinality n are reported in Tables 1, 2, 3. Each entry corresponds to one input instance P, and contains: the number of dummy arcs in P, its jump number $s(P)$, the sequence of approximate solutions generated by `TS-OptimizeJumpNumber`, the number of iterations, the running time in seconds, and the error $\frac{s_{APX}(P)-s(P)}{s(P)}$ of the best found linear extension. If n iterations did not suffice to find the optimum, the time of last improvement in the tabu search process is also reported.

Observation 4.1.7. The proposed tabu search algorithm, applied to n-element interval orders, generates linear extensions with no more jumps than 105% of optimum, in n iterations.

4.2. Two-dimensional posets

The complexity status of the jump number problem on 2D posets is unsettled. Even though it has not been classified as NP-hard, we are attempting to solve the following Problem 4.2.1.

	d.a.	$s(P)$	$s_{APX}(P)$	iterations	time	error
P_1	28	22	30 . . . 22	32	7 s	0.0000
P_2	32	30	39 . . . 30	48	24 s	0.0000
P_3	29	24	29 . . . 24	89	25 s	0.0000
P_4	32	26	33 . . . 26	75	19 s	0.0000
P_5	28	21	25 . . . 21	18	7 s	0.0000
P_6	26	22	26 . . . 22	20	6 s	0.0000
P_7	31	28	32 . . . 28	26	10 s	0.0000
P_8	32	26	34 . . . 26	17	8 s	0.0000
P_9	35	29	35 . . . 30	100 (36)	17 s (8 s)	0.0345
P_{10}	37	31	36 . . . 31	59	38 s	0.0000

Table 1. Performance of tabu search on 100-element interval orders

	d.a.	$s(P)$	$s_{APX}(P)$	iterations	time	error
P_1	34	30	36 . . . 30	47	19 s	0.0000
P_2	46	32	39 . . . 32	14	19 s	0.0000
P_3	45	34	42 . . . 35	150 (26)	81 s (18 s)	0.0294
P_4	50	44	55 . . . 46	150 (90)	123 s (85 s)	0.0455
P_5	37	32	39 . . . 32	36	17 s	0.0000
P_6	51	42	52 . . . 43	150 (27)	35 s	0.0238
P_7	38	36	45 . . . 36	36	52 s	0.0000
P_8	42	37	45 . . . 37	52	66 s	0.0000
P_9	45	35	42 . . . 35	39	20 s	0.0000
P_{10}	43	37	42 . . . 38	150 (29)	143 s (31 s)	0.0270

Table 2. Performance of tabu search on 150-element interval orders

	d.a.	$s(P)$	$s_{APX}(P)$	iterations	time (s)	error
P_1	50	45	52 . . . 46	200 (164)	380 (336)	0.0222
P_2	54	50	61 . . . 52	200 (36)	352 (101)	0.0400
P_3	57	49	61 . . . 51	200 (13)	258 (40)	0.0408
P_4	54	49	62 . . . 51	200 (158)	415 (362)	0.0408
P_5	56	50	62 . . . 52	200 (16)	306 (43)	0.0400
P_6	56	48	57 . . . 49	200 (126)	328 (228)	0.0208
P_7	63	49	57 . . . 50	200 (66)	327 (132)	0.0204
P_8	54	48	54 . . . 50	200 (6)	160 (10)	0.0417
P_9	54	49	59 . . . 50	200 (73)	235 (109)	0.0204
P_{10}	65	53	64 . . . 54	200 (50)	413 (144)	0.0189

Table 3. Performance of tabu search on 200-element interval orders

Problem 4.2.1. Give an approximation algorithm for the jump number problem on two-dimensional posets.

For this purpose it is useful to analyze the performance of general-purpose algorithms on 2D orders. For $n > 50$ we usually fail to compute $s(P)$ exactly in reasonable time. Hence, we compare each tabu search result with a lower bound on the $s(P)$, inferred from the results of Ceroi.

The bump number of two-dimensional posets

As observed by Ceroi [4], for two-dimensional posets the jump number can be interpreted as the problem of finding a maximum weight independent set of a family of axis-parallel rectangles corresponding to certain chains of P. Let P be a two-dimensional poset with realizer $\{L_1, L_2\}$. With $p \in P$ we associate a point $(x, y) \in \mathbb{R}^2$ such that x is the position of p in L_1 and y is the position of p in L_2. Then, each chain of P is easily seen as a rectangle in $\mathbb{R}^2$. Linear extensions are formed only from chains C which are convex, i.e., $\forall p <_P q <_P r \in P$, if $p \in C$ and $r \in C$ then $q \in C$. If by $R(P) = (V_R, E_R)$ we denote the graph of rectangle intersections, in which a weight for each vertex $v \leftrightarrow C_v$ is equal to its bumps, i.e., $w(v) = |C_v| - 1$, then we have the following result.

Lemma 4.2.2 (Ceroi [4])**.** *The maximum bump number $b(P)$ is equal to the maximum weight of an independent set in $R(P)$.*

Hence, to calculate $b(P)$ it suffices to solve an integer linear program for maximum weight independent set (MWIS), $\max \sum_v w(v) \cdot x(v)$ s.t. $x(v) \in \{0, 1\}$ for each $v \in V_R$, and $x(u) + x(v) \leqslant 1$ for each $(u, v) \in E_R$. However, $|V_R|$ is usually greater than $|P|$ and exact computation of MWIS quickly becomes infeasible. Therefore, we only get an upper bound $b_{UB}(P) \geqslant b(P)$ by solving an LP-relaxation of this formulation, i.e., with $0 \leqslant x(v) \leqslant 1$ (and so a lower bound $s_{LB}(P) = \lceil n - 1 - b_{UB}(P) \rceil$ for the jump number).

Random two-dimensional posets

In the first experiment concerning two-dimensional posets n was set to 30. For 30-element posets optimal solution can be computed by `OptLinExt`. On larger instances we have to resort to the LP-relaxation which provides an upper bound for $b(P)$.

Tens of random 30-element two-dimensional orders were generated containing a varied number of dummy arcs (from 10 to 50). First, $s(P)$

$b(P)$	16	17	18	17	16
$b_{UB}(P)$	17.25	18.66	19.0	17.5	17.5
$b(P)$	17	19	18	16	17
$b_{UB}(P)$	18.33	19.75	19.88	17.33	18.11

Table 4. Discrepancy between the bump number and its LP relaxation, $n = 30$

	d.a.	$\lvert V_R \rvert$	$b_{UB}(P)$	$s_{LB}(P)$	$s_{APX}(P)$	error
P_1	84	341	41.5	18	23 . . . 19	0.0556
P_2	81	339	39.5	20	25 . . . 22	0.1000
P_3	74	375	38.5	21	24 . . . 22	0.0476
P_4	66	349	41.5	18	23 . . . 20	0.1111
P_5	100	337	40.75	19	22 . . . 19	0.0000
P_6	90	297	39.125	20	26 . . . 23	0.1500

Table 5. Performance of tabu search on 60-element 2D posets

was computed. Then, the tabu search procedure was started with a limit of n iterations. Optimum was found by `TS-OptimizeJumpNumber` in all cases beside one poset with 50 dummies, for which a linear extension was generated having 12 jumps instead of 11. On all inputs, the algorithm required at most 21 iterations and no more than 2 seconds.

By this occasion, $b(P)$ was additionally compared with the linear programming relaxation $b_{UB}(P)$ of the equivalent MWIS problem. Some typical discrepancies between these values are shown in Table 4.2. The number of vertices in $R(P)$, i.e., the number of convex chains in P, varied from 96 to 138.

In the next experiment, 150 posets were generated with $n = 60$. This time, the approximated number of jumps was compared only with the lower bound obtained via an LP relaxation of MWIS on $R(P)$, that is, $s_{LB}(P) = \lceil n - 1 - b_{UB}(P) \rceil$. The results for a sample of 6 posets are reported in Table 5.

An average error of $s_{APX}(P)$, when compared against $s_{LB}(P)$, was 0.12. In the worst spotted case, it was 0.29. The number of dummy arcs in these posets ranged from 60 to 100. The number of vertices in $R(P)$ ranged from 290 to 403. The running time was always below 12 seconds.

Finally, 30 two-dimensional posets with $n = 90$ were taken. The error of s_{APX} obtained with n iterations of our algorithm was on average 0.15 and at most 0.29. The time of optimization was always below 45 seconds

	d.a.	$\lvert V_R\rvert$	$b_{UB}(P)$	$s_{LB}(P)$	$s_{APX}(P)$	error
P_1	166	593	63.57	26	35 . . . 29	0.1154
P_2	150	566	62.44	27	33 . . . 30	0.1111
P_3	158	570	60.96	29	38 . . . 31	0.0690
P_4	167	599	61.58	28	39 . . . 36	0.2857
P_5	163	631	63.00	26	33 . . . 29	0.1154
P_6	173	557	60.17	29	39 . . . 34	0.1724

TABLE 6. Performance of tabu search on 90-element 2D posets

(it took much longer to compute the lower bound $s_{LB}(P)$ with linear programming). 6 representative cases are reported in Table 6.

Observation 4.2.3. The proposed tabu search algorithm, applied to n-element 2D posets, generates linear extensions with no more jumps than 130% of optimum, in n iterations.

The parameters of the tabu search in all the benchmarks were set as follows: $TabuSize = 10$, $CheckedNeighbours = 7$, $MaxDummies = 15$. These values have been established experimentally as a good compromise between the running time and convergence of the algorithm.

The running time of our tabu search algorithm could be improved by incorporating more involved methods to rebuild an arc diagram in every step. For instance, it was observed in [25] that in the case of interval orders, an arc diagram can be generated with Algorithm 5. This is much faster than Algorithm 1. We have tried this construction and it turned out that the running times reported in Section 4.1 would decrease by a factor of 3.

All the experiments have been performed on a 64-bit computer with Intel® Core™ i5-2500K CPU clocked at 3.30 GHz, and 24 GB of RAM. The algorithms have been implemented in the C# language for the .NET Framework 4. For linear programming, the `LPSolve` function from the `Optimization` package of Maple™ was employed.

4.3. An approximation ratio for two-dimensional posets

A way to measure the quality of approximation algorithms is to assess their approximation ratio.

Definition 4.3.1. An algorithm $\mathcal{A}$ is an ϵ-approximation algorithm (with $\epsilon > 1$) for a problem $\mathcal{P}$ if it runs in time polynomial in the input

size and always generates a solution of value APX $\geqslant \frac{\text{OPT}}{\epsilon}$ when $\mathcal{P}$ is a maximization problem, or of cost APX $\leqslant \epsilon$OPT when $\mathcal{P}$ is a minimization problem.

We now look again at the jump number problem on 2D orders via the Ceroi reduction, described in Lemma 4.2.2 of Section 4.2. The bump maximization problem can be solved by computing a maximum weight independent set of rectangles in $R(P) = (V_R, E_R)$. Some approximation algorithms are known for this problem.

Theorem 4.3.2 (Agarwal, van Kreveld, Suri [1])**.** *There exists a* $\log |V|$*-approximation algorithm for the maximum weight independent set problem on intersection graphs of rectangles.*

The original paper of Agarwal et al. focuses on maximum independent set problem, but their algorithm extends to the weighted variant in a straightforward way, see [11, pp. 136–140]. All logarithms are in base 2.

So this implies an approximation of the bump number. Can this result be used to approximate the jump number, too? We claim that the answer is positive.

Theorem 4.3.3. *There exists an* $(n/\log\log n)$*-approximation algorithm for the jump number of two-dimensional posets.*

In the proof, we use the following result.

Theorem 4.3.4 (McCartin [15])**.** *There exists an algorithm to decide for any poset* P *whether* $s(P) \leqslant h$*, running in time* $\mathcal{O}(h^2 h! n)$*.*

Proof of Theorem 4.3.3. The algorithm computes an approximate bump number $b_A(P) \geqslant \frac{b(P)}{\log |V|}$, where $|V|$ is the number of vertices in $G(R)$. We obtain an approximate jump number $s_A(P) = n - 1 - b_A(P)$, so the approximation ratio is

$$\frac{s_A(P)}{s(P)} \leqslant \frac{n - 1 - \frac{b(P)}{\log |V|}}{n - 1 - b(P)}.$$

We verify when this ratio is worse than $\frac{n}{\log\log n}$:

$$\frac{n - 1 - \frac{b(P)}{\log |V|}}{n - 1 - b(P)} > \frac{n}{\log\log n},$$

$$\log\log n \cdot (n - 1 - \frac{b(P)}{\log |V|}) > n \cdot s(P),$$

$$n \cdot s(P) < (n-1) \cdot \log\log n - \frac{\log\log n \cdot b(P)}{\log |V|} < n \cdot \log\log n,$$

$$s(P) < \log\log n.$$

When the jump number is very small, i.e., $s(P) < \log\log n$, the algorithm of Agarwal et al. [1] does not provide the claimed approximation of $s(P)$, but in such cases we may compute the jump number to optimality by the McCartin algorithm [15] in polynomial time, since one quickly verifies that $(\log\log n)! < n$. □

5. Conclusions and future work

In this paper a new tabu search algorithm for the jump number problem has been proposed and benchmarked. There are some remarks concerning the approximation, exact computation and hardness of the problem.

The results of our algorithm on interval orders have been compared with optimal values. The experiments reveal that in relatively short time linear extensions can be obtained with no more jumps than 105% of $s(P)$. It is easy to spot interval orders for which a 50% suboptimal semi-strongly greedy linear extension exists [12]. Hence, it is definitely worth to apply tabu search to improve the quality of generated solutions.

Previously, no approximation algorithms have been given for the class of two-dimensional posets. Our work is an attempt to fulfill this demand. In comparison with a lower bound on the jump number, linear extensions generated by our tabu search procedure turn out to be at most 30% suboptimal. Since the LP relaxation is usually inexact, the real error is probably lower. We have proved by other techniques that there is an $(n/\log\log n)$-approximation algorithm for the jump number problem on two-dimensional posets We hypothesize that even stronger approximation ratio could be proved for this class, perhaps by a detailed analysis of the semi-strongly greedy algorithm. Addressing this questions is our main objective in the future work. More precisely, it is conceivable that such an algorithm may be based on the rectangle intersection MWIS lower bound.

We have also empirically verified the number of all linear extensions generated by the exact algorithm `OptLinExt` of Sysło. It turns out that their number is far from the theoretical bound of $k!$. Considering that we have proved, based on the results of Mitas, that the complexity of computing $s(P)$ on interval orders is dominated by 2^n, and is lower in

practice, we believe that an estimation of the running time of `OptLinExt` can be improved in this class. Another question for further research is, whether or not the time complexity bound of this algorithm can be lowered on arbitrary posets.

The complexity status of computing $s(P)$ on two dimensional orders is open. Our experiments show that the space of semi-strongly greedy linear extensions is explicitly greater than in the case of interval orders. Hence, from this point of view, the class of two dimensional posets seems harder than interval orders for the jump number problem.

References

[1] P.K. Agarwal, M. van Kreveld, S. Suri, *Label placement by maximum independent set in rectangles*, Comput. Geom. 11 (1998), 209–218.

[2] A. Arnim, C. Higuera, *Computing the jump number on semi-orders is polynomial*, Discrete Appl. Math. **51**, 219–232 (1994).

[3] V. Bouchitté, M. Habib, *NP-completeness properties about linear extensions*, Order **4** (1987), 143–154.

[4] S. Ceroi, *A weighted version of the jump number problem on two-dimensional orders is NP-complete*, Order **20** (2003), 1–11.

[5] S. Felsner, *A 3/2-approximation algorithm for the jump number of interval orders*, Order **6** (1990), 325–334.

[6] P.C. Fishburn, *Interval orders and interval graphs. A study of partially ordered sets*, Wiley, New York, 1985.

[7] F. Glover, M. Laguna, *Tabu Search*, Kluwer Academic Publishers, 1997.

[8] M. Habib, *Comparability invariants*, in: M. Pouzet, D. Richard, ed., *Orders: Description and Roles*, Annals of Discrete Mathematics 23 (1984), pp. 371–386.

[9] M. Held, R.M. Karp, *A dynamic programming approach to sequencing problems*, J. Soc. Ind. Appl. Math. **10** (1962), 196–210.

[10] C. Higuera, L. Nourine, *Drawing and encoding two-dimensional posets*, Theor. Comput. Sci. **175** (1997), 293–308.

[11] C. Iturriaga, *Map labeling problems*, PhD Thesis, University of Waterloo, 1999.

[12] P. Krysztowiak, M.M. Sysło, *An experimental study of approximation algorithms for the jump number problem on interval orders*, Discrete Appl. Math., submitted (2012).

[13] P. Krysztowiak, *An improved approximation ratio for the jump number problem on interval orders*, Theor. Comput. Sci **513** (2013), 77–84.

[14] P. Krysztowiak, *Improved approximation algorithm for the jump number of interval orders*, Electron. Notes Discrete Math. **40** (2013), 193–198.

[15] C. McCartin, *An improved algorithm for the jump number problem*, Inform. Process. Lett. **79** (2001), 87–92.

[16] J. Mitas, *Tackling the jump number of interval orders*, Order **8** (1991), 115–132.

[17] A. Ngom, *Genetic algorithm for the jump number scheduling problem*, Order **15** (1998), 59–73.

[18] W.R. Pulleyblank, *On minimizing setups in precedence-constrained scheduling*, Report No. 81185 – OR (unpublished), May 1981.

[19] I. Rival, *Optimal linear extensions by interchanging chains*, Proc. Am. Math. Soc. **89** (1983), 387–394.

[20] J. Spinrad, J. Valdes, *Recognition and isomorphism of two dimensional partial orders*, Automata, Languages and Programming, LNCS 154 (1983), 676–686.

[21] M.M. Sysło, *A graph-theoretic approach to the jump-number problem*, in: I. Rival, ed., *Graphs and Orders*, D. Reidel, Dodrecht 1985, pp. 185–215.

[22] M.M. Sysło, *Minimizing the jump number for partially ordered sets: a graph-theoretic approach*, Order **1** (1984), 7–19.

[23] M.M. Sysło, *Minimizing the jump number for partially-ordered sets: a graph-theoretic approach, II*, Discrete Math. **63** (1987), 279–295.

[24] M.M. Sysło, *An algorithm for solving the jump number problem*, Discrete Math. **72** (1988), 337–346.

[25] M.M. Sysło, *The jump number problem on interval orders: A 3/2 approximation algorithm*, Discrete Math. **144** (1995), 119–130.

[26] M.H. El-Zahar, J.H. Schmerl, *On the size of jump-critical ordered sets*, Order **1** (1984), 3–5.

[27] M.H. El-Zahar, I. Rival, *Examples of jump-critical ordered sets*, SIAM J. Algebr. Discrete Meth. **6** (1985), 713–720.

[28] M.H. El-Zahar, *On jump-critical posets with jump-number equal to width*, Order **17** (2000), 93–101.

CONTACT INFORMATION

P. Krysztowiak, M. M. Sysło

Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, Toruń, Poland

E-Mail(*s*): pk@mat.umk.pl, syslo@mat.umk.pl

Received by the editors: 29.11.2013
and in final form 29.11.2013.

Algebra and Discrete Mathematics
Volume **20** (2015). Number 1, pp. 115–125

RESEARCH ARTICLE

Serial group rings of finite groups. General linear and close groups

Andrei Kukharev and Gena Puninski

Communicated by V. V. Kirichenko

Abstract. For a given p, we determine when the p-modular group ring of a group from $\mathrm{GL}(n,q)$, $\mathrm{SL}(n,q)$ and $\mathrm{PSL}(n,q)$-series is serial.

Introduction

There is a recent progress in classifying finite groups G whose group ring FG over a modular field F is serial. It is shown in [15] that the crucial point in this description is making a list of simple finite groups (and fields of finite characteristics) with this property.

For instance in [14] such a classification is given for symmetric and alternating groups; and [15] provides a list of sporadic simple groups and simple Suzuki groups with this property. Furthermore the first author described in [12] groups in the $\mathrm{PSL}(2,q)$-series whose modular group rings are serial.

In this paper we will continue this line of research by including into considerations all projective special linear groups $\mathrm{PSL}(n,q)$. Despite these groups are the main target of this paper, we have to make a bypass by considering general linear groups $\mathrm{GL}(n,q)$, and also special linear

The authors are grateful to Alexandre Zalesski for a helpful discussion. The research of the first author was supported by BRFFI grant F15RM-025.

2010 MSC: 20C05,20G40.

Key words and phrases: serial ring, group ring, general linear group, special linear group, projective special linear group.

groups $\mathrm{SL}(n,q)$. The reason for such a detour is that for general linear groups the structure of Brauer trees of blocks is best known, due to results of Fong and Srinivasan [8, 9]. Namely it is shown there that the Brauer tree of any block of $\mathrm{GL}(n,q)$ is an interval whose exceptional vertex is located at its end.

From general theory it is known (see [1, Sect. 5]) that a block B of a group algebra is serial if and only if its Brauer tree is a star with the exceptional vertex at the center. Thus in the case of the serial p-modular group ring of $\mathrm{GL}(n,q)$ we obtain that all Brauer trees of blocks are intervals with at most two edges and, if a tree has two edges, then the exceptional vertex should have multiplicity one. Furthermore the number of edges in a particular block can be calculated using centralizers and normalizers of defect subgroups. There are rather few cases which are left to analyze, which is achieved in this paper without difficulty.

In most cases descending from $\mathrm{GL}(n,q)$ to $\mathrm{SL}(n,q)$ and then to $\mathrm{PSL}(n,q)$ is a straightforward normal subgroup business, the only difficulty is when p divides $q-1$. In this case more groups with serial group rings occur, and our analysis is based on [12] or directly by looking at character tables.

There is no doubt that a similar approach applies to all classical groups but, because a myriad of details should be taken into account, we will postpone this to a future paper.

1. Preliminaries

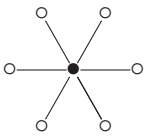
Recall that a module M over a ring R is said to be *uniserial*, if all submodules of M are linearly ordered by inclusion; and M is *serial* if it is a direct sum of uniserial modules. Furthermore R is called a *serial* ring, if R is serial as a right and left module over itself. It is known (see [2, Sect. 32]) that R is serial if and only if there exists a collection $e_1, \ldots, e_n$ of orthogonal idempotents such that each right module e_iR is serial, and the same is true for each left module Re_j. For a general theory of serial rings the reader is referred to [19] or recent [4]. Within the class of artinian algebras over a field, the serial rings are also known as *Nakayma algebras* - see [3, Sect. 4.2].

Let G be a finite group and let F be a field of finite characteristic p. If p does not divide the order of G then, by Maschke's theorem, the ring FG is semisimple artinian, hence serial. In this paper we will always assume that p *divides* $|G|$.

Let P denote a p-Sylow subgroup of G. Since (see [2, Theorem 32.3]) artinian serial rings are of finite representation type, it follows from Higman [10] that, if FG is serial, then P is a cyclic group. This gives a necessary condition for seriality, which is not always sufficient: for instance (see [1, p. 123]) the group $\mathrm{SL}(2,5)$ for $p = 5$ gives a counterexample.

Furthermore, the seriality of the group ring FG depends on characteristic of F only [6, 16]. Thus in this paper (to ease references) we will always assume that F *is algebraically closed*. For instance, it is known (see [18, 20] or [13]) that a p-modular group ring of a p-solvable group is serial.

We say that the Brauer tree of a block is a *star* if it has no path of length more than 2. Here is a typical shape of a star with the exceptional vertex in the center:



A useful criterion for checking seriality is given by the following.

Fact 1 (see [1, Sect. 5] or [7, Corollary VII.2.22])**.** A modular group ring $R = FG$ is serial if and only if for each block B of R its Brauer tree is a star whose exceptional vertex (if any) is located in the center.

Thus a satisfactory description of groups with serial group rings depends on the supply of information on Brauer trees of blocks, which is not always readily available.

In some cases the seriality can be lifted from normal subgroups. Suppose that B is a block of the group algebra FG; H is a normal subgroup of G and b is a block of FH. A definition of the notion that B *covers* b can be found in [1, Sect. 14]. For instance if H contains a p-Sylow subgroup of B, then the principal block B_0 of G covers the principal block b_0 of H.

Fact 2 (see [7, Theorem 6.2.7])**.** 1) Suppose that a block B of G covers a block b of H where H contains a defect group of B. Then B is serial if and only if b is serial.

2) Suppose that F is a field of characteristic p and let H be a normal subgroup of G whose index $|G/H|$ is coprime to p. Then the ring FG is serial if and only if FH is serial.

Suppose that B is a block of a modular group ring FG with a cyclic defect group D and let e denote the number of edges in the Brauer tree of B. For instance the defect group of the principal block B_0 equals P. By $C_G(D)$ we denote the centralizer of D in G; and $N_G(D)$ is the normalizer of D.

Fact 3 (see [1, Sect. 5, Theorem 1])**.** The number of edges e in the Brauer tree of a block B divides the order of the factor group $N_G(D)/C_G(D)$, hence divides $p-1$. Furthermore the multiplicity of the exceptional vertex equals $(|D|-1)/e$.

For the principal block B_0 the number of edges e equals to the order $|N_G(P)/C_G(P)|$.

We will need one more technical result. Recall that $O_{p'}$ denotes the largest normal subgroup of G consisting of elements whose order is coprime to p. We say that an element $g \in G$ is in the *kernel* of a block B if g acts trivially on every indecomposable projective module in B.

Fact 4 (see [7, Lemma IV.4.12])**.** The kernel of the principal block of G equals $O_{p'}$.

2. General linear group

In this section we will describe serial rings of general linear groups $\mathrm{GL}(n,q)$ over finite fields with q elements.

Theorem 1. *Let $G = \mathrm{GL}(n,q)$, $n \geqslant 2$ and let F be a field of characteristic p dividing the order of G. Then the group ring FG is serial if and only if one of the following holds.*

1) $n = 2$ and $p = q$ equal 2 or 3.

2) $n = 2, 3$, $p = 3$ and $q \equiv 2, 5 \pmod 9$.

For instance $\mathrm{GL}(3,2) \cong \mathrm{PSL}(2,7)$ and, for any field of characteristic 3, the group ring of this group is serial.

Recall that the order of $\mathrm{GL}(n,q)$ equals $q^{n(n-1)/2} \cdot (q-1) \cdot \ldots \cdot (q^n-1)$. Thus if p divides the order of G, then either $p \mid q$ or p divides $q^k - 1$ for some $k = 1, \ldots, n$.

We will divide the proof of Theorem 1 in two parts. The case of the defining characteristic $p \mid q$ is easy.

Lemma 1. *Let $q = p^r$, $G = \mathrm{GL}(n,q)$ and F is a field of characteristic p. The group ring FG is serial if and only if $n = 2$, $r = 1$ and p equals 2 or 3.*

Proof. If $n = 3$ then the matrices $\left(\begin{smallmatrix}1&1&0\\0&1&0\\0&0&1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}1&0&0\\0&1&0\\0&1&1\end{smallmatrix}\right)$ generate a subgroup $C_p \times C_p$, hence P is not cyclic; and we argue similarly for $n \geqslant 4$.

Thus it remains to consider the case $n = 2$.

If $r \geqslant 2$, it is easily checked that P is not cyclic, hence we may assume that $p = q$. Because $p - 1$, the index of $\mathrm{SL}(2, p)$ in $\mathrm{GL}(2, p)$, is coprime to p, it follows from Fact 2 that the seriality of group rings of $\mathrm{GL}(2, p)$ and $\mathrm{SL}(2, p)$ is equivalent.

If $p \geqslant 5$ we conclude from [1, p. 124] that the Brauer tree of the principal block B_0 of the group $H = \mathrm{SL}(2, p)$ is an interval with at least 3 edges, hence the ring FH (and then FG) is not serial. It remains to consider the case $p = 2, 3$.

If $p = 2$, then $G = \mathrm{GL}(2, 2) \cong S_3$ is 2-nilpotent, hence the ring FG is serial.

Similarly for $p = 3$ the group $\mathrm{GL}(2, 3)$ has order 48 and is 3-solvable, hence FG is serial. □

Thus we may assume that p does not divide q. Let d be the order of q modulo p, i.e. the least d such that $p \mid q^d - 1$. By the assumption we have $1 \leqslant d \leqslant n$, and clearly $d \mid p - 1$.

We will show that d cannot be very small (otherwise the p-Sylow subgroup P of G is not cyclic) and cannot be very large (otherwise the Brauer tree of the principal block has too many edges).

The description of normalizers and centralizers of p-Sylow subgroups of $\mathrm{GL}(n, q)$ is well known (see [21, 22]). We will add some explanations to ease reader's task.

Lemma 2. *1) P is cyclic if and only if $n < 2d$.*

2) If $n < 2d$ then the factor group $N_G(P)/C_G(P)$ has order d.

Proof. Consider the Galois field $\mathbb{F}_{q^d}$ as a vector space (of dimension d) over $\mathbb{F}_q$ with a basis $v_1, \ldots, v_d$. Let z be nonzero element of $\mathbb{F}_{q^d}$. Then $zv_i = \sum_j z_{ij} v_j$ for some $z_{ij} \in \mathbb{F}_q$. The mapping $z \mapsto (z_{ij})$ defines an embedding of the multiplicative group of $\mathbb{F}_{q^d}$ into $\mathrm{GL}(d, q)$. The image of a generator of $\mathbb{F}^*_{q^d}$ gives us a matrix $x \in \mathrm{GL}(d, q)$ of order $q^d - 1$.

Write $q^d - 1 = p^a \cdot s$ such that p and s are coprime, hence $y = x^s$ generates the p-Sylow subgroup P of order p^a.

1) If $n \geqslant 2d$, then one could insert in $\mathrm{GL}(n, q)$ two copies of $\mathrm{GL}(d, q)$ as 1 through d, and $d + 1$ through $2d$ diagonal blocks. It follows easily that P is not cyclic.

On the other hand, if $n < 2d$ then, comparing the sizes, we see that P can be chosen inside $\mathrm{GL}(d,q)$ embedded in the upper left 1 through d corner of $\mathrm{GL}(n,q)$, and therefore is generated by y.

2) It is known (see [22]) that the centralizer of P is generated by x, hence has order q^d-1. Furthermore (see [21, Lemma 4.6]) the normalizer of P is generated over $C_G(P)$ by an element of order d.

For our purposes it suffices to find an element which normalizes P and has order d modulo the centralizer. This can be achieved as follows.

Suppose that the action of x on the basis is given by a matrix $A=(a_{ij})$, $a_{ij}\in\mathbb{F}_q$: $xv_i=\sum_j a_{ij}v_j$. Applying the Frobenius morphism $x\mapsto x^q$ on $\mathbb{F}_{q^d}$ we obtain $x^q v_i^q=\sum_j a_{ij}v_j^q$. It follows that the action of x^q in the basis v_i^q is given by the same matrix A.

Because in the original basis this action is given by A^q, we conclude that $UAU^{-1}=A^q$, where U is the transition (from v_i^q to v_i) matrix. Then the conjugation by U defines an automorphism of order d on the subgroup generated by x. It follows that this action induces on P an automorphism ψ of the same order.

Namely, let $\psi(y)=y^q$ and suppose that $y^{q^k}=y$ for some k. Plugging $y=x^s$ we obtain $x^{(q^k-1)s}=1$, therefore $q^d-1=p^a\cdot s$ divides $(q^k-1)s$. It follows that p^a divides q^k-1, and hence d divides k, by the choice of d. □

Now we complete the proof of Theorem 1 by showing the following.

Proposition 1. *Let $G=\mathrm{GL}(n,q)$ and F is a field of characteristic p dividing the order of G but not dividing q. Then the group ring FG is serial if and only if $n=2,3$, $p=3$ and $q\equiv 2,5 \pmod 9$.*

Proof. We may assume that the p-Sylow subgroup P of G is cyclic. By the item 1) of Lemma 2 it follows that $n/2<d\leqslant n$, where d is the order of q modulo p.

Suppose first that $d>2$. If $p=2$ it follows (since p does not divide q) that q is odd, therefore p divides $q-1$ and $d=1$, a contradiction. Thus we may assume that $p>2$.

By Fact 3 and the item 2) of Lemma 2 the Brauer tree of the principal block B_0 of G has $e=d$ edges. Furthermore [8, Prop. 4] implies that this tree is an interval. Since $d>2$, this block is not serial.

Thus we are left with the case $d=2$, in particular p divides q^2-1. The definition of d yields that p does not divide $q-1$ and hence divides $q+1$. Again the Brauer tree of the principal block of G is an interval with $e=2$ edges, whose exceptional vertex is located at its end. By Fact 3 the multiplicity of this vertex is $(|P|-1)/2$. If $|P|>3$ this bock is not serial.

Thus we may assume that $|P| = 3$, which clearly yields $p = 3$ and $q \equiv 2, 5 \pmod 9$ (otherwise 9 divides the order of P). It follows that the principal block is serial.

We prove that, in this case, any non-principal block B of G is also serial. Namely, by Fact 3 the number of edges, e, of this block divides $p - 1 = 2$. If $e = 1$ then this block contains only one Brauer character, hence serial. If $e = 2$ then the multiplicity of the exceptional vertex equals $(3 - 1)/2 = 1$, hence this block is also serial. □

Note that in the proof of the implication $\Rightarrow$ in Theorem 1 we used only that the principal block B_0 of $\mathrm{GL}(n, q)$ is serial.

3. Special linear and projective special linear groups

In this section we will consider the seriality of group rings of special linear groups $\mathrm{SL}(n, q)$ and projective special linear groups $\mathrm{PSL}(n, q)$. The answer turns out to be the same for both series; and the proofs go in parallel.

Recall that $\mathrm{SL}(n, q)$ is a normal subgroup of $\mathrm{GL}(n, q)$ of index $q - 1$. Furthermore $\mathrm{PSL}(n, q)$ is obtained from $\mathrm{SL}(n, q)$ by factoring out the center Z whose order equals $(n, q - 1)$. Note also that, except of $\mathrm{PSL}(2, 2)$ and $\mathrm{PSL}(2, 3)$, $\mathrm{PSL}(n, q)$ is a simple group.

To avoid long sentences we will divide the classification theorem in two cases: when p divides $q - 1$ and when it is not. In the former case the answer is the same as in Theorem 1.

Proposition 2. *Let G is one of the groups $\mathrm{SL}(n, q)$ or $\mathrm{PSL}(n, q)$, $n \geqslant 2$. Let F be a field of characteristic p such that p does not divide $q - 1$. Then the ring FG is serial if and only if one of the following holds.*

1) $n = 2$ and $p = q$ equal 2 or 3.

2) $n = 2, 3$, $p = 3$ and $q \equiv 2, 5 \pmod 9$.

Proof. Since p does not divide $q - 1$, by Fact 2, we conclude that the seriality of group rings of $\mathrm{SL}(n, q)$ and $\mathrm{GL}(n, q)$ is equivalent. Applying Theorem 1 we obtain the desired conclusion for $\mathrm{SL}(n, q)$.

Thus we may assume that $G = \mathrm{PSL}(n, q)$. If 1) or 2) holds true then the group ring R of $\mathrm{SL}(n, q)$ is serial. Since G is a factor group of this group, it follows that the group ring of G is a factor ring of R, therefore is also serial.

Thus we may assume that the group ring of $\mathrm{PSL}(n, q)$ is serial and we need to show that either 1) or 2) holds true.

By Fact 4 the principal block b_0 of $\mathrm{SL}(n,q)$ has Z in its kernel, and therefore coincides with the principal block of $\mathrm{PSL}(n,q)$. Furthermore, because $\mathrm{SL}(n,q)$ contains the p-Sylow subgroup of $\mathrm{GL}(n,q)$ it follows by Fact 2 that the principal block B_0 of $\mathrm{GL}(n,q)$ is serial.

Now the result follows from the proof of Theorem 1 (see a remark at the end of Section 2). □

Now we consider the remaining case $p \mid q-1$. In this case serial rings occur more often than in the GL-case (cp. Theorem 1).

Proposition 3. *Let G be one of the group $\mathrm{SL}(n,q)$ or $\mathrm{PSL}(n,q)$, $n \geqslant 2$ and let F be a field of characteristic p dividing $q-1$. The group ring FG is serial if and only if $n=2$ and $p \neq 2$.*

Proof. If $n \geqslant 3$ then it is easily seen that p-Sylow subgroups of G are not cyclic. Thus we may assume that $n=2$.

If $G = \mathrm{PSL}(2,q)$ then FG is serial if and only if $p \neq 2$ [12].

Thus we may assume that $G = \mathrm{SL}(2,q)$. If $p=2$ then the group ring FG is not serial. Indeed, otherwise, being a factor ring of FG, the group ring of $\mathrm{PSL}(2,q)$ would be serial, a contradiction.

It remains to consider the case $p>2$ and we have to prove that the group ring of FG is serial. Observe that, if q is even, then $\mathrm{SL}(2,q) \cong \mathrm{PSL}(2,q)$, hence the ring is serial. Thus we assume that q is odd. In this case the center Z of $\mathrm{SL}(2,q)$ consists of matrices $\pm I$, where $I = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.

For the remaining part of the proof we need the character table of $G = \mathrm{SL}(2,q)$ — see Table 1.

In the table, $1 \leqslant l \leqslant (q-3)/2$, $1 \leqslant m \leqslant (q-1)/2$, $\varepsilon = (-1)^{(q-1)/2}$, ρ is a primitive $(q-1)$-th root of 1, and σ is a primitive $(q+1)$-th root of 1.

Let ν be a generator of the group $\mathbb{F}_q^*$. Denote $\gamma = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$, $\delta = \left(\begin{smallmatrix} 1 & 0 \\ \nu & 1 \end{smallmatrix}\right)$, $\alpha = \left(\begin{smallmatrix} \nu & 0 \\ 0 & \nu^{-1} \end{smallmatrix}\right)$. So, the order of α is $q-1$. The group G contains also an element β of order $q+1$. Moreover, two columns for the classes of $\gamma' = -I \cdot \gamma$ and $\delta' = -I \cdot \delta$ are omitted (to save space in the table). The values of any irreducible character χ of G on these classes are obtained by the formulas $\chi(\gamma') = \chi(\gamma)\chi(-I)/\chi(I)$ and $\chi(\delta') = \chi(\delta)\chi(-I)/\chi(I)$. Since $p \mid q-1$, only the sixth column of the table contain p-singular elements.

In particular, the cyclic group $\langle \alpha \rangle$ contains a generator y of a p-Sylow subgroup P of G.

It is easy to show (see [5, p. 230]) that $C_G(y) = \langle \alpha \rangle$ and $N_G(y) = \langle \alpha, \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) \rangle$. Hence $|N_G(P)/C_G(P)| = 2$. In particular, the number of

Classes	I	$-I$	γ	δ	α^l	β^m
Number of classes	1	1	1	1	$\frac{q-3}{2}$	$\frac{q-1}{2}$
Size of classes	1	1	$\frac{q^2-1}{2}$	$\frac{q^2-1}{2}$	$q(q+1)$	$q(q-1)$
1_G	1	1	1	1	1	1
ψ	q	q	0	0	1	-1
χ_i $(i = 1, \ldots, \frac{q-3}{2})$	$q+1$	$(-1)^i \times$ $\times(q+1)$	1	1	$\rho^{il} + \rho^{-il}$	0
$\theta_j (j = 1, \ldots, \frac{q-1}{2})$	$q-1$	$(-1)^j \times$ $\times(q-1)$	-1	-1	0	$-(\sigma^{jm} + \sigma^{jm})$
$\xi_1,\ \xi_2$	$\frac{q+1}{2}$	$\frac{\varepsilon(q+1)}{2}$	$\frac{1\pm\sqrt{\varepsilon q}}{2}$	$\frac{1\mp\sqrt{\varepsilon q}}{2}$	$(-1)^l$	0
$\eta_1,\ \eta_2$	$\frac{q-1}{2}$	$-\frac{\varepsilon(q-1)}{2}$	$\frac{-1\pm\sqrt{\varepsilon q}}{2}$	$\frac{-1\mp\sqrt{\varepsilon q}}{2}$	0	$(-1)^{m+1}$

TABLE 1. The character table of $\mathrm{SL}(2, q)$, q is odd [5, p. 228]

edges in the principal block B_0 of G equals 2, furthermore the number of edges in any block of G divides 2.

Observe that θ_j, η_1 and η_2 have value 0 on the class of α. By [17, Theorem 4.4.14], these characters belong to blocks of defect zero. It follows that these blocks contain only one irreducible ordinary character, hence is serial.

Furthermore it is easily checked (using [11, Theorem 2.1.8]) that the Steinberg character ψ belongs to the principal block B_0. Looking at the values on p-singular elements (and using cross-naught business — see [11, Chap. 2]) we see that the Brauer tree of B_0 is an interval with 2 edges having 1_G and ψ at its ends. Thus if there is an exceptional vertex it should be located at the center of this interval (in fact certain characters χ_i will occupy the center making an exceptional vertex there).

Because each character χ_i has the largest possible degree, it follows from [11, Lemma 2.1.22] that such a character cannot occur at the end of an interval of length 2. Thus the only possibility for such an interval is to have ξ_1 at one end, ξ_2 at another end, and some characters χ_i in between. But this block is clearly serial.

In fact such a block exists if $q \equiv 1 \pmod 4$; otherwise each non-principal block contains at most one modular character (i.e. its Brauer tree has at most one edge).

By this we have established that the group ring of $\mathrm{SL}(2, q)$ is serial if $2 \neq p \mid q - 1$, hence finished the proof of the proposition. □

Prepositions 2 and 3 completely describe groups of $\mathrm{SL}(n,q)$ and $\mathrm{PSL}(n,q)$-series whose p-modular group rings are serial.

References

[1] J.L. Alperin, *Local Representation Theory*, Cambridge University Press, 1981.

[2] F.W. Anderson, K.R. Fuller, *Rings and Categories of Modules*, 2d edition, Springer Graduate Texts in Math., Vol. **13**, 1992.

[3] M. Auslander, I. Reiten, S. Smalø, *Representation Theory of Artin Algebras*, Cambridge Studies in Advanced Mathematics, Vol. **36**, Cambridge, 1995.

[4] Y. Baba, K. Oshiro, *Classical Artinian Rings*, World Scient. Publ., 2009.

[5] L. Dornhoff, *Group representation theory. Part A: Ordinary Representation Theory* (Pure and applied mathematics 7), New York, 1971.

[6] D. Eisenbud, P. Griffith, *Serial rings*, J. Algebra, **17** (1971), pp. 389–400.

[7] W. Feit, *The Representation Theory of Finite Groups*, North Holland Mathematical Library, Vol. **25**, 1982.

[8] P. Fong, B. Srinivasan, *Blocks with cyclic defect groups in* $\mathrm{GL}(n,q)$, Bull. Amer. Math. Soc., **3** (1980), pp. 1041–1044.

[9] P. Fong, B. Srinivasan, *Brauer trees in* $\mathrm{GL}(n,q)$, Math. Z., **187** (1984), pp. 81–88.

[10] D.G. Higman, *Indecomposable representations at characteristic p*, Duke J. Math., **21** (1954), pp. 377–381.

[11] G. Hiss, K. Lux, *Brauer Trees of Sporadic Groups*, Oxford, 1989.

[12] A. Kukharev, *Seriality of group rings of unimodular projective groups*, in: Proc. of the 71th Scientific Conf. of Students and PhD students of Belarusian State University, Minsk, May 18–21, 2014, Part 1, pp. 11–14.

[13] A. Kukharev, G. Puninski, *Serial group rings of finite groups. p-solvability*, Algebra Discrete Math., **16** (2013), pp. 201–216.

[14] A. Kukharev, G. Puninski, *The seriality of group rings of alternating and symmetric groups*, Vestnik of Belarusian State University, Mathemathics and Informatics series, **2** (2014), pp. 61–64.

[15] A. Kukharev, G. Puninski, *Serial group rings of finite groups. Sporadic simple groups and Suzuki groups*, Notes Research Semin. Steklov Institute Sanct-Petersb., **435** (2015), pp. 73–94.

[16] A. Kukharev, G. Puninski, Yu. Volkov, *The seriality of the group ring of a finite group depends only of characteristic of the field*, Notes Research Semin. Steklov Institute Sanct-Petersb., **423** (2014), pp. 57–66.

[17] K. Lux, H. Pahlings, *Representations of Groups: a Computational Approach*, Cambrodge Studies in dvanced Mathmeatics, Vol. **124**, 2010.

[18] K. Morita, *On group rings over a modular field which possess radicals expressible as principal ideal*, Sci. Repts. Tokyo Daigaku, **4** (1951), pp. 177–194.

[19] G. Puninski, *Serial Rings*, Kluwer, 2001.

[20] B. Srinivasan, *On the indecomposable representations of a certain class of groups*, Proc. Lond. Math. Soc., **10** (1960), pp. 497–513.

[21] M. Stather, *Constructive Sylow theorems for the classical groups*, J. Algebra, **316** (2007), pp. 536–559.

[22] A.J. Weir, *Sylow p-subgroups of the classical groups over finite fields with chacteristic prime to p*, Proc. Amer. Math. Soc. **6** (1955), pp. 529–533.

Contact information

A. Kukharev,
G. Puninski

Faculty of Mechanics and Mathematics,
Belarusian State University, 4,
Nezavisimosti Ave., Minsk, 220030, Belarus
E-Mail(*s*): kukharev.av@mail.ru,
punins@mail.ru
Web-page(*s*): www.mmf.bsu.by

Received by the editors: 11.05.2015
and in final form 12.07.2015.

Algebra and Discrete Mathematics
Volume **20** (2015). Number 1, pp. 126–141

RESEARCH ARTICLE

Lattice groups

L. A. Kurdachenko, V. S. Yashchuk, I. Ya. Subbotin

Dedicated to Professor Efim Zelmanov in occasion of his 60th birthday

Abstract. In this paper, we introduce some algebraic structure associated with groups and lattices. This structure is a semigroup and it appeared as the result of our new approach to the fuzzy groups and L-fuzzy groups where L is a lattice. This approach allows us to employ more convenient language of algebraic structures instead of currently accepted language of functions.

The purpose of this work is to look with a somewhat different angle at algebraic structures related to the functions defined on a group. For every subset M of a set S there exists its characteristic function, that is the mapping $\chi_M \colon S \to \{0,1\}$ such that $\chi_M(y) = 1$ for all $y \in M$ and $\chi_M(y) = 0$ for all $y \notin M$. In many commonly used cases, a subset of M is identified with its characteristic function. In 1965, L.A. Zadeh [6] based on his generalization of the characteristic function introduced the fuzzy mathematics. Thus, a fuzzy set on a set S is a sort of generalized "characteristics function" on S, for whose "degrees of membership" we can use more diverse set than simple {yes, no}. In fact, we can consider the set L of degrees of membership. In the optimization problems, L may express the degree of optimality of the choice (in S); in the classification problems, it may express the degree of membership in a pattern class; in other contexts other terminologies appear. In fuzzy mathematics, a habitual step was to review the situation when $L = [0,1]$ is the usual closed interval of real numbers with its natural order. The following

2010 MSC: 20M10, 06D99, 20M99.
Key words and phrases: group, lattice, distributive lattice, fuzzy group, semigroup.

interpretation justifies this approach: we can consider a value of the generalized characteristic function as a probability of the fact that the given element belongs to the given subset. In this way, the algebraic fuzzy structures were constructed as follows. With every algebraic structure A, a corresponding fuzzy structure which characterized by a specific functions of A on $[0,1]$ associated with this conventional algebraic structure A, was connected (see, for example, [4]). For instance, in fuzzy group theory the objects of study are the functions $\gamma\colon G \to [0,1]$, G is a group, satisfying the following conditions:

$\gamma(xy) \geqslant \gamma(x) \wedge \gamma(y)$ for all $x, y \in G$; and $\gamma(x^{-1}) \geqslant \gamma(x)$ for every $x \in G$

(see, for example, [5] S 1.2). Some generalizations have appeared immediately. More concretely, considerations of the function $\gamma\colon G \to \mathfrak{L}$ where $\mathfrak{L}$ is a distributive lattice [1] were initiated. The theory of fuzzy groups was developed quite rapidly. However it was upswing in breadth rather than depth development. A variety of results obtained there was not planned properly. Even in the book [5], there were no attempts to systematize these results. A large array of results on fuzzy groups just has been collected in this book with no proper arrangement. In the L-fuzzy groups, regardless of the most common results, there was no serious progress at all.

Perhaps the key obstacle here is in the interpretation of an algebraic structure as a function, which is not very convenient most of the time. Because of that, very often the function γ is interpreted as an all point function $\chi(g, \gamma(g)), g \in G$. Here $\chi(g, \mathfrak{a})$ is a function such that $\chi(g, \mathfrak{a})(g) = \mathfrak{a}$, $\chi(g, \mathfrak{a})(y) = 0$ whenever $y \neq g$. However, in some cases we need to consider the function γ as an union of all point functions $\chi(g, \mathfrak{a})$ for all $g \in G$ and $\mathfrak{a} \leqslant \gamma(g)$ (see, for example, [2], [3]). Actually speaking, the point functions $\chi(g, \mathfrak{a})$ play here the role of elements, formally the subfunctions of γ, so that each time it is necessary to implement keep in mind some special reservations.

In the current article we offer the interpretation of L-fuzzy groups as sets with operations. With this algebraic approach, the basic concepts and results of algebraic nature acquire its natural form, and the process of their appearance becomes more meaningful. We present the basic concepts of the theory of L-fuzzy groups, as well as the results in the form in which they are needed to be for our transformation. The resulting structure is formally different, and therefore the term for it to be used is different. In the article we are concerned only with the basic concepts, but nevertheless, our approach will make it possible to see the general structural picture.

As for the term L-fuzzy group, it seems it does not reflect the essence of the case, so we will use the term group function. We do not seek maximize generality, it seems more natural to consider the case, when lattice L is distributive and finite, although the obtained results can be extended on the case of an arbitrary complete distributive lattice.

Let $\mathfrak{L}$ be a lattice and G be a group. To avoid misunderstandings, the identity element of G is denoted by e. We will consider a set $\mathfrak{L}^G$ of all functions $\lambda\colon G \to \mathfrak{L}$. On this set we define the operations $\wedge$ and $\vee$ by the following rules: if $\lambda, \mu \in \mathfrak{L}^G$, then put

$$(\lambda\wedge\mu)(x) = \lambda(x)\wedge\mu(x) \quad \text{and} \quad (\lambda\vee\mu)(x) = \lambda(x)\vee\mu(x) \quad \text{for each } x \in G.$$

Clearly the operations $\wedge$ and $\vee$ are commutative and associative,

$$(\lambda \wedge (\lambda \vee \mu))(x) = \lambda(x) \wedge (\lambda \vee \mu)(x) = \lambda(x) \wedge (\lambda(x) \vee \mu(x)) = \lambda(x)$$

and

$$(\lambda \vee (\lambda \wedge \mu))(x) = \lambda(x) \vee (\lambda \wedge \mu)(x) = \lambda(x) \vee (\lambda(x) \wedge \mu(x)) = \lambda(x),$$

so that $\lambda \wedge (\lambda \vee \mu) = \lambda$ and $\lambda \vee (\lambda \wedge \mu) = \lambda$. Clearly $\lambda \wedge \lambda = \lambda$ and $\lambda \vee \lambda = \lambda$. Hence a set $\mathfrak{L}^G$ is a lattice.

If $a, b \in \mathfrak{L}$, then $a\vee b = b$ is equivalent to $a \leqslant b$. Therefore we can define an order on $\mathfrak{L}^G$: for $\lambda, \mu \in \mathfrak{L}^G$ will put $\lambda \leqslant \mu$ if and only if $\lambda(x) \leqslant \mu(x)$ for each element $x \in G$.

Suppose now that a lattice $\mathfrak{L}$ is *distributive and finite*. Being finite, it has the greatest element $\mathfrak{m}$ and the least element $\mathfrak{o}$. For every function $f\colon G \to \mathfrak{L}$ define $\mathrm{Supp}(f)$ as a subset of all elements $x \in G$ such that $f(x) \neq \mathfrak{o}$.

Let Y be a subset of G and $\mathfrak{a} \in \mathfrak{L}$. We define the function $\chi(Y, \mathfrak{a})$ as follows:

$$\chi(Y, \mathfrak{a}) = \begin{cases} \mathfrak{a}, & \text{if } x \in Y \\ \mathfrak{o}, & \text{if } x \notin Y. \end{cases}$$

If $Y = \{y\}$, then instead of $\chi(\{y\}, \mathfrak{a})$ we will write $\chi(y, \mathfrak{a})$. The function $\chi(y, \mathfrak{a})$ is called the *point function* or shorter the *point*. By its definition, $\chi(y, \mathfrak{a}) \in \mathfrak{L}^G$. Furthermore, let $f \in \mathfrak{L}^G$. If $\mathrm{Supp}(f) = \{g_1, \ldots, g_n\}$ is finite and $f(g_j) = \mathfrak{a}_\mathfrak{j}$, $1 \leqslant j \leqslant n$, then clearly $f = \chi(g_1, \mathfrak{a}_\mathfrak{1}) \vee \ldots \vee \chi(g_n, \mathfrak{a}_\mathfrak{n})$.

Define now the binary operation $\odot$ on $\mathfrak{L}^G$ by the following rule. Let $\mu, \nu \in \mathfrak{L}^G$, and x be an arbitrary element of a group G. Consider the subset of the lattice $\mathfrak{L}$

$$\{\mu(y) \wedge \nu(z) | u, v \text{ are the elements of } G \text{ such that } yz = x\}.$$

Since $\mathfrak{L}$ is finite, this subset really is finite. Therefore we can define about its least upper bound. Put

$$(\mu \odot \nu)(x) = \vee_{y,z\in G, yz=x}(\mu(y) \wedge \nu(z)).$$

We remark that

$$(\mu \odot \nu)(x) = \vee_{y\in G}(\mu(y) \wedge \nu(y^{-1}x)) = \vee_{z\in G}(\mu(xz^{-1}) \wedge \nu(z)).$$

Consider now some basic properties of this product.

Proposition 1. *The following assertions hold:*

(i) *The operation $\odot$ is associative.*

(ii) *The function $\chi(e, \mathfrak{m})$ is an identity element of the operation $\odot$.*

(iii) $\lambda \odot (\mu \vee \nu) = (\lambda \odot \mu) \vee (\lambda \odot \nu)$ *and* $(\mu \vee \nu) \odot \lambda = (\mu \odot \lambda) \vee (\nu \odot \lambda)$ *for all functions $\lambda, \mu, \nu \in \mathfrak{L}^G$.*

(iv) *If $x, y \in G, \mathfrak{a} \in \mathfrak{L}$, then $(\chi(y, \mathfrak{a}) \odot \lambda)(x) = \mathfrak{a} \wedge \lambda(y^{-1}x)$; in particular, if $\mathfrak{a} = \vee_{x\in G}\lambda(x)$, then $((\chi(y, \mathfrak{a}) \odot \lambda)(x) = \lambda(y^{-1}x)$.*

(v) $(\lambda \odot (\chi(y, \mathfrak{a})))(x) = \mathfrak{a} \wedge \lambda(xy^{-1})$*; in particular, if* $\mathfrak{a} = \vee_{x\in G}\lambda(x)$, *then* $(\lambda \odot \chi(y, \mathfrak{a}))(x) = \lambda(xy^{-1})$.

(vi) *if $x, y, u \in G$, $\mathfrak{a}, \mathfrak{b} \in \mathfrak{L}$ then $(\chi(y, \mathfrak{a}) \odot \chi(u, \mathfrak{b}))(yu) = \mathfrak{a} \wedge \mathfrak{b}$ and $(\chi(y, \mathfrak{a}) \odot \chi(u, \mathfrak{b}))(x) = \mathfrak{o}$ if $x \neq yu$. In other words, $\chi(y, \mathfrak{a}) \odot \chi(u, \mathfrak{b}) = \chi(yu, \mathfrak{a} \wedge \mathfrak{b})$; in particular, $\chi(y, \mathfrak{a}) \odot \chi(u, \mathfrak{a}) = \chi(yu, \mathfrak{a})$.*

(vii) $(\chi(x, \mathfrak{a}) \odot \lambda \odot \chi(x^{-1}, \mathfrak{a}))(y) = \mathfrak{a} \wedge \lambda(x^{-1}yx)$.

Proof. (i) Let $\lambda, \mu, \nu \in \mathfrak{L}^G$. Put $\kappa = \lambda \odot \mu$ and $\eta = \mu \odot \nu$. We have

$$\begin{aligned}((\lambda \odot \mu) \odot \nu)(x) &= (\kappa \odot \nu)(x) = \vee_{y,z\in G, yz=x}(\kappa(y) \wedge \nu(z)) \\ &= \vee_{y,z\in G, yz=x}(\vee_{u,v\in G, uv=y}(\lambda(u) \wedge \mu(v)) \wedge \nu(z)) \\ &= \vee_{u,v,z\in G, uvz=x}((\lambda(u) \wedge \mu(v)) \wedge \nu(z)).\end{aligned}$$

$$\begin{aligned}(\lambda \odot (\mu \odot \nu))(x) &= (\lambda \odot \eta)(x) = \vee_{u,y\in G, uy=x}(\lambda(u) \wedge \eta(y)) \\ &= \vee_{u,y\in G, uy=x}(\lambda(u) \wedge (\vee_{v,z\in G, vz=y}(\mu(v) \wedge \nu(z)))) \\ &= \vee_{u,v,z\in G, uvz=x}(\lambda(u) \wedge (\mu(v) \wedge \nu(z))).\end{aligned}$$

Since $(\lambda(u) \wedge \mu(v)) \wedge \nu(z) = \lambda(u) \wedge (\mu(v) \wedge \nu(z))$ for all $u, v, z \in G$,

$$((\lambda \odot \mu) \odot \nu)(x) = (\lambda \odot (\mu \odot \nu))(x)$$

for each $x \in G$. It implies that $(\lambda \odot \mu) \odot \nu = \lambda \odot (\mu \odot \nu)$.

(ii) Let $\lambda \in \mathfrak{L}^G$ and consider the product $\lambda \odot \chi(e, \mathfrak{m})$. By its definition, $(\chi(e, \mathfrak{m}))(e) = \mathfrak{m}$ and $(\chi(e, \mathfrak{m}))(x) = \mathfrak{o}$ whenever $x \neq 1$. We have now

$$\begin{aligned} &\lambda(x) \wedge (\chi(e, \mathfrak{m}))(e) = \lambda(x) \wedge \mathfrak{m} = \lambda(x) \\ &\text{and } \lambda(y) \wedge (\chi(e, \mathfrak{m}))(z) = \mathfrak{o} \quad \text{if } z \neq 1, \end{aligned}$$

so that

$$\begin{aligned} (\lambda \odot \chi(e, \mathfrak{m}))(e) &= \vee_{y,z \in G, yz=1}(\lambda(y) \wedge \chi(e, \mathfrak{m})(z)) \\ &= \lambda(e) \wedge \chi(e, \mathfrak{m})(e) = \lambda(e), \\ (\lambda \odot \chi(e, \mathfrak{m}))(x) &= \vee_{y,z \in G, yz=x}(\lambda(y) \wedge \chi(e, \mathfrak{m})(z)) \\ &= \lambda(x) \wedge \chi(e, \mathfrak{m})(e) = \lambda(x). \end{aligned}$$

Since it is valid for all $x \in G$, $\lambda \odot \chi(e, \mathfrak{m}) = \lambda$. In a similar way we can prove that $\chi(e, \mathfrak{m}) \odot \lambda = \lambda$.

(iii) We have

$$\begin{aligned} \lambda \odot (\mu \vee \nu)(x) &= \vee_{y \in G}(\lambda(y) \wedge ((\mu \vee \nu)(y^{-1}x))) \\ &= \vee_{y \in G}(\lambda(y) \wedge (\mu(y^{-1}x) \vee \nu(y^{-1}x))) \\ &= \vee_{y \in G}(\lambda(y) \wedge \mu(y^{-1}x)) \vee (\lambda(y) \wedge \nu(y^{-1}x)) \\ &= (\vee_{y \in G}(\lambda(y) \wedge \mu(y^{-1}x))) \vee (\vee_{y \in G}(\lambda(y) \wedge \nu(y^{-1}x))) \\ &= (\lambda \odot \mu)(x) \vee (\lambda \odot \nu)(x) \\ &= ((\lambda \odot \mu) \vee (\lambda \odot \nu))(x). \end{aligned}$$

It proves that

$$\lambda \odot (\mu \vee \nu) = (\lambda \odot \mu) \vee (\lambda \odot \nu).$$

Using similar arguments, we obtain that and

$$(\mu \vee \nu) \odot \lambda = (\mu \odot \lambda) \vee (\nu \odot \lambda).$$

(iv) Let x be an arbitrary element of G. If $z \neq y$, then $\chi(y, \mathfrak{a})(z) = \mathfrak{o}$, so we have

$$\begin{aligned} (\chi(y, \mathfrak{a}) \odot \lambda)(x) &= \vee_{z \in G}(\chi(y, \mathfrak{a})(z) \wedge \lambda(z^{-1}x)) \\ &= \chi(y, \mathfrak{a})(y) \wedge \lambda(y^{-1}x)) = \mathfrak{a} \wedge \lambda(y^{-1}x). \end{aligned}$$

The proof of (v) is similar.

(vi) If $u \in G$, $\mathfrak{b} \in \mathfrak{L}$, then $(\chi(y, \mathfrak{a}) \odot \chi(u, \mathfrak{b}))(x) = \mathfrak{a} \wedge \chi(u, \mathfrak{b})(y^{-1}x)$. Recall that $\chi(u, \mathfrak{b})(y^{-1}x) = \mathfrak{b}$ if $y^{-1}x = u$ or $x = yu$ and $\chi(u, \mathfrak{b})(y^{-1}x) = \mathfrak{o}$ if $y^{-1}x \neq u$ or $x \neq yu$. Thus

$$(\chi(y, \mathfrak{a}) \odot \chi(u, \mathfrak{b}))(x) = \begin{cases} \mathfrak{a} \wedge \mathfrak{b}, & \text{if } x = yu \\ \mathfrak{o}, & \text{if } x \neq yu. \end{cases}$$

Hence we obtain (vi).

(vi) Using the above arguments we obtain

$$\begin{aligned} &(\chi(x, \mathfrak{a}) \odot (\gamma \odot \chi(x^{-1}, \mathfrak{a})))(y) \\ &\quad = \vee_{u,v,z \in G, uvz = y} \chi(x, \mathfrak{a})(u) \wedge (\gamma(v) \wedge \chi(x^{-1}, \mathfrak{a}))(z) \\ &\quad = \chi(x, \mathfrak{a})(x) \wedge \gamma(x^{-1}yx) \wedge \chi(x^{-1}, \mathfrak{a})(x^{-1}) \\ &\quad = \mathfrak{a} \wedge \gamma(x^{-1}yx) \wedge \mathfrak{a} = \mathfrak{a} \wedge \gamma(x^{-1}yx). \end{aligned}$$ □

Let G be a group and $\gamma \in \mathfrak{L}^G$. Then a surjective function γ is said to be a *group function on* G if it satisfies the following conditions:

(GF 1) $\gamma(xy) \geqslant \gamma(x) \wedge \gamma(y)$ for all $x, y \in G$,

(GF 2) $\gamma(x^{-1}) \geqslant \gamma(x)$ for every $x \in G$.

Let γ, κ group functions on G. If $\gamma \leqslant \kappa$, then we will say that γ *is a subgroup function of* κ. This fact we will denote $\gamma \preceq \kappa$.

Proposition 2. *Let G be a group, $\mathfrak{L}$ be a finite distributive lattice, $\gamma \in \mathfrak{L}^G$, and suppose that γ is a group function on G. Then the following assertions hold:*

(i) $\gamma(x^{-1}) = \gamma(x)$ *for every $x \in G$ (in order words, a function γ is even).*

(ii) $\gamma(xy^{-1}) \geqslant \gamma(x) \wedge \gamma(y)$ *for all $x, y \in G$.*

(iii) $\gamma(x^n) \geqslant \gamma(x)$ *for every $x \in G$ and every integer n.*

(iv) $\gamma(e) \geqslant \gamma(x)$ *for every $x \in G$.*

(v) *Let λ, $\kappa \leqslant \gamma$, then $\lambda \odot \kappa \leqslant \gamma$, in particular, $\gamma \odot \gamma \leqslant \gamma$.*

Proof. (i) We have $x = (x^{-1})^{-1}$, so (GF 2) implies that $\gamma(x) \geqslant \gamma(x^{-1})$, which together with $\gamma(x^{-1}) \geqslant \gamma(x)$ gives $\gamma(x) = \gamma(x^{-1})$ for every element $x \in G$.

(ii) Let x, y be arbitrary elements of G. By (GF 1) $\gamma(xy^{-1}) \geqslant \gamma(x) \wedge \gamma(y^{-1})$, and by (i) $\gamma(y^{-1}) = \gamma(y)$, so that $\gamma(xy^{-1}) \geqslant \gamma(x) \wedge \gamma(y)$.

(iii) Let $x \in G$. By (GF 1) $\gamma(x^2) = \gamma(xx) \geqslant \gamma(x) \wedge \gamma(y) = \gamma(x)$. Using ordinary induction, we obtain that $\gamma(x^n) \geqslant \gamma(x)$ for every $n \in \mathbb{N}$. Suppose now that $n = -k$ where $k \in \mathbb{N}$. Then $x^n = (x^{-1})^k$. By proved above

$$\gamma(x^n) = \gamma((x^{-1})^k) \geqslant \gamma(x^{-1}) = \gamma(x).$$

(iv) Let $x \in G$. By (GF 1) we have

$$\gamma(e) = \gamma(xx^{-1}) \geqslant \gamma(x) \wedge \gamma(x^{-1}) = \gamma(x) \wedge \gamma(x) = \gamma(x).$$

(v) Let x be an arbitrary element of G. The inclusions λ, $\kappa \leqslant \gamma$ imply $\lambda(y) \wedge \kappa(z) \leqslant \gamma(y) \wedge \gamma(z)$. Since γ is a group function, $\gamma(y) \wedge \gamma(z) \leqslant \gamma(yz)$, thus

$$(\lambda \odot \kappa)(x) = \vee_{y,z \in G, yz=x}(\gamma(y) \wedge \kappa(z)) \leqslant \vee_{y,z \in G, yz=x}\gamma(yz) = \gamma(x). \quad \square$$

Proposition 3 (A criterion of group function)**.** *Let G be a group, $\mathfrak{L}$ be a finite distributive lattice and $\gamma \in \mathfrak{L}^G$. Then γ is a group function on G if and only if the following assertions hold:*

(GF 3) $\chi(x, \gamma(x)) \odot \chi(y, \gamma(y)) \subseteq \gamma$ *for all* $x, y \in G$.

(GF 4) $\chi(x^{-1}, \gamma(x)) \subseteq \gamma$ *for every* $x \in G$.

Proof. Suppose first that γ is a group function. Clearly $\chi(x, \gamma(x)) \subseteq \gamma$ and $\chi(y, \gamma(y)) \subseteq \gamma$ for all elements $x, y \in G$. Using Proposition 2 (v) we obtain that

$$\chi(x, \gamma(x)) \odot \chi(y, \gamma(y)) \subseteq \gamma.$$

Let x be an arbitrary element of G. We have $(\chi(x^{-1}, \gamma(x)))(x^{-1}) = \gamma(x)$. Since γ is a group function, $\gamma(x) \leqslant \gamma(x^{-1})$. We note that if $y \neq x^{-1}$, then $(\chi(x, \gamma(x)))(y) = \mathfrak{o}$, so that $(\chi(x^{-1}), \gamma(x)))(y) \leqslant \gamma(y)$ for every $y \in G$. This means that $\chi(x^{-1}, \gamma(x)) \subseteq \gamma$.

Conversely, suppose that γ satisfies both conditions (GF 3) and (GF 4). Let x, y be arbitrary elements of G. Then (GF 3) shows that $\chi(x, \gamma(x)) \odot \chi(y, \gamma(y)) \subseteq \gamma$. By Proposition 1 (vi),

$$(\chi(x, \gamma(x)) \odot \chi(y, \gamma(y)))(xy) = \gamma(x) \wedge \gamma(y).$$

The inclusion $\chi(x, \gamma(x)) \odot \chi(y, \gamma(y)) \subseteq \gamma$ implies that $(\chi(x, \gamma(x)) \odot \chi(y, \gamma(y)))(xy) \leqslant \gamma(xy)$, thus we obtain $\gamma(x) \wedge \gamma(y) \leqslant \gamma(xy)$, and γ satisfies (GF 1).

Let $x \in G$. Since $\chi(x^{-1}, \gamma(x)) \subseteq \gamma$, $(\chi(x^{-1}, \gamma(x)))(y) \leqslant \gamma(y)$ for every $y \in G$. In particular, $(\chi(x^{-1}, \gamma(x)))(x^{-1}) = \gamma(x) \leqslant \gamma(x^{-1})$, so that γ satisfies (GF 2). $\square$

Let G be a group and $\mathfrak{L}$ be a finite distributive lattice. Consider the Cartesian product $A = G \times \mathfrak{L}$. Define the operation (multiplication) on A by the following rule: $(u, \mathfrak{a})(v, \mathfrak{b}) = (uv, \mathfrak{a} \wedge \mathfrak{b})$ for all $u, v \in G$, $\mathfrak{a}, \mathfrak{b} \in \mathfrak{L}$. This operation is associative because multiplication in G and the operation $\wedge$ in $\mathfrak{L}$ are associative. The pair $(e, \mathfrak{m})$ is the identity element for this operation. The obtained above criterion allows us to transform the definition of the group function in the following.

A nonempty subset Λ of $G \times \mathfrak{L}$ is called a *lattice group over* $\mathfrak{L}$ if it satisfies the following conditions:

(LG 1) if $(x, \mathfrak{a}) \in \Lambda$ and $\mathfrak{b} \leqslant \mathfrak{a}$, then $(x, \mathfrak{b}) \in \Lambda$;

(LG 2) if $(x, \mathfrak{a})$, $(y, \mathfrak{b}) \in \Lambda$, then $(x, \mathfrak{a})(y, \mathfrak{b}) \in \Lambda$;

(LG 3) if $(x, \mathfrak{a}) \in \Lambda$, then $(x^{-1}, \mathfrak{a}) \in \Lambda$.

For every element $x \in \mathrm{pr}_G(\Lambda)$ put $\mathfrak{C}_\Lambda(x) = \{\mathfrak{a} \in \mathfrak{L} | (x, \mathfrak{a}) \in \Lambda\}$.

Observe at once that a lattice group Λ defines a group function on G. Indeed, for every element $x \in \mathrm{pr}_G(\Lambda)$ the set $\mathfrak{C}_\Lambda(x)$ is not empty. Put $\lambda(x) = \vee \mathfrak{C}_\Lambda(x)$. If $x \notin \mathrm{pr}_G(\Lambda)$, then put $\lambda(x) = \mathfrak{o}$. Then λ is a function. If $u, v \in G$ and $\lambda(u) = \mathfrak{a}$, $\lambda(v) = \mathfrak{b}$, then $(uv, \mathfrak{a} \wedge \mathfrak{b}) \in \Lambda$ by condition (LG 2). It follows that $\lambda(uv) \geqslant \mathfrak{a} \wedge \mathfrak{b} = \lambda(u) \wedge \lambda(v)$, so that λ satisfies (GF 1). Similarly, let $\lambda(u) = \mathfrak{a}$, then $(u^{-1}, \mathfrak{a}) \in \Lambda$ by condition (LG 3). It follows that $\lambda(u^{-1}) \geqslant \mathfrak{a} = \lambda(u)$, so that λ satisfies (GF 2).

Let Λ, Γ be the lattice groups over $\mathfrak{L}$. If Λ includes Γ, then we will say that Γ is a *lattice subgroup* of Λ, and will denote this by $\Gamma \leqslant \Lambda$.

If γ is a defined by Γ group function, then $\gamma \preceq \lambda$.

Clearly $G \times \mathfrak{L}$ is the greatest lattice group over $\mathfrak{L}$, and $E = \{(e, \mathfrak{o})\}$ is the least lattice group over $\mathfrak{L}$; the last lattice group is called *trivial*. Furthermore, if $\mathfrak{a} \in \mathfrak{L}$, then $\{(e, \mathfrak{b}) | \mathfrak{b} \leqslant \mathfrak{a}\}$ is a lattice group over $\mathfrak{L}$.

Every lattice group Λ includes $\mathrm{pr}_G(\Lambda) \times \{\mathfrak{o}\}$. For every subgroup H of G the subset $H \times \{\mathfrak{o}\}$ is a lattice group. Recall that a subset $\mathfrak{M}$ of $\mathfrak{L}$ is called a *lower* (respectively *upper*) *segment* of $\mathfrak{L}$, if from $\mathfrak{a} \in \mathfrak{M}$ and $\mathfrak{b} \leqslant \mathfrak{a}$ (respectively $\mathfrak{a} \leqslant \mathfrak{b}$) it follows that $\mathfrak{b} \in \mathfrak{M}$.

If $\mathfrak{a} \in \mathfrak{L}$, then the subset $\{\mathfrak{x} | \mathfrak{x} \in \mathfrak{L}$ and $\mathfrak{x} \leqslant \mathfrak{a}\}$ (respectively $\{\mathfrak{x} | \mathfrak{x} \in \mathfrak{L}$ and $\mathfrak{x} \geqslant \mathfrak{a}\}$) is a lower segment (respectively upper segment) of $\mathfrak{L}$. It called the *principal lower* (respectively *upper*) *segment* of $\mathfrak{L}$ generated by $\mathfrak{a}$.

Consider some preliminary properties of the lattice groups.

Proposition 4. *Let G be a group, $\mathfrak{L}$ be a finite distributive lattice and $\mathfrak{S}$ be a family of lattice subgroups over $\mathfrak{L}$. Then intersection $\cap\mathfrak{S}$ is a lattice subgroup.*

Proof. The proof is almost obvious. □

Proposition 5. *Let G be a group, $\mathfrak{L}$ be a finite distributive lattice and Λ a lattice group. Then:*

(i) $\mathrm{pr}_{\mathfrak{L}}(\Lambda)$ *is a semigroup by operation* $\wedge$ *with identity* $\mathfrak{e}(\Lambda) = \vee\mathfrak{C}_{\Lambda}(1)$ *and zero* $\mathfrak{o}$. *Moreover,* $\mathrm{pr}_{\mathfrak{L}}(\Lambda)$ *is the principal lower segment of* $\mathfrak{L}$, *generated by* $\mathfrak{e}(\Lambda)$.

(ii) $\mathrm{pr}_G(\Lambda)$ *is a subgroup of* G. *Conversely, for every subgroup* H *of* $\mathrm{pr}_G(\Lambda)$ *the subset* $\{(x,\mathfrak{a})|(x,\mathfrak{a}) \in \Lambda$ *and* $x \in H\} = \mathrm{pr}_G^{-1}(H)$ *is a lattice subgroup of* Λ.

(iii) *If* $\mathfrak{M}$ *is a lower segment of* $\mathfrak{L}$, *then* $\{(x,\mathfrak{a})|(x,\mathfrak{a}) \in \Lambda$ *and* $\mathfrak{a} \in \mathfrak{M}\}$ *is a lattice subgroup of* Λ. *In particular,* $\mathrm{pr}_{\mathfrak{L}}^{-1}(\mathfrak{M})$ *is a lattice group.*

Proof. (i) Indeed, if $\mathfrak{a}, \mathfrak{b} \in \mathrm{pr}_{\mathfrak{L}}(\Lambda)$, then there are elements $u, v \in G$ such that $(u,\mathfrak{a}), (v,\mathfrak{b}) \in \Lambda$. Since Λ is a lattice group, $(uv, \mathfrak{a}\wedge\mathfrak{b}) = (u,\mathfrak{a})(v,\mathfrak{b}) \in \Lambda$. It follows that $\mathfrak{a} \wedge \mathfrak{b} \in \mathrm{pr}_{\mathfrak{L}}(\Lambda)$. In particular, $\mathfrak{e}(\Lambda) = \vee\mathfrak{C}_{\Lambda}(e) \in \mathrm{pr}_{\mathfrak{L}}(\Lambda)$.

Let $\mathfrak{a} \in \mathrm{pr}_{\mathfrak{L}}(\Lambda)$ and u be an element of G such that $(u,\mathfrak{a}) \in \Lambda$. Since Λ is a lattice group, $(u^{-1},\mathfrak{a}) \in \Lambda$ by condition (LG 3). Using (LG 2), we obtain that $(e,\mathfrak{a}) = (uu^{-1},\mathfrak{a}) = (uu^{-1}, \mathfrak{a}\wedge\mathfrak{a}) = (u,\mathfrak{a})(u^{-1},\mathfrak{a}) \in \Lambda$. Hence $\mathfrak{a} \in \mathfrak{C}(e)$, which follows that $\mathfrak{a} \leqslant \mathfrak{e}(\Lambda)$. In other words, $\mathfrak{e}(\Lambda)$ is the greatest element of $\mathrm{pr}_{\mathfrak{L}}(\Lambda)$.

Let $\mathfrak{c}$ be an arbitrary element of $\mathfrak{L}$ such that $\mathfrak{c} \leqslant \mathfrak{e}(\Lambda)$. Since $(e,\mathfrak{e}(\Lambda)) \in \Lambda$. $(e,\mathfrak{c}) \in \Lambda$ by condition (LG 1). It follows that $\mathrm{pr}_{\mathfrak{L}}(\Lambda)$ is the principal lower segment of $\mathfrak{L}$, generated by $\mathfrak{e}(\Lambda)$.

(ii) Let $K = \mathrm{pr}_G(\Lambda), u, v \in K$. Then there are the elements $\mathfrak{a}, \mathfrak{b} \in \mathfrak{L}$ such that $(u,\mathfrak{a}), (v,\mathfrak{b}) \in \Lambda$. Since Λ is a lattice group, $(uv, \mathfrak{a}\wedge\mathfrak{b}) = (u,\mathfrak{a})(v,\mathfrak{b}) \in \Lambda$. It follows that $uv \in K$. If $(u,\mathfrak{a}) \in \Lambda$, then $(u^{-1},\mathfrak{a}) \in \Lambda$ by condition (LG 3), which follows that $u^{-1} \in K$. Hence K is a subgroup of G.

Let now H be a subgroup of $\mathrm{pr}_G(\Lambda)$, $(u,\mathfrak{a}), (v,\mathfrak{b}) \in \mathrm{pr}_G^{-1}(H)$. Since Λ is a lattice group, $(uv, \mathfrak{a}\wedge\mathfrak{b}) = (u,\mathfrak{a})(v,\mathfrak{b}) \in \Lambda$. The fact that H is a subgroup implies that $uv \in H$, so that $(uv, \mathfrak{a}\wedge\mathfrak{b}) \in \mathrm{pr}_G^{-1}(H)$. Since H is a subgroup, then from $u \in H$ it follows that $u^{-1} \in H$. Since Λ is a lattice group, $(u,\mathfrak{a}) \in \Lambda$ implies that $(u^{-1},\mathfrak{a}) \in \Lambda$. Hence $(u^{-1},\mathfrak{a}) \in \mathrm{pr}_G^{-1}(H)$, so that $\mathrm{pr}_G^{-1}(H)$ satisfies the conditions (LG 2), (LG 3), and $(uv, \mathfrak{a}\wedge\mathfrak{b}) = (u,\mathfrak{a})(v,\mathfrak{b}) \in \Lambda$. Hence K is a subgroup of G. Let $(u,\mathfrak{a}) \in \mathrm{pr}_G^{-1}(H)$ and $\mathfrak{b}$ be an element of $\mathfrak{L}$ such that $\mathfrak{b} \leqslant \mathfrak{a}$. Then $(u,\mathfrak{b}) \in \Lambda$ and hence $(u,\mathfrak{b}) \in \mathrm{pr}_G^{-1}(H)$.

(iii) Let $\mathfrak{M}$ is a lower segment of $\mathfrak{L}$, K a subgroup of G and $M = K \times \mathfrak{M}$. Then M is a lattice group. Indeed, if $(x,\mathfrak{a}) \in M$ and $\mathfrak{b} \leqslant \mathfrak{a}$, then $\mathfrak{b} \in \mathfrak{M}$,

because $\mathfrak{M}$ is a lower segment of $\mathfrak{L}$. It follows that $(x, \mathfrak{b}) \in M$, so that M satisfies (LG 1). Suppose that $(x, \mathfrak{a}), (y, \mathfrak{b}) \in M$. Since $\mathfrak{a} \wedge \mathfrak{b} \leqslant \mathfrak{b}, \mathfrak{a} \wedge \mathfrak{b} \in \mathfrak{M}$. The fact that K is a subgroup of G implies $xy \in K$, and hence $(xy, \mathfrak{a} \wedge \mathfrak{b}) \in M$. We note that $(xy, \mathfrak{a} \wedge \mathfrak{b}) = (x, \mathfrak{a})(y, \mathfrak{b})$, which shows that M satisfies (LG 2). Finally, let $(x, \mathfrak{a}) \in M$. Since K is a subgroup of G, $x^{-1} \in K$. Therefore $(x^{-1}, \mathfrak{a}) \in M$, and M satisfies (LG 3).

Let again $H = \mathrm{pr}_G(\Lambda)$, then it is not hard to see that

$$\{(x, \mathfrak{a}) | (x, \mathfrak{a}) \in \Lambda \text{ and } \mathfrak{a} \in \mathfrak{M}\} = H \times \mathfrak{M} \cap \Lambda.$$

Proposition 4 shows that this subset is a lattice subgroup of Λ. □

Let Λ be a lattice group. Unlike abstract groups, a lattice group can contains more than one idempotent. Moreover, Λ contains a pair $(1, \mathfrak{a})$ for each element $\mathfrak{a} \in \mathrm{pr}_{\mathfrak{L}}(\Lambda)$. Indeed, let u be an element of G such that $(u, \mathfrak{a}) \in \Lambda$. Since Λ is a lattice group, $(u, \mathfrak{a})(u^{-1}, \mathfrak{a}) \in \Lambda$. But $(u, \mathfrak{a})(u^{-1}, \mathfrak{a}) = (e, \mathfrak{a} \wedge \mathfrak{a}) = (e, \mathfrak{a})$. It shows that a semigroup Λ can be a group only in the case when $\mathrm{pr}_{\mathfrak{L}}(\Lambda)$ contains only one element $\mathfrak{a}$. Let $\mathfrak{b} \in \Lambda$ and $\mathfrak{b} \leqslant \mathfrak{a}$, then condition (LG 1) implies that $(u, \mathfrak{b}) \in \Lambda$. Hence $\mathfrak{a} = \mathfrak{b}$. In other words, $\mathfrak{a}$ is the least element of $\mathfrak{L}$, i.e. $\mathfrak{a} = \mathfrak{o}$. Consequently, a lattice group Λ is a group if and only if $\mathrm{pr}_{\mathfrak{L}}(\Lambda) = \{\mathfrak{o}\}$. In this regard, we note that the semigroup Λ may include many subsemigroups, which are groups by multiplication. Indeed, let H be a subgroup of G and $\mathfrak{a} \in \mathfrak{L}$, then it is not hard to see that the subset $H \times \{\mathfrak{a}\}$ is a group by multiplication. Furthermore, for every $\mathfrak{a} \in \mathfrak{L}$ the subset $\{(u, \mathfrak{a}) | (u, \mathfrak{a}) \in \Lambda\}$ is also a group by multiplication.

If Λ is a lattice subgroup over $\mathfrak{L}$, then put $E(\Lambda) = \{(e, \mathfrak{b}) | \mathfrak{b} \leqslant \mathfrak{e}(\Lambda)\}$. Clearly $E(\Lambda)$ is a lattice subgroup of Λ.

Let Γ be a lattice subgroup of Λ. The pair $(e, \mathfrak{e}(\Lambda))$ is an identity element of Λ and $(e, \mathfrak{e}(\Gamma))$ is an identity element of Γ. Since $\Gamma \leqslant \Lambda$, Proposition 5 shows that $\mathfrak{e}(\Gamma) \leqslant \mathfrak{e}(\Lambda)$. We say that Γ is an *unitary lattice subgroup* of Λ, if $(e, \mathfrak{e}(\Lambda)) \in \Gamma$. Every lattice subgroup of Λ can be extended to an unitary lattice subgroup. Indeed, put $\Gamma^{u(\Lambda)} = \Gamma \cup \{(e, \mathfrak{b}) | \mathfrak{b} \leqslant \mathfrak{e}(\Lambda)\} = \Gamma \cup E(\Lambda)$, then $\Gamma^{u(\Lambda)}$ is a lattice group. In fact, if $(u, \mathfrak{a}) \in \Lambda$, then $(u, \mathfrak{a})(e, \mathfrak{b}) = (u, \mathfrak{a} \wedge \mathfrak{b})$. Since $\mathfrak{a} \wedge \mathfrak{b} \leqslant \mathfrak{a}$, $(u, \mathfrak{a} \wedge \mathfrak{b}) \in \Gamma$. It shows that $\Gamma^{u(\Lambda)}$ satisfies all conditions (LG 1)–(LG 3).

Let M be a subset of $G \times \mathfrak{L}$ and $\mathfrak{S}$ be a family of all lattice groups, including M. By Proposition 4, the intersection $\cap\mathfrak{S}$ is a lattice group. It called the lattice group generated by M and will be denoted by $\langle M \rangle$.

Let $(x, \mathfrak{a}) \in G \times \mathfrak{L}$. If Λ is a lattice group containing $(x, \mathfrak{a})$, then it is not hard to prove that $(x, \mathfrak{a})^n = (x^n, \mathfrak{a} \wedge \ldots \wedge \mathfrak{a}) = (x^n, \mathfrak{a}) \in \Lambda$ for each positive

integer n. By (LG 3), $(x^{-1}, \mathfrak{a}) \in \Lambda$, and hence $(e, \mathfrak{a}) = (x, \mathfrak{a})(x^{-1}, \mathfrak{a}) \in \Lambda$. From $(x^{-1}, \mathfrak{a}) \in \Lambda$ we obtain that $(x, \mathfrak{a})^{-n} = (x^{-n}, \mathfrak{a}) \in \Lambda$, so that $\{(x^n, \mathfrak{a}) | n \in \mathbb{Z}\} \subseteq \Lambda$. Let $\mathfrak{A}$ be the principal lower segment of $\mathfrak{L}$, generated by $\mathfrak{a}$. If $\mathfrak{b} \leqslant \mathfrak{a}$, then (LG 1) implies that $(x^n, \mathfrak{b}) \in \Lambda$ for each integer n. Thus $\{(x^n, \mathfrak{b}) | \mathfrak{b} \leqslant \mathfrak{a}, n \in \mathbb{Z}\} \subseteq \Lambda$. It is not hard to check that the subset $\{(x^n, \mathfrak{b}) | \mathfrak{b} \leqslant \mathfrak{a}, n \in \mathbb{Z}\}$ is a lattice group. It follows that $\langle (x, \mathfrak{a}) \rangle = \{(x^n, \mathfrak{b}) | \mathfrak{b} \leqslant \mathfrak{a}, n \in \mathbb{Z}\}$.

Let Λ, Γ be the lattice subgroups. Define its product in the usual way: put

$$\Lambda\Gamma = \{(x, \mathfrak{a})(y, \mathfrak{b}) = (xy, \mathfrak{a} \wedge \mathfrak{b}) | (x, \mathfrak{a}) \in \Lambda, (y, \mathfrak{b}) \in \Gamma\}.$$

The following result is a rationale for this determination.

Proposition 6. *Let G be a group, $\mathfrak{L}$ be a finite distributive lattice and $\gamma, \kappa \colon G \to \mathfrak{L}$ be functions. Then*

$$\gamma \odot \kappa = \cup_{y \in \mathrm{Supp}(\gamma), z \in \mathrm{Supp}(\kappa)} \chi(y, \gamma(y)) \odot \chi(z, \kappa(z)).$$

Proof. By definition we have

$$(\gamma \odot \kappa)(x) = \vee_{y,z \in G, yz = x} (\gamma(y) \wedge \kappa(z)).$$

If $y \notin \mathrm{Supp}(\gamma)$, then $\gamma(y) = \mathfrak{o}$ and $\gamma(y) \wedge \kappa(z) = \mathfrak{o}$. Similarly, if $z \notin \mathrm{Supp}(\kappa)$, then $\kappa(z) = \mathfrak{o}$, and again $\gamma(y) \wedge \kappa(z) = \mathfrak{o}$. It follows that

$$(\gamma \odot \kappa)(x) = \vee_{y \in \mathrm{Supp}(\gamma), z \in \mathrm{Supp}(\kappa), yz = x} (\gamma(y) \wedge \kappa(z)).$$

On the other hand, let $\xi = \cup_{y \in \mathrm{Supp}(\gamma), z \in \mathrm{Supp}(\kappa)} \chi(y, \gamma(y)) \odot \chi(z, \kappa(z))$. By Proposition 1, $\chi(y, \gamma(y)) \odot \chi(z, \kappa(z)) = \chi(yz, (\gamma(y) \wedge \kappa(z)))$. If $x \in G$ and $x = yz$, then $\chi(yz, (\gamma(y) \wedge \kappa(z)))(x) = \gamma(y) \wedge \kappa(z)$, otherwise $\chi(yz, (\gamma(y) \wedge \kappa(z)))(x) = \mathfrak{o}$. Therefore

$$\begin{aligned} \xi(x) &= \vee_{y \in \mathrm{Supp}(\gamma), z \in \mathrm{Supp}(\kappa)} (\chi(yz, (\gamma(y) \wedge \kappa(z))))(x) \\ &= \vee_{y \in \mathrm{Supp}(\gamma), z \in \mathrm{Supp}(\kappa), yz = x} (\gamma(y) \wedge \kappa(z)) = (\gamma \odot \kappa)(x). \end{aligned}$$

Since it is true for each $x \in G$,

$$\gamma \odot \kappa = \cup_{y \in \mathrm{Supp}(\gamma), z \in \mathrm{Supp}(\kappa)} \chi(y, \gamma(y)) \odot \chi(z, \kappa(z)). \qquad \square$$

Corollary. *Let G be a group, $\mathfrak{L}$ be a finite distributive lattice, $\mathfrak{a} \in \mathfrak{L}$, and $\kappa \colon G \to \mathfrak{L}$ be functions. Then for every $x \in G$*

$$\chi(x, \mathfrak{a}) \odot \kappa = \cup_{z \in \mathrm{Supp}(\kappa)} \chi(x, \mathfrak{a}) \odot \chi(z, \kappa(z)),$$

$$\kappa \odot \chi(x, \mathfrak{a}) = \cup_{z \in \mathrm{Supp}(\kappa)} \chi(z, \kappa(z)) \odot \chi(x, \mathfrak{a}).$$

Let $\lambda\colon G \to \mathfrak{L}$ be a function defined by Λ and $\gamma\colon G \to \mathfrak{L}$ be a function defined by Γ. Consider a function $\kappa\colon G \to \mathfrak{L}$ defined by the product $\Lambda\Gamma$. Let g be an arbitrary element of G. If $g \notin \mathrm{pr}_G(\Lambda\Gamma)$, then $\kappa(g) = \mathfrak{o}$. On the other hand, let u, v be an arbitrary elements of G such that $g = uv$. Since $g \notin \mathrm{pr}_G(\Lambda\Gamma) = \mathrm{pr}_G(\Lambda)\,\mathrm{pr}_G(\Gamma)$, then either $u \notin \mathrm{pr}_G(\Lambda)$, $v \notin \mathrm{pr}_G(\Gamma)$, or $u \in \mathrm{pr}_G(\Lambda)$ but $v \notin \mathrm{pr}_G(\Gamma)$ or $u \notin \mathrm{pr}_G(\Lambda)$ but $v \in \mathrm{pr}_G(\Gamma)$. In each of these cases either $\lambda(u) = \mathfrak{o}$ or $\gamma(v) = \mathfrak{o}$, so that

$$\vee_{u,v\in G, uv=g}(\lambda(u) \wedge \gamma(v)) = \mathfrak{o} = \kappa(g).$$

Suppose now that $g \in \mathrm{pr}_G(\Lambda\Gamma)$, then $\kappa(g) = \vee\mathfrak{C}_{\Lambda\Gamma}(g)$. Let again u, v be arbitrary elements of G such that $g = uv$. If $u \notin \mathrm{pr}_G(\Lambda)$ or $v \notin \mathrm{pr}_G(\Gamma)$, then $(\lambda(u)\wedge\gamma(v)) = \mathfrak{o}$. Suppose that $u \in \mathrm{pr}_G(\Lambda)$ and $v \in \mathrm{pr}_G(\Gamma)$ and let $\mathfrak{a}$, $\mathfrak{b}$ be the elements of $\mathfrak{L}$ such that $(u, \mathfrak{a}), (v, \mathfrak{b}) \in \mathfrak{L}$. We have $(u, \mathfrak{a})(v, \mathfrak{b}) = (uv, \mathfrak{a} \wedge \mathfrak{b})$. This shows that $\mathfrak{C}_{\Lambda\Gamma}(g) = \{\mathfrak{a} \wedge \mathfrak{b} | \mathfrak{a} \in \mathfrak{C}_\Lambda(u), \mathfrak{b} \in \mathfrak{C}_\Gamma(v)\}$. Since $\lambda(u) = \vee\mathfrak{C}_\Lambda(u)$, $\gamma(v) = \vee\mathfrak{C}_\Gamma(v)$, $\mathfrak{C}_{\Lambda\Gamma}(g) = \lambda(u) \wedge \gamma(v)$. In other words, in this case we have also

$$\kappa(g) = \vee_{u,v\in G, uv=g}(\lambda(u) \wedge \gamma(v)).$$

Thus $\kappa = \lambda \odot \gamma$. Thus, from the bulky and not very transparent product of functions we come to the intuitively clear and convenient product of subsets.

Let us now see how another important concept, the concept of normal fuzzy subgroup can be transformed. Again, it should be recalled that we use different terminology.

Let $\lambda, \kappa\colon G \to \mathfrak{L}$ be a group functions and $\kappa \preceq \lambda$. We say that κ *is a normal subgroup function of* λ, if $\kappa(yxy^{-1}) \geqslant \kappa(x) \wedge \lambda(y)$ for every elements $x, y \in G$.

We will need the following criteria of normality.

Proposition 7. *Let G be a group, $\mathfrak{L}$ be a finite distributive lattice and $\lambda, \kappa\colon G \to \mathfrak{L}$ be group functions such that $\kappa \preceq \gamma$. Then the following assertions are equivalent:*

(i) *κ is a normal subgroup function of γ;*

(ii) *$\chi(x, \gamma(x)) \odot \kappa \odot \chi(x^{-1}, \gamma(x)) \preceq \kappa$ for every element $x \in G$;*

(iii) *$\chi(x, \gamma(x))\odot\chi(y, \kappa(y))\odot\chi(x^{-1}, \gamma(x)) \subseteq \kappa$ for every elements $x, y \in G$;*

(iv) *$\chi(x, \mathfrak{a}) \odot \chi(y, \mathfrak{b}) \odot \chi(x^{-1}, \mathfrak{a}) \subseteq \kappa$ for every elements $x, y \in G$, $\mathfrak{a} \leqslant \gamma(x)$, $\mathfrak{b} \leqslant \kappa(y)$.*

Proof. (i) $\Rightarrow$ (ii). Suppose that κ is a normal subgroup function of λ. For arbitrary element $y \in G$ we consider the product $\chi(y, \gamma(y)) \odot \kappa \odot \chi(y^{-1}, \gamma(y))$. Let x be an arbitrary element of G. From Proposition 1 we obtain

$$(\chi(y, \gamma(y)) \odot \kappa \odot \chi(y^{-1}, \gamma(y)))(x) = \gamma(y) \wedge \kappa(y^{-1}xy).$$

Put $u = y^{-1}xy$, then $x = y(y^{-1}xy)y^{-1} = yuy^{-1}$, so that

$$(\chi(y, \gamma(y)) \odot \kappa \odot \chi(y^{-1}, \gamma(y)))(yuy^{-1}) = \gamma(y) \wedge \kappa(u).$$

Since $\kappa(u) \wedge \gamma(y) \leqslant \kappa(yuy^{-1})$, we obtain

$$(\chi(y, \gamma(y)) \odot \kappa \odot \chi(y^{-1}, \gamma(y)))(yuy^{-1}) \leqslant \kappa(yuy^{-1}),$$

that is

$$(\chi(y, \gamma(y)) \odot \kappa \odot \chi(y^{-1}, \gamma(y)))(x) \leqslant \kappa(x).$$

Since this is valid for every element $x \in G$,

$$\chi(y, \gamma(y)) \odot \kappa \odot \chi(y^{-1}, \gamma(y)) \preceq \kappa.$$

(ii) $\Rightarrow$ (iii). Indeed, Corollary to Proposition 6 shows that

$$\begin{aligned}&\chi(y, \gamma(y)) \odot \kappa \odot \chi(y^{-1}, \gamma(y)) \\ &\quad = \cup_{z \in \mathrm{Supp}(\kappa)} \chi(y, \gamma(y)) \odot \chi(z, \kappa(z)) \odot \chi(y^{-1}, \gamma(y)).\end{aligned}$$

Hence the inclusion $\chi(y, \gamma(y)) \odot \kappa \odot \chi(y^{-1}, \gamma(y)) \preceq \kappa$ implies that

$$\chi(y, \gamma(y)) \odot \chi(z, \kappa(z)) \odot \chi(y^{-1}, \gamma(y)) \subseteq \kappa \quad \text{for every elements } y, z \in G.$$

(iii) $\Rightarrow$ (iv). Indeed, Proposition 1 shows that

$$\chi(x, \gamma(x)) \odot \chi(y, \kappa(y)) \odot \chi(x^{-1}, \gamma(x)) = \chi(xyx^{-1}, \gamma(x) \wedge \kappa(y)).$$

We have

$$\chi(x, \mathfrak{a}) \odot \chi(y, \mathfrak{b}) \odot \chi(x^{-1}, \mathfrak{a}) = \chi(xyx^{-1}, \mathfrak{a} \wedge \mathfrak{b}) \subseteq \chi(xyx^{-1}, \gamma(x) \wedge \kappa(y)).$$

(iv) $\Rightarrow$ (i). Using again Proposition 1, we obtain that

$$\chi(x, \gamma(x)) \odot \chi(y, \kappa(y)) \odot \chi(x^{-1}, \gamma(x)) = \chi(xyx^{-1}, \gamma(x) \wedge \kappa(y)).$$

Now (vi) shows that $\chi(xyx^{-1}, \gamma(x) \wedge \kappa(y)) \subseteq \kappa$. Then

$$\gamma(x) \wedge \kappa(y) = \chi(xyx^{-1}, \gamma(x) \wedge \kappa(y))(xyx^{-1}) \leqslant \kappa(xyx^{-1}).$$

This means that κ is a normal subgroup function of γ. □

Proposition 7 leads us to the following analogue of normality in lattice groups.

Let Γ be a lattice subgroup of Λ. We say that Γ is a *normal lattice subgroup* of Λ, if $(y^{-1}, \mathfrak{b})(x, \mathfrak{a})(y, \mathfrak{b}) \in \Gamma$ for all pairs $(y, \mathfrak{b}) \in \Lambda$, $(x, \mathfrak{a}) \in \Gamma$.

We remark that $(y^{-1}, \mathfrak{b})(x, \mathfrak{a})(y, \mathfrak{b}) = (y^{-1}xy, \mathfrak{a}\wedge\mathfrak{b})$. At once this shows that if Γ a normal lattice subgroup of Λ, then $\mathrm{pr}_G(\Gamma)$ is a normal subgroup of $\mathrm{pr}_G(\Lambda)$. Conversely, suppose that H is a normal subgroup of G and $\Lambda_H = \{(x, \mathfrak{a}) \in \Lambda | x \in H\}$. Then $(y^{-1}, \mathfrak{b})(x, \mathfrak{a})(y, \mathfrak{b}) = (y^{-1}xy, \mathfrak{a} \wedge \mathfrak{b}) \in \Lambda$ for each pair $(y, \mathfrak{b}) \in \Lambda$. Since H is normal in G, $y^{-1}xy \in H$, so that $(y^{-1}, \mathfrak{b})(x, \mathfrak{a})(y, \mathfrak{b}) \in \Lambda_H$.

Let $\mathfrak{M}$ be a lower segment of $\mathfrak{L}$. Then Proposition 5 proves that $\Lambda[\mathfrak{M}] = \{(x, \mathfrak{a}) | (x, \mathfrak{a}) \in \Lambda$ and $\mathfrak{a} \in \mathfrak{M}\}$ is a lattice subgroup of Λ. $\Lambda[\mathfrak{M}]$ is called an $\mathfrak{M}$-*layer* of Λ. We note that $\Lambda[\mathfrak{M}]$ is a normal lattice subgroup of Λ. In fact, let $(x, \mathfrak{a}) \in \Lambda[\mathfrak{M}]$ and $(y, \mathfrak{b}) \in \Lambda$, then $(y^{-1}, \mathfrak{b})(x, \mathfrak{a})(y, \mathfrak{b}) = (y^{-1}xy, \mathfrak{a} \wedge \mathfrak{b})$. Since $\mathfrak{a} \wedge \mathfrak{b} \leqslant \mathfrak{a}, \mathfrak{a} \in \mathfrak{M}$ and $\mathfrak{M}$ is a lower segment of $\mathfrak{L}$, $\mathfrak{a} \wedge \mathfrak{b} \in \mathfrak{M}$. Thus $(y^{-1}, \mathfrak{b})(x, \mathfrak{a})(y, \mathfrak{b}) \in \Lambda[\mathfrak{M}]$.

If Γ is a normal lattice subgroup of Λ, then $\Gamma^{u(\Lambda)}$ is a normal lattice subgroup of Λ. Indeed, let $(x, \mathfrak{a}) \in \Gamma^{u(\Lambda)}$ and $(y, \mathfrak{b}) \in \Lambda$. If $x \neq e$, then $(x, \mathfrak{a}) \in \Gamma$ and $(y^{-1}, \mathfrak{b})(x, \mathfrak{a})(y, \mathfrak{b}) \in \Gamma$. If $x = e$, then $(y^{-1}, \mathfrak{b})(e, \mathfrak{a})(y, \mathfrak{b}) = (1, \mathfrak{a} \wedge \mathfrak{b}) \in E(\Lambda)$.

The layers of lattice group play a very important role. Especially it is useful in the case when $\mathrm{pr}_{\mathfrak{L}}(\Lambda)$ is a chain. This case arises in theory of fuzzy group when a group G is finite. Suppose that $|\mathrm{pr}_{\mathfrak{L}}(\Lambda)| = k$. Then $\mathrm{pr}_{\mathfrak{L}}(\Lambda)$ is isomorphic (as an ordered set) to the set $\mathrm{Ch}[1, k] = \{1, 2, \ldots, k\}$ with the natural ordering $1 \leqslant 2 \leqslant \ldots \leqslant k$. In this case, we will say that Λ is a lattice group over $\mathrm{Ch}[1, k]$.

For this case we construct some natural series of subgroups both in the lattice group Λ and in $\mathrm{pr}_G(\Lambda)$. The subset $\{1\}$ is the lower segment of $\mathrm{Ch}[1, k]$, and therefore the $\{1\}$-layer $\Lambda[1]$ of Λ is a lattice subgroup of Λ. If $(u, m) \in \Lambda$, then $(u, 1) \in \Lambda$ by condition (LG 1). This implies that $\mathrm{pr}_G(\Lambda) = \mathrm{pr}_G(\Lambda[1])$. For every m, $1 \leqslant m \leqslant k$, the subset $K_m = \{(u, m) | (u, m) \in \Lambda\}$ is the subgroup by multiplication, so that $\mathrm{H}(m) = \mathrm{pr}_G(K_m)$ is a subgroup of $\mathrm{H}(1) = \mathrm{pr}_G(\Lambda)$. A subgroup $\mathrm{H}(m)$ is called the m-*hoop* of Λ. From $(u, m) \in \Lambda$ we obtain $(u, m - 1) \in \Lambda$ by condition (LG 1). This implies the inclusion $\mathrm{H}(m) \leqslant \mathrm{H}(m - 1)$, so we obtain the following descending series of subgroups

$$\mathrm{H}(1) \geqslant \mathrm{H}(2) \geqslant \ldots \geqslant \mathrm{H}(k).$$

Clearly the mapping $u \to (u, m)$, $u \in \mathrm{H}(m)$, is an isomorphism of $\mathrm{H}(m)$ on K_m for each m, $1 \leqslant m \leqslant k$.

Figuratively speaking, the pictured structure of a lattice group over $\mathrm{Ch}[1,k]$ reminds the cake "Napoleon". Here the groups play the role of the cakes lays, and the idempotents play the role of cream lays. Indeed, in the first step, by above remarked the $\Lambda[1]$ is a normal lattice subgroup of Λ. We have seen also that $\Lambda[1]$ is a group by multiplication (moreover, it is isomorphic to $\mathrm{pr}_G(\Lambda)$). Now add the cream: put $\Lambda_1 = \Lambda[1] \cup \{(e,2)\}$. It is not hard to see, that Λ_1 is a normal lattice subgroup of Λ. Next step: consider the $\{1,2\}$-layer $\Lambda[1,2]$ of Λ, which is a normal lattice subgroup of Λ. We note that $\Lambda_1 \leqslant \Lambda[1,2]$, moreover Λ_1 is a normal lattice subgroup of Λ. For every element $(x,j) \in \Lambda[1,2]$ denote by $(x,j)\Lambda_1$ the product $\{(x,j)\}\Lambda_1$. This subset is called a *coset* by Λ_1. Since $(x,j) \in \Lambda[1,2]$, $j \leqslant 2$, so that $(x,j) = (xe, j \wedge 2) = (x,j)(e,2) \in (x,j)\Lambda_1$. It follows that $\Lambda[1,2]$ is an union of all subsets $(x,j)\Lambda_1$. Suppose that $(x,j)\Lambda_1 \neq \Lambda_1$. Then $x \neq e$ and $j = 2$. Thus we can see that the equality $(x,2) = (y,2)(z,m)$ where $(z,m) \in \Lambda_1$ is possible only in the case when $m = 2$. In turn, the single pair of Λ_1, whose second component is equal to 2, is the pair $(e,2)$. Hence $(x,2) = (y,2)(e,2)$, so that $x = y$. In other words, the equality $(x,2)\Lambda_1 = (y,2)\Lambda_1$ is possible only in the case, when $x = y$. Consider the product of subsets $((x,2)\Lambda_1)((y,2)\Lambda_1)$. Its arbitrary element has a form $(x,2)(u,j)(y,2)(v,m)$ where $(u,j),(v,m) \in \Lambda_1$. If $j = 1$ or $m = 1$, then $(x,2)(u,j)(y,2)(v,m) = (xuyv,1) \in \Lambda_1$. Hence if $(x,2)(u,j)(y,2)(v,m) \notin \Lambda_1$, then $j = m = 2$. But it is possible only if $u = v = e$. In this case, $(x,2)(u,j)(y,2)(v,m) = (xy,2)$. In turn it follows that $((x,2)\Lambda_1)((y,2)\Lambda_1) = (xy,2)\Lambda_1$. Hence the set of all cosets by Λ_1 becomes a semigroup. Moreover, this semigroup is a group, because it has an identity element $(e,2)\Lambda_1 = \Lambda_1$, and for every coset $(x,2)\Lambda_1$ we have $(x^{-1},2)\Lambda_1(x,2)\Lambda_1 = (e,2)\Lambda_1 = (x,2)\Lambda_1(x^{-1},2)$. Therefore we can talk here about a factor-group of a lattice group $\Lambda[1,2]$ by the normal lattice subgroup Λ_1. For it we will use a common notation $\Lambda[1,2]/\Lambda_1$. We emphasize that here we are talking about a factor-group, rather than a lattice factor-group. It is our special selection provides such an opportunity; in general, is not always the family of cosets by normal lattice subgroup is a group or a lattice group.The mapping Φ, defined by the rule $\Phi((x,2)) = (x,2)\Lambda_1$, $(x,2) \in K_2$, is an epimorphism. As we have seen early, the equality $(x,2)\Lambda_1 = \Lambda_1$ is possible only in the case when $x = e$, which shows that Φ is an isomorphism. Since $K_2 \cong \mathrm{H}(2)$, we obtain that $\Lambda[1,2]/\Lambda_1$ is isomorphic to the 2-hoop of Λ.

Adding the next lay of the cream $\{(e,3)\}$ to $\Lambda[1,2]$, we come to the normal lattice subgroup $\Lambda_2 = \Lambda[1,2] \cup \{(e,3)\}$, and then we cover it with the next lay of cake, i.e. extend Λ_2 to the $\{1,2,3\}$-layer $\Lambda[1,2,3]$ of Λ,

which is a normal lattice subgroup of Λ. Using the above arguments, we shows that a family of cosets $(x,3)\Lambda_2$ is a group by multiplication and this group is isomorphic to the 2-hoop of Λ. And so on. As the result we obtain the sequences

$$\Lambda_0 = \{(e,1)\} \leqslant \Lambda[1] \leqslant \Lambda_1 \leqslant \Lambda[1,2] \leqslant \Lambda_2 \leqslant \Lambda[1,2] \leqslant \ldots \leqslant \Lambda_{k-1} \leqslant \Lambda$$

of normal lattice subgroups such that $\Lambda_m = \Lambda[1,\ldots,m] \cup \{(e,m+1)\}$, and $\Lambda[1,\ldots,m+1]/\Lambda_m \cong \mathrm{H}(m+1)$, $0 \leqslant m \leqslant k-1$.

Note, that in the theory of fuzzy groups we could not find any similar description of a general structure of a fuzzy group γ for the case when $\mathrm{Im}(\gamma)$ is finite.

References

[1] Goguen J.A., *L-Fuzzy Sets*, Journal of Math. Analysis and Applications., N.**18**,1967, pp.145-174.

[2] Kurdachenko L.A., Grin K.O., Turbay N.A., *On normalizers in fuzzy groups*, Algebra and Discrete Mathematics., N.**15**,2013, pp.23-36.

[3] Kurdachenko L.A., Otal J., Subbotin I.Ya., *On permutable fuzzy subgroups*, Serdica Mathematical Journal., N.**39**,2013, pp.83-102.

[4] Mordeson J.N., Nair P.S., *Fuzzy Mathematics*, Springer: Berlin., 2001.

[5] Mordeson J.N., Bhutani K.R., Rosenfeld A., *Fuzzy Group Theory*, Springer: Berlin, 2005.

[6] Zadeh L.A., *Fuzzy sets*, Information Control., N.**8**, 1965, pp.338-353.

Contact information

Leonid A. Kurdachenko, Viktoriia S. Yashchuk — Department of Algebra, Oles Honchar Dnipropetrovsk National University, 72 Gagarin Av., Dnepropetrovsk, Ukraine 49010
E-Mail(*s*): lkurdachenko@i.ua, ViktoriiaYashchuk@mail.ua

Igor Ya. Subbotin — Department of Mathematics and Natural Sciences, National University, 5245 Pacific Concourse Drive, LA, CA 90045, USA
E-Mail(*s*): isubboti@nu.edu

Received by the editors: 02.04.2015
and in final form 02.04.2015.

Algebra and Discrete Mathematics
Volume **20** (2015). Number 1, pp. 142–151

RESEARCH ARTICLE

On the units of integral group ring of $C_n \times C_6$

Ömer Küsmüş

Communicated by I. Ya. Subbotin

Abstract. There are many kind of open problems with varying difficulty on units in a given integral group ring. In this note, we characterize the unit group of the integral group ring of $C_n \times C_6$ where $C_n = \langle a : a^n = 1\rangle$ and $C_6 = \langle x : x^6 = 1\rangle$. We show that $\mathcal{U}_1(\mathbb{Z}[C_n \times C_6])$ can be expressed in terms of its 4 subgroups. Furthermore, forms of units in these subgroups are described by the unit group $\mathcal{U}_1(\mathbb{Z}C_n)$. Notations mostly follow [11].

1. Introduction

Let G given as a finite group. Its integral group ring is denoted by $\mathbb{Z}G$. Invertible elements in $\mathbb{Z}G$ is called by units and the set of units forms a group according to the multiplication and is shown by $\mathcal{U}(\mathbb{Z}G)$. The group of units with augmentation 1 is displayed by $\mathcal{U}_1(\mathbb{Z}G)$. If one pay attention to the corresponding literature, that can easily see that the obtained results mostly arises from finite groups especially finite abelian groups. Fundamentals of the unit problem have come from the thesis of G. Higman in 1940. Higman stated and proved the following [4]:

Lemma 1. *If* $U(\mathbb{Z}G) = \pm G$, *then* $U(\mathbb{Z}[G \times C_2]) = \pm[G \times C_2]$.

Also, the following useful lemma was shown in [4] and [3].

2010 MSC: 16U60, 16S34.
Key words and phrases: group ring, integral group ring, unit group, unit problem.

Lemma 2. *$\mathcal{U}(\mathbb{Z}G)$ has a torsion-free complement of finite rank $\rho = \frac{1}{2}(|G| + n_2 + 1 - 2l)$ where n_2 shows the number of elements of order 2 in G and l is the number of all distinct cyclic subgroups of G.*

On the other hand, in [7], Li considered the question: If $\mathcal{U}(\mathbb{Z}G)$ has a normal complement generated by bicyclic units, does $\mathcal{U}(\mathbb{Z}[G \times C_2])$ has also a normal complement generated by bicyclic units? Jespers showed that the answer for this question is yes while $G = D_6$ or D_8 [8–10]. Li gave a counterexample for showing this is not true in general by considering the group $D_8 \times C_2 \times C_2$ [7]. However, Li proved that if $\mathcal{U}(\mathbb{Z}G)$ is generated by unitary units, then $\mathcal{U}(\mathbb{Z}[G \times C_2])$ is also generated by unitary units [7]. Another description of $\mathcal{U}(\mathbb{Z}[G \times C_2])$ was given by Low in [6] by linearly extending some group epimorphisms to the group ring homomorphisms. He also tried to generalize the problem for $\mathcal{U}(\mathbb{Z}[G \times C_p])$ where p is a prime integer. In [6], He showed that

$$\mathcal{U}(\mathbb{Z}[G \times C_p]) = K \rtimes \mathcal{U}(\mathbb{Z}G) \cong M \rtimes \mathcal{U}(\mathbb{Z}G)$$

where K is the kernel of the natural group homomorphism: $\pi : \mathcal{U}(\mathbb{Z}[G \times C_p]) \longrightarrow \mathcal{U}(\mathbb{Z}G)$ and M is a subgroup of finite index in $\mathcal{U}(\mathbb{Z}[\zeta]G)$ such that ζ is a primitive p^{th} root of unity. Low also explicitly proved the following 4 lemmas [6]:

Lemma 3. *Let $G^* = G \times \langle x : x^2 = 1\rangle$. Then, $\mathcal{U}(\mathbb{Z}G^*)$ is obtained as*

$$\{u = 1 + (x-1)\alpha : \alpha \in \mathbb{Z}G, u \in \mathcal{U}(\mathbb{Z}G^*)\} \rtimes \mathcal{U}(\mathbb{Z}G).$$

Further, $1 + (x-1)\alpha \in \mathcal{U}(\mathbb{Z}G^*) \Leftrightarrow 1 - 2\alpha \in \mathcal{U}(\mathbb{Z}G)$.

Lemma 4. *Let $P = \langle a, b : a^4 = b^4 = 1, [b, a] = a^2\rangle$ be the indecomposable group of order 16. Then,*

$$\mathcal{U}(\mathbb{Z}[P \times C_2]) = \pm[F_{65} \rtimes F_9] \rtimes (P \times C_2)$$

where F_i denotes a free group of rank i.

Lemma 5. *Let $C_5^* = \langle c : c^5 = 1\rangle \times \langle x : x^2 = 1\rangle$. Then, the unit group*

$$\mathcal{U}(\mathbb{Z}C_5^*) = \langle 1 + (x-1)P\rangle \times \langle v\rangle \times C_5^*$$

where $P = -3 - c + 3c^2 + 3c^3 - c^4$ and $v = (c+1)^2 - \widehat{c}$.

Lemma 6. *Let $C_8^* = \langle c : c^8 = 1\rangle \times \langle x : x^2 = 1\rangle$. Then, the unit group*

$$\mathcal{U}(\mathbb{Z}C_8^*) = \langle 1 + (x-1)P\rangle \times \langle v\rangle \times C_8^*$$

where $P = -4 - 3c + 3c^3 + 4c^4 + 3c^5 - 3c^7$ and $v = 3 - \widehat{c} + 2(c + c^7) + (c^2 + c^6)$.

Kelebek and Bilgin considered the finite abelian group $C_n \times K_4$ where K_4 is the Klein 4-group and characterized the unit group of its integral group ring in terms of 4 components as follows [1]:

Theorem 1. $\mathcal{U}_1(\mathbb{Z}[C_n \times K_4]) = \mathcal{U}_1(\mathbb{Z}C_n) \times (1+K^x) \times (1+K^y) \times (1+K^{xy})$ *where*

$$1 + K^x = \{1 + (x-1)P : 1 - 2P \in \mathcal{U}_1(\mathbb{Z}C_n)\}$$
$$1 + K^y = \{1 + (y-1)P : 1 - 2P \in \mathcal{U}_1(\mathbb{Z}C_n)\}$$
$$1 + K^{xy} = \{1 + (x-1)(y-1)P : 1 + 4P \in \mathcal{U}_1(\mathbb{Z}C_n)\}$$

2. Motivation for construction of $\mathcal{U}_1(\mathbb{Z}[C_n \times C_6])$

Now, let us begin with some remarks.

Remark 1. The following maps are group epimorphisms:

$$\pi_{x^2} : C_n \times C_6 \longrightarrow C_n \times \langle x^2 \rangle$$
$$a \mapsto a$$
$$x \mapsto x^2$$

$$\pi_{x^3} : C_n \times C_6 \longrightarrow C_n \times \langle x^3 \rangle$$
$$a \mapsto a$$
$$x \mapsto x^3$$

Remark 2. $\mathrm{Ker}(\pi_{x^2}) = \langle x^3 \rangle$ and $\mathrm{Ker}(\pi_{x^3}) = \langle x^2 \rangle$.

Since $C_n \times \langle x^2 \rangle \hookrightarrow C_n \times C_6$, $C_n \times \langle x^3 \rangle \hookrightarrow C_n \times C_6$ and i denotes the inclusion map, we get the following short exact sequences at group level:

$$0 \longrightarrow \langle x^3 \rangle \xrightarrow{i} C_n \times C_6 \xrightarrow{\pi_{x^2}} C_n \times \langle x^2 \rangle \longrightarrow 0$$
$$0 \longrightarrow \langle x^2 \rangle \xrightarrow{i} C_n \times C_6 \xrightarrow{\pi_{x^3}} C_n \times \langle x^3 \rangle \longrightarrow 0$$
$$0 \longrightarrow \langle x \rangle \xrightarrow{i} C_n \times C_6 \xrightarrow{\pi_{x^2}\pi_{x^3}} C_n \longrightarrow 0$$

If we linearly extend π_{x^2} and π_{x^3} to integral group rings over $\mathbb{Z}$, we obtain the following ring homomorphisms:

$$\overline{\pi}_{x^2} : \mathbb{Z}[C_n \times C_6] \longrightarrow \mathbb{Z}[C_n \times \langle x^2 \rangle]$$
$$\sum_{j=0}^{5} P_j x^j \mapsto (P_0 + P_3) + (P_1 + P_4)x^2 + (P_2 + P_5)x^4$$

and

$$\overline{\pi}_{x^3} : \mathbb{Z}[C_n \times C_6] \longrightarrow \mathbb{Z}[C_n \times \langle x^3 \rangle]$$
$$\sum_{j=0}^{5} P_j x^j \mapsto (P_0 + P_2 + P_4) + (P_1 + P_3 + P_5)x^3$$

Lemma 7. $K^{x^2} := \mathrm{Ker}(\overline{\pi}_{x^2}) = (x^3 - 1)\mathbb{Z}[C_n \times \langle x^2 \rangle]$

Proof.

$$\begin{aligned}
\mathrm{Ker}(\overline{\pi}_{x^2}) &= \Big\{\sum_{i=0}^{5} P_i x^i : \overline{\pi}_{x^2}(\sum_{i=0}^{5} P_i x^i) = 0, P_i \in \mathbb{Z}C_n\Big\} \\
&= \Big\{\sum_{i=0}^{5} P_i x^i : P_0 + P_3 = P_1 + P_4 = P_2 + P_5 = 0\Big\} \\
&= \Big\{\sum_{i=0}^{5} P_i x^i : P_0 = -P_3, P_1 = -P_4, P_2 = -P_5\Big\} \\
&= \{-P_3 - P_4 x - P_5 x^2 + P_3 x^3 + P_4 x^4 + P_5 x^5\} \\
&= \{(x^3 - 1)P_3 + (x^4 + x)P_4 + (x^5 - x^2)P_5\} \\
&= (x^3 - 1)[\mathbb{Z}C_n \oplus x^2\mathbb{Z}C_n \oplus x^4\mathbb{Z}C_n] \\
&= (x^3 - 1)\mathbb{Z}[C_n \times \langle x^2 \rangle].
\end{aligned}$$

□

Lemma 8. $K^{x^3} := \mathrm{Ker}(\overline{\pi}_{x^3}) = (x^2 - 1)[\mathbb{Z}C_n \oplus x\mathbb{Z}C_n \oplus x^2\mathbb{Z}C_n \oplus x^3\mathbb{Z}C_n]$

Proof.

$$\begin{aligned}
\mathrm{Ker}(\overline{\pi}_{x^3}) &= \Big\{\sum_{i=0}^{5} P_i x^i : \overline{\pi}_{x^3}(\sum_{i=0}^{5} P_i x^i) = 0, P_i \in \mathbb{Z}C_n\Big\} \\
&= \Big\{\sum_{i=0}^{5} P_i x^i : P_0 + P_2 + P_4 = P_1 + P_3 + P_5 = 0\Big\} \\
&= \Big\{\sum_{i=0}^{5} P_i x^i : P_0 = -(P_2 + P_4), P_1 = -(P_3 + P_5)\Big\} \\
&= \{(x^2 - 1)[P_2 + xP_3 + (x^2 + 1)P_4 + (x^2 + 1)xP_5]\} \\
&= (x^2 - 1)[\mathbb{Z}C_n \oplus x\mathbb{Z}C_n \oplus x^2\mathbb{Z}C_n \oplus x^3\mathbb{Z}C_n]
\end{aligned}$$

Similarly, we can write the following ring homomorphism:

$$\begin{aligned} \overline{\pi}_{x^2}\overline{\pi}_{x^3} : \mathbb{Z}[C_n \times C_6] &\longrightarrow \mathbb{Z}C_n \\ \sum_{j=0}^{5} P_j x^j &\mapsto \sum_{j=0}^{5} P_j. \end{aligned}$$

□

Lemma 9. $K^{x^2x^3} := \mathrm{Ker}(\overline{\pi}_{x^2}\overline{\pi}_{x^3}) = \oplus_{j=1}^{5}(x^j - 1)\mathbb{Z}C_n$

Proof.

$$\begin{aligned} \mathrm{Ker}(\overline{\pi}_{x^2}\overline{\pi}_{x^3}) &= \Big\{\sum_{i=0}^{5} P_i x^i : \overline{\pi}_{x^2}\overline{\pi}_{x^3}(\sum_{i=0}^{5} P_i x^i) = 0, P_i \in \mathbb{Z}C_n\Big\} \\ &= \Big\{\sum_{i=0}^{5} P_i x^i : \sum_{i=0}^{5} P_i = 0, P_i \in \mathbb{Z}C_n\Big\} \\ &= \Big\{\sum_{i=0}^{5} P_i x^i : P_0 = -\sum_{i=1}^{5} P_i, P_i \in \mathbb{Z}C_n\Big\} \\ &= \Big\{-\sum_{i=1}^{5} P_i + \sum_{i=1}^{5} P_i x^i : P_i \in \mathbb{Z}C_n\Big\} \\ &= \Big\{\sum_{j=1}^{5} (x^j - 1)P_j : P_j \in \mathbb{Z}C_n\Big\} \\ &= \oplus_{j=1}^{5}(x^j - 1)\mathbb{Z}C_n. \end{aligned}$$

By Remarks 1 and 2, we get the following short exact sequences at group ring level:

$$0 \longrightarrow K^{x^2} \xrightarrow{i} \mathbb{Z}[C_n \times C_6] \xrightarrow{\overline{\pi}_{x^2}} \mathbb{Z}[C_n \times \langle x^2 \rangle] \longrightarrow 0$$

$$0 \longrightarrow K^{x^3} \xrightarrow{i} \mathbb{Z}[C_n \times C_6] \xrightarrow{\overline{\pi}_{x^3}} \mathbb{Z}[C_n \times \langle x^3 \rangle] \longrightarrow 0$$

$$0 \longrightarrow K^{x^2x^3} \xrightarrow{i} \mathbb{Z}[C_n \times C_6] \xrightarrow{\overline{\pi}_{x^2}\overline{\pi}_{x^3}} \mathbb{Z}C_n \longrightarrow 0$$

If we restrict $\overline{\pi}_{x^2}$ and $\overline{\pi}_{x^3}$ to the unit level, we conclude that the followings are also short exact sequences:

$$1 \longrightarrow \mathcal{U}_1(1 + K^{x^2}) \xrightarrow{i} \mathcal{U}_1(\mathbb{Z}[C_n \times C_6]) \xrightarrow{\overline{\pi}_{x^2}} \mathcal{U}_1(\mathbb{Z}[C_n \times \langle x^2 \rangle]) \longrightarrow 1$$

$$1 \longrightarrow \mathcal{U}_1(1 + K^{x^3}) \xrightarrow{i} \mathcal{U}_1(\mathbb{Z}[C_n \times C_6]) \xrightarrow{\overline{\pi}_{x^3}} \mathcal{U}_1(\mathbb{Z}[C_n \times \langle x^3 \rangle]) \longrightarrow 1$$

$$1 \longrightarrow \mathcal{U}_1(1 + K^{x^2x^3}) \xrightarrow{i} \mathcal{U}_1(\mathbb{Z}[C_n \times C_6]) \xrightarrow{\overline{\pi}_{x^2}\overline{\pi}_{x^3}} \mathcal{U}_1(\mathbb{Z}C_n) \longrightarrow 1$$

Since we can consider embeddings $\mathcal{U}_1(\mathbb{Z}[C_n \times \langle x^2 \rangle]) \hookrightarrow \mathcal{U}_1(\mathbb{Z}[C_n \times C_6])$ and $\mathcal{U}_1(\mathbb{Z}[C_n \times \langle x^3 \rangle]) \hookrightarrow \mathcal{U}_1(\mathbb{Z}[C_n \times C_6])$, the following split extensions hold:

$$\begin{aligned}\mathcal{U}_1(\mathbb{Z}[C_n \times C_6]) &= \mathcal{U}_1(1+K^{x^2}) \times \mathcal{U}_1(\mathbb{Z}[C_n \times \langle x^2 \rangle]) \\ &= \mathcal{U}_1(1+K^{x^3}) \times \mathcal{U}_1(\mathbb{Z}[C_n \times \langle x^3 \rangle]) \\ &= \mathcal{U}_1(1+K^{x^2x^3}) \times \mathcal{U}_1(\mathbb{Z}C_n). \qquad \square\end{aligned}$$

Remark 3. In $\mathcal{U}_1(\mathbb{Z}[C_n \times C_6])$ the normal subgroups $\mathcal{U}_1(1+K^{x^2})$, $\mathcal{U}_1(1+K^{x^3})$ and $\mathcal{U}_1(1+K^{x^2x^3})$ are determined as in the following forms respectively:

(i) $\{u = 1 + (x^3-1)[P_0 + P_2x^2 + P_4x^4] : u \text{ is a unit}\}$;

(ii) $\{u = 1 + (x^2-1)[P_0 + P_1x + P_2x^2 + P_3x^3] : u \text{ is a unit}\}$;

(iii) $\{u = 1 + \sum_{j=1}^{5}(x^j-1)P_j : u \text{ is a unit}\}$.

3. An explicit characterization of $\mathcal{U}_1(\mathbb{Z}[C_n \times C_6])$

In this section, an explicit characterization of $\mathcal{U}_1(\mathbb{Z}[C_n \times C_6])$ is given with the help of the results in the previous section. First, we should give some restrictions of the maps $\overline{\pi}_{x^2}$, $\overline{\pi}_{x^3}$ and $\overline{\pi}_{x^2}\overline{\pi}_{x^3}$. Let $\overline{\pi}_{x^3}|_{\mathcal{U}_1(1+K^{x^2})}$ denote the restriction of $\overline{\pi}_{x^3}$ on $\mathcal{U}_1(1+K^{x^2})$.

Lemma 10. $W_1 := Im(\overline{\pi}_{x^3}|_{\mathcal{U}_1(1+K^{x^2})}) = 1 + (x^3-1)\mathbb{Z}C_n$.

Proof. Let us take an element from $\mathcal{U}_1(1+K^{x^2})$ as $\gamma = 1 + (x^3-1)[P_0 + P_2x^2 + P_4x^4]$ where $P_i \in \mathbb{Z}C_n$.Then,

$$\overline{\pi}_{x^3} : \gamma \mapsto 1 + (x^3-1)[P_0 + P_2 + P_4].$$

Say $P_0 + P_2 + P_4 = P$. Thus, $Im(\overline{\pi}_{x^3}|_{\mathcal{U}_1(1+K^{x^2})})$ consists of elements of the form $1 + (x^3-1)P$. $\square$

Lemma 11. $W_2 := \mathrm{Ker}(\overline{\pi}_{x^3}|_{\mathcal{U}_1(\mathbb{Z}[C_n \times \langle x^2 \rangle])}) = 1 + (x^2-1)\mathbb{Z}C_n \oplus (x^4-1)\mathbb{Z}C_n$.

Proof. Let us take an element from $\mathcal{U}_1(\mathbb{Z}[C_n \times \langle x^2 \rangle])$ as $\sigma = P_0 + P_2x^2 + P_4x^4$. Here, we can manipulate the parameter $P_0 = 1 + P_0'$. Then, we get

$$\overline{\pi}_{x^3} : \sigma \mapsto 1 + P_0' + P_2 + P_4 = 1 \Longleftrightarrow P_0' = -P_2 - P_4.$$

This means that the kernel consists of elements of the form

$$1 + (-P_2 - P_4) + P_2x^2 + P_4x^4 = 1 + (x^2-1)P_2 + (x^4-1)P_4.$$

Hence the required is obtained. $\square$

Lemma 12.

$$W_3 := \mathrm{Ker}(\overline{\pi}_{x^3}|_{\mathcal{U}_1(1+K^{x^2})}) = 1+(x^3-1)(x^2-1)\mathbb{Z}C_n \oplus (x^3-1)(x^4-1)\mathbb{Z}C_n.$$

Proof. Again, let us consider an element from $\mathcal{U}_1(1+K^{x^2})$ as $\eta = 1 + (x^3-1)[P_0 + P_2x^2 + P_4x^4]$. Then,

$$\overline{\pi}_{x^3} : \eta \mapsto 1 + (x^3-1)[P_0+P_2+P_4] = 1 \Longleftrightarrow P_0 = -P_2 - P_4$$

Thus, $\mathrm{Ker}(\overline{\pi}_{x^3}|_{\mathcal{U}_1(1+K^{x^2})})$ consists of

$$\begin{aligned} 1 + (x^3-1)[P_0 + P_2x^2 + P_4x^4] &= 1 + (x^3-1)[-P_2 - P_4 + P_2x^2 + P_4x^4] \\ &= 1 + (x^3-1)[(x^2-1)P_2 + (x^4-1)P_4]. \ \square \end{aligned}$$

Therefore, by Lemma 10, Lemma 11 and Lemma 12, we can construct the following commutative diagram:

$$\begin{array}{ccccc} W_3 & \xrightarrow{i} & \mathcal{U}_1(1+K^{x^3}) & \xrightarrow{\overline{\pi}_{x^2}} & W_2 \\ \downarrow i & & \downarrow i & & \downarrow i \\ \mathcal{U}_1(1+K^{x^2}) & \xrightarrow{i} & \mathcal{U}_1(\mathbb{Z}[C_n \times C_6]) & \xrightarrow{\overline{\pi}_{x^2}} & \mathcal{U}_1(\mathbb{Z}[C_n \times \langle x^2 \rangle]) \\ \downarrow \overline{\pi}_{x^3} & & \downarrow \overline{\pi}_{x^3} & & \downarrow \overline{\pi}_{x^3} \\ W_1 & \xrightarrow{i} & \mathcal{U}_1(\mathbb{Z}[C_n \times \langle x^3 \rangle]) & \xrightarrow{\overline{\pi}_{x^2}} & \mathcal{U}_1(\mathbb{Z}C_n) \end{array}$$

Since we can take embeddings as the inverses of $\overline{\pi}_{x^2}$ and $\overline{\pi}_{x^3}$, this diagram splits as follows:

$$\mathcal{U}_1(\mathbb{Z}[C_n \times C_6]) = W_1 \times W_2 \times W_3 \times \mathcal{U}_1(\mathbb{Z}C_n).$$

Now, let us characterize explicitly W_1, W_2 and W_3.

Proposition 1. *$u = 1+(x^3-1)P \in W_1$ is a unit $\Longleftrightarrow 1-2P \in \mathcal{U}_1(\mathbb{Z}C_n)$*

Proof.

$$\begin{aligned} u = 1 + (x^3-1)P \text{ is unit} &\Longleftrightarrow \exists v = 1 + (x^3-1)Q : uv = 1 \\ &\Longleftrightarrow 1 + (x^3-1)[P+Q-2PQ] = 1 \\ &\Longleftrightarrow P + Q - 2PQ = 0 \\ &\Longleftrightarrow 1 - 2P - 2Q + 4PQ = 1 \\ &\Longleftrightarrow (1-2P)(1-2Q) = 1 \\ &\Longleftrightarrow 1 - 2P \in \mathcal{U}_1(\mathbb{Z}C_n). \qquad \square \end{aligned}$$

Proposition 2. *$u = 1 + (x^2 - 1)P + (x^4 - 1)Q \in W_2$ is a unit $\Longleftrightarrow$ $P^2 + Q^2 - PQ - P - Q = 0$*

Proof. First, we need to define a closed operation. If we define $\alpha = x^2 - 1$ and $\beta = x^4 - 1$, we get the following straightforward computations:

$$\begin{aligned}
\alpha^2 &= (x^2-1)^2 = -2(x^2-1) + (x^4-1) = -2\alpha + \beta \\
\alpha\beta &= (x^2-1)(x^4-1) = -(x^2-1) - (x^4-1) = -\alpha - \beta \\
\beta^2 &= (x^4-1)^2 = -2(x^4-1) + (x^2-1) = \alpha - 2\beta
\end{aligned}$$

Let us state this operation in a table as follows:

$\bullet$	α	β
α	$-2\alpha + \beta$	$-\alpha - \beta$
β	$-\alpha - \beta$	$\alpha - 2\beta$

Now, we can give a necessary and sufficient condition to be a unit for the element u. $u = 1 + (x^2 - 1)P + (x^4 - 1)Q \in W_2$ is a unit if and only if $\exists v = 1 + (x^2 - 1)P' + (x^4 - 1)Q'$ such that $uv = 1$. Hence,

$$1 + \alpha P + \beta Q + \alpha P' + \beta Q' + \alpha^2 PP' + \beta^2 QQ' + \alpha\beta(PQ' + P'Q) = 1.$$

By the above operation, we can arrange this equation as

$$\begin{aligned}
1 + \alpha(P + P') + \beta(Q + Q') + (-2\alpha + \beta)PP' \\
+ (\alpha - 2\beta)QQ' + (-\alpha - \beta)(PQ' + P'Q) = 1
\end{aligned}$$

That is,

$$\begin{aligned}
1 + \alpha(P + P' - 2PP' + QQ' - PQ' - P'Q) \\
+ \beta(Q + Q' + PP' - 2QQ' - PQ' - P'Q) = 1.
\end{aligned}$$

This equation holds if and only if the following system of matrix has a unique solution:

$$\begin{bmatrix} 1 - 2P - Q & Q - P \\ P - Q & 1 - 2Q - P \end{bmatrix} \begin{bmatrix} P' \\ Q' \end{bmatrix} = \begin{bmatrix} -P \\ -Q \end{bmatrix}$$

Therefore,

$$\begin{bmatrix} 1 - 2P - Q & Q - P \\ P - Q & 1 - 2Q - P \end{bmatrix} \in SL_2(\mathbb{Z}C_n)$$

A straightforward calculation shows that $P^2 + Q^2 - PQ - P - Q = 0$. $\square$

Proposition 3. *$u = 1 + (x^3-1)(x^2-1)P + (x^3-1)(x^4-1)Q \in W_3$ is a unit if and only if the following equation holds:*

$$2P^2 + 2Q^2 - 2PQ - P - Q = 0$$

Proof. First, let $\lambda = (x^3-1)(x^2-1)$ and $\mu = (x^3-1)(x^4-1)$. One can easily compute the followings:

$$\begin{aligned}\lambda^2 &= (x^3-1)^2(x^2-1)^2 = 4\lambda - 2\mu,\\ \lambda\mu &= (x^3-1)^2(x^2-1)(x^4-1) = 2\lambda + 2\mu,\\ \mu^2 &= (x^3-1)^2(x^4-1)^2 = -2\lambda + 4\mu.\end{aligned}$$

In a better expression, we write

$\bullet$	λ	μ
λ	$4\lambda - 2\mu$	$2\lambda + 2\mu$
μ	$2\lambda + 2\mu$	$-2\lambda + 4\mu$

Now, let us determine the necessary and sufficient condition to be a unit for an element u. $u = 1 + \lambda P + \mu Q \in W_3$ is a unit if and only if $\exists v = 1 + \lambda P' + \mu Q' : uv = 1$. Thus, a straight forward computation shows us that

$$\begin{aligned}1 + \lambda(P + P' + 4PP' + 2P'Q + 2PQ' - 2QQ')\\ + \mu(Q + Q' - 2PP' + 2P'Q + 2PQ' + 4QQ') = 1.\end{aligned}$$

This equation holds if and only if the following system of matrix has a unique solution:

$$\begin{bmatrix} 1+4P+2Q & 2P-2Q \\ 2Q-2P & 1+2P+4Q \end{bmatrix} \begin{bmatrix} P' \\ Q' \end{bmatrix} = \begin{bmatrix} -P \\ -Q \end{bmatrix}$$

Then, the required result comes from the following:

$$\begin{bmatrix} 1+4P+2Q & 2P-2Q \\ 2Q-2P & 1+2P+4Q \end{bmatrix} \in SL_2(\mathbb{Z}C_n). \qquad \square$$

Consequently, we can summarize all the obtained results as follows:

Corollary 1. *$\mathcal{U}_1(\mathbb{Z}[C_n \times C_6]) = \mathcal{U}_1(\mathbb{Z}C_n) \times U \times V \times W$ where*

$$\begin{aligned}U &= \{1 + (x^3-1)P : 1 - 2P \in \mathcal{U}_1(\mathbb{Z}C_n)\}\\ V &= \{1 + \alpha P + \beta Q : P^2 + Q^2 - PQ - P - Q = 0\}\\ W &= \{1 + \lambda P + \mu Q : 2P^2 + 2Q^2 - 2PQ - P - Q = 0\}\end{aligned}$$

such that

$$\alpha = x^2 - 1, \;\; \beta = x^4 - 1, \;\; \lambda = (x^3-1)(x^2-1), \;\; \mu = (x^3-1)(x^4-1).$$

Acknowledgements

The authors would like to thank to all the members of the journal Algebra and Discrete Mathematics.

References

[1] Kelebek I. G. and T. Bilgin, *Characterization of* $\mathcal{U}_1(\mathbb{Z}[C_n \times K_4])$, Eur. J. Pure and Appl. Math., 7(4), 462-471, 2014.

[2] T. Bilgin, *Characterization of* $\mathcal{U}_1(\mathbb{Z}C_{12})$, Int. J. Pure Appl. Math., 14, 531-535, 2004.

[3] Ayoub R. G. and Ayoub C., *On The Group Ring of a Finite Abelian Group*, Bull. Aust. Math. Soc., 1, 245-261, 1969.

[4] Higman G., *The Units of Group Rings*, Proc. London Math. Soc., 46(2), 1940.

[5] Karpilovsky G., *Commutative Group Algebras*. Marcel Dekker, New York, 1983.

[6] Low R. M., *On The Units of Integral Group Ring* $\mathbb{Z}[G \times C_p]$, J. Algebra Appl., 7, 393-403, 2008.

[7] Y Li. *Units of* $\mathbb{Z}(G \times C_2)$, Quaest. Math., 21(3-4), 201-218, 1998.

[8] Jespers E., *Bicyclic Units in Some Integral Group Rings. Canad. Math. Bull*, 38(1), 80-86, 1995.

[9] Jespers E. and Leal G., *Describing Units of Integral Group Rings of Some 2-groups*, Comm. Algebra. 19, 1809-1827, 1991.

[10] Jespers E. and Parmenter M. M., *Bicyclic Units in* $\mathbb{Z}S_3$, Bull. Soc. Math. Belg. Ser. B., 44, 141-146, 1992.

[11] Milies C. P. and Sehgal S. K., *An Introduction to Group Ring*, Kluwer Academic Publishers, London, 2002.

CONTACT INFORMATION

Ö. Küsmüş Department of Mathematics, Faculty of Science, Yuzuncu Yil University, 65080, Van, TURKEY
E-Mail(s): omerkusmus@yyu.edu.tr

Received by the editors: 21.02.2015
and in final form 05.03.2015.

Algebra and Discrete Mathematics
Volume **20** (2015). Number 1, pp. 152–170

RESEARCH ARTICLE

On algebraic graph theory and non-bijective multivariate maps in cryptography

Vasyl Ustimenko

Communicated by R. I. Grigorchuk

This paper is dedicated to the glorious 60-th anniversary of Efim Zelmanov whose research is an inspirational example of continuous fruitful serving to Algebra

Abstract. Special family of non-bijective multivariate maps F_n of ${Z_m}^n$ into itself is constructed for $n = 2, 3, \ldots$ and composite m. The map F_n is injective on $\Omega_n = \{\mathrm{x} | x_1 + x_2 + \ldots x_n \in {Z_m}^*\}$ and solution of the equation $F_n(\mathrm{x}) = \mathrm{b}, \mathrm{x} \in \Omega_n$ can be reduced to the solution of equation $z^r = \alpha$, $z \in {Z_m}^*$, $(r, \phi(m)) = 1$. The "hidden RSA cryptosystem" is proposed.

Similar construction is suggested for the case $\Omega_n = {Z_m}^{*n}$.

1. Introduction

The RSA is one of the most popular cryptosystems. It is based on a number factorisation problem and Euler Theorem. Peter Shor discovered that factorisation problem can be effectively solved with the usage of theoretical quantum computer. It means that RSA could not be a security tool in the future postquantum era. One of the research directions which can lead to a postquantum secure public key is the Multivariate Cryptography which uses polynomial maps of affine space K^n defined over a finite commutative ring into itself as encryption tools (see [1]). This is a young promising research area with the current lack of known cryptosystems

Key words and phrases: multivariate cryptography, linguistic graphs, hidden Eulerian equation, hidden discrete logarithm problem.

with the proven resistence against attacks with the use of Turing machines. Other important direction of Postquantum Cryptography is a study of Super-elliptic Curves cryptosystems.

Applications of Algebraic Graph Theory to Multivariate Cryptography were observed in my talk at Central European Conference on Cryptology 2014 (Alfred Renyi Institute, Budapest) [2]. This talk was dedicated to algorithms based on bijective maps of affine spaces into themselves. Applications of algebraic graphs to cryptography started from symmetric algorithms based on explicit constructions of extremal graph theory and their directed analogs (see survey [3], [4]). The main idea is to convert an algebraic graph in finite automaton and use the preudorandom walks on the graph as encryption tools. This approach can be also used for the key exchange protocols. Nowadays the idea of "symbolic walks" on algebraic graphs when the walk on the graph depends on parameters given as special multivariate polynomials in variables depending on plainspace vector brings several public key cryptosystems. Other source of graphs suitable for cryptography is connected with finite geometries and their flag system (see [3], [5], [6] and further references).

This paper presents new cryptoalgorithm in terms of Algebraic Combinatorics which use non-bijective transformations of K^n.

Multivariate cryptography started from studies of potential for the special quadratic encryption multivariate bijective map of K^n, where K is an extention of finite field F_q of characteristic 2. One of the first such cryptosystems was proposed by Imai and Matsumoto, cryptanalysis for this system was invented by J. Patarin [1], [7]. The survey on various modifications of this algorithm and corresponding cryptanalysis the reader can find in [1]. Bijective multivariate sparse encryption maps of rather high degree based on walks in algebraic graphs were proposed in [8].

One of the first usage of non bijective map of multivariate cryptography was in *oil and vinegar* crptosystem proposed in [9] and analysed in [10]. Nowadays this general idea is strongly supported by the publication [11] dedicated to security analysis of direct attacks on modified unbalanced oil and vinegar systems. It looks like such systems and rainbow signatures schemes may lead to promising Public Key Schemes of Multivariate Encryption defined over finite fields. Non bijective multivariate sparse encryption maps of degree 3 and $\geqslant 3$ based on walks on algebraic graphs $D(n, K)$ defined over general commutative ring and their homomorphic images were proposed in [12].

The paper is dedicated to other constructions of non bijective maps. We introduce the concept of family of multivariate maps $F = F_n$ of

the free modules K^n onto itself decomposed into transition functions $F^1, F^2, \ldots, F^{s(n)}$ of special symbolic vertex automata of linguistic graphs. In case $K = Z_m$, where m is composite, it allows us to construct partially invertible F_n respectively to subsets Ω_n of ${Z_m}^n$. It means that the restriction of F on Ω_n is injective and the decomposition above allows us to solve the equation $F(\mathrm{x})) = \mathrm{b}$ for unknown $(\mathrm{x}) \in \Omega_n$ and $\mathrm{b} \in F(\Omega_n)$ in polynomial time. We are interested in the case of Eulerian maps F_n when the solution of the equation can be reduced to the study of equations of kind $z^r = d$, where z in ${Z_m}^*$ and $(r, \phi(m)) = 1$. We construct infinite families of maps of kind $H_n = \tau_1 F_n \tau_2$, where τ_i are bijective affine transformations of ${Z_m}^n$, with Eulerian F_n of bounded degree such that H_n is partially invertible for $\Omega_n = {Z_m}^{*n}$ and $\Omega_n = \{\mathrm{x} \in {Z_m}^n | x_1 + x_2 + \ldots x_n \in {Z_m}^*\}$.

So the following scheme of a cryptosystem can be used. Alice (the public key owner) uses special linguistic graph $L_n(Z_m)$, its symbolic automaton with a special symbolic key to generate the Eulerian map F_n and the list of transition functions $F^1, F^2, \ldots, F^{s(n)}$ of the symbolic computation. She chooses appropriate bijective affine transformations τ_1 and τ_2 and creates a deformation $H_n = \tau_1 F_n \tau_2$ which is partially invertible for Ω_n as above. Alice writes the following standard form for H_n:

$$x_1 \to h_1(x_1, x_2, \ldots, x_n), \quad x_2 \to h_2(x_1, x_2, \ldots, x_n), \quad \ldots,$$
$$x_n \to h_n(x_1, x_2, \ldots, x_n)$$

where polynomials $h_i(x_1, x_2, \ldots, x_n)$, $i = 1, 2, \ldots, n$ are given by their lists of monomial terms with respect to the chosen order.

She announces the form and the plainspace Ω_n in public way.

Notice that Alice keeps the transition functions generating F_n and *deformation rule* $H_n = \tau_1 F_n \tau_2$ in secret. Cryptanalytic knows only the list of h_i and the graph $L_n(Z_m)$.

Public user (Bob) writes his message $(p_1, p_2, \ldots, p_n)$ from the plainspace Ω_n. He computes the ciphertext $\mathrm{c} = (c_1, c_2, \ldots, c_n)$, $c_i = h_i(p_1, p_2, \ldots, p_n)$, $i = 1, 2, \ldots, n$ and sends it to Alice.

Alice solves the equation $F_n(x_1, x_2, \ldots, x_n) = (c_1, c_2, \ldots, c_n)$ due to her knowledge of symbolic key of the automaton. So she reads the plaintext.

Notice that to make this scheme feasible we need to care about polynomiality of generation time, bound for the degree of H_n, Eulerian nature of the map F_n. We achieve it via special choice of linguistic graph (well known graphs $D(n, K)$) and some restriction on symbolic keys.

Section 2 is dedicated to linguistic graphs and related to them automata. In Section 3 the reader can find information on chosen linguistic

graph $D(n,K)$. The properties af chosen computation of vertex automaton for graph $D(n,Z_m)$ are justified in section 4. Last section gives precise descryption of cryptosystem.

2. Linguistic graphs and their vertex automata

The missing definitions of graph-theoretical concepts which appear in this paper can be found in [13]. All graphs we consider are *simple* , i.e. undirected without loops and multiple edges. Let $V(G)$ and $E(G)$ denote the set of vertices and the set of edges of G respectively. Then $|V(G)|$ is called the *order* of G, and $|E(G)|$ is called the *size* of G. When it is convenient we shall identify G with the corresponding anti-reflexive binary relation on $V(G)$, i.e. $E(G)$ is a subset of $V(G)\times V(G)$ and write $v\,G\,u$ for the adjacent vertices u and v (or neighbours). We assume that $V(G)$ is a finite or an infinite set. The majority of examples will be *locally finite graphs* G, i.e. each vertex v has finite number of neighbours ($x\in V(G)$, such that $x\,G\,v$). We refer to $|\{x\in V(G)|x\,G\,v\}|$ as *degree of the vertex* v.

The sequence of distinct vertices $v_0,v_1,\dots,v_t$, such that $v_i\,G\,v_{i+1}$ for $i=1,\dots,t-1$ is a *path* in the graph. The path in G is called *simple* if all its vertices are distinct. The graph is *connected* if each two of its vertices are joined by some path. The length of the path is a number of its edges. The *distance* between two vertices u and v of the graph, denoted by $\mathrm{dist}(u,v)$, is the length of the shortest path between them. The *diameter* of the graph, denoted by $\mathrm{diam}(G)$, is the maximal distance between two vertices u and v of the graph. Let C_m denote the cycle of length m, i.e. the sequence of distinct vertices $v_0,\dots,v_m$ such that $v_i\,G\,v_{i+1}$, $i=1,\dots,m-1$ and $v_m\,G\,v_1$. The *girth* of a graph G, denoted by $g=g(G)$, is the length of the shortest cycle in G.

The *incidence structure* is the set V with partition sets P (*points*) and L (*lines*) and symmetric binary relation I such that the incidence of two elements implies that one of them is a point and another one is a line. We shall identify I with the simple graph of this incidence relation (*bipartite graph*).

We refer to a triple consisting of set V, its partition $V=P\cup L$ and symmetric and antireflexive binary relation I (incidence) on the set V, such that $\mathrm{x}I\mathrm{y}$ implies $\mathrm{x}\in P$, $\mathrm{y}\in L$ or $\mathrm{x}\in L$ and $\mathrm{y}\in P$ as *incidence structure*. The pair $\{\mathrm{x},\mathrm{y}\}$, $\mathrm{x}\in P$, $\mathrm{y}\in L$ such that $\mathrm{x}I\mathrm{y}$ is called a *flag* of incidence structure I.

Let K be a finite commutative ring. We refer to an incidence structure with a point set $P=P_{s,m}=K^{s+m}$ and a line set $L=L_{r,m}=K^{r+m}$ as

linguistic incidence structure I_m if point

$$(\mathrm{x}) = (x_1, x_2, \ldots, x_s, x_{s+1}, x_{s+2}, \ldots, x_{s+m})$$

is incident to line

$$[\mathrm{y}] = [y_1, y_2, \ldots, y_r, y_{r+1}, y_{r+2} \ldots, y_{r+m}]$$

if and only if the following relations hold

$$\begin{aligned}
\xi_1 x_{s+1} + \zeta_1 y_{r+1} &= f_1(x_1, x_2, \ldots, x_s, y_1, y_2, \ldots, y_r) \\
\xi_2 x_{s+2} + \zeta_2 y_{r+2} &= f_2(x_1, x_2, \ldots, x_s, x_{s+1}, y_1, y_2, \ldots, y_r, y_{r+1}) \\
&\ldots \\
\xi_m x_{s+m} + \zeta_m y_{r+m} &= f_m(x_1, x_2, \ldots, x_{s+m-1}, y_1, y_2, \ldots, y_{r+m-1})
\end{aligned}$$

where ξ_j and ζ_j, $j = 1, 2, \ldots, m$ are not zero divisors, and f_j are multivariate polynomials with coefficients from K. Brackets and parenthesis allow us to distinguish points from lines (see [14]).

The colour $\rho(\mathrm{x}) = \rho((\mathrm{x}))$ ($\rho(\mathrm{y}) = \rho([\mathrm{y}])$) of point (x) (line [y]) is defined as projection of an element (x) ([y]) from a free module on its initial s (relatively r) coordinates. As it follows from the definition of linguistic incidence structure for each vertex of incidence graph there exists unique neighbour of a chosen colour. We also consider a linguistic incidence structures defined by infinite number of equations.

We refer to $\rho((\mathrm{x})) = (x_1, x_2, \ldots, x_s)$ for $(\mathrm{x}) = (x_1, x_2, \ldots, x_{s+m})$ and $\rho([\mathrm{y}]) = (y_1, y_2, \ldots, y_r)$ for $[y_1, y_2, \ldots, y_{s+m}]$ as the colour of the point and the colour of the line respectively. For each $\mathrm{b} \in K^r$ and $(\mathrm{p}) = (p_1, p_2, \ldots, p_{s+m})$ there is a unique neighbour of the point $[\mathrm{l}] = N_{\mathrm{b}}(\mathrm{p})$ with the colour b. Similarly for each $\mathrm{b} \in K^s$ and $[\mathrm{l}] = [l_1, l_2, \ldots, l_{r+m}]$ there is a unique neighbour of the line $[\mathrm{p}] = N_{\mathrm{b}}([\mathrm{l}])$ with the colour b. Let $S(K^n)$ be the semigroup of all polynomial maps from K^n into K^n, where K is a commutative ring.

Assume that the transformation $F(n) \in S(K^n)$ is written in the form $x_j \to f(n)_j(x_1, x_2, \ldots, x_n)$ where each $f(n)_j$, $j = 1, 2, \ldots, n$ is determined by the list of all monomial terms with the respect to some chosen order.

Let us refer to the sequence of maps $F(n)$ from $S(K^n)$, $n = 2, 3, \ldots$ as a family of bounded degree, if the degree of each transformation $F(n)$ is bounded by some constant d, $d > 0$.

Let $\tau(n)_L$ and $\tau(n)_R$ be affine transformations of kind $\mathrm{x} \to \mathrm{x}A + \mathrm{b}$, where $\mathrm{x} \in K^n$, $\mathrm{b} \in K^n$, $A = (a_{ij})$, $1 \leqslant i, j \leqslant n$.

We assume, that the transformations $\tau_L(i)$ and $\tau_R(i)$ are invertible.

We refer to the sequence of $G(n) = \tau_L(n)F(n)\tau_R(n)$ as the deformation of the family $F(n)$, $n = 2, 3, \ldots$.

Notice that $\deg g(n) = \deg f(n)$, but densities of the maps can be different. In fact densities of $g(n)$ heavily depend on the choices of an affine transformation τ_L.

Let us convert the bipartite graph of incidence relation $I = I_m$ to vertex automaton $VA(I_m)$ in the following way. We announce that vertices of the graph are states of $VA(I_m)$. If $(\mathrm{p})I[\mathrm{l})]$ and $[\mathrm{l}] = N_{\mathrm{b}}(\mathrm{p})$ then we draw an arrow from (p) to [l] with the weight $\mathrm{b} \in K^r$. If $(\mathrm{p})I[\mathrm{l})]$ and $[\mathrm{p}] = N_{\mathrm{b}}(\mathrm{p})$ then we draw an arrow from [l] to (p) with the weight $\mathrm{b} \in K^s$. We assume that all vertices of the bipartite graph are accepting states.

Let us assume that $r = s = 1$ in all further considerations. We assume that graph I_m has connectivity invariants $d_1(\mathrm{x}), d_2(\mathrm{x}), \ldots, d_t(\mathrm{x})$ which are multivariate functions from K^{s+m} into K such that for two vertices v_1 and v_2 (points or lines) from the same connected component of the graph equalities $d_i(\mathrm{v}_2) = d_i(\mathrm{v}_1)$, $i = 1, 2, \ldots, t$ hold.

We consider symbolic vertex automaton $SV(I_m)$ corresponding to I_m defined in the following way. Its states are divided into points $(f_1, f_2, \ldots, f_{m+1})$ and lines $[g_1, g_2, \ldots, g_{m+1}]$ where $f_i \in K[x_1, x_2, \ldots, x_{1+m}]$ and $g_i \in K[x_1, x_2, \ldots, x_{1+m}]$, $i = 1, 2, \ldots, m+1$. There are two options for an by initial state: symbolic point $(x_1, x_2, \ldots, x_{1+m})$ or symbolic line $[x_1, x_2, \ldots, x_{1+m}]$. The computation of $SV(I_m)$ is given by its symbolic key $\mathrm{h}_j \in K[z_1, z_2, \ldots, z_{1+t}]$, $j = 1, 2, \ldots, k$ and its initial state (point for example) in the following way. One has to form the specialisation of a symbolic key $\tilde{h}_j = h(x_1, d_1(\mathrm{x}), d_2(\mathrm{x}), \ldots, d_t(\mathrm{x})) \in K[x_1, x_2, \ldots x_{1+m}]$ and compute the chain $(x_1, x_2, \ldots, x_{1+m})$,

$$\begin{aligned} N_{\tilde{h}_1(x_1,x_2,\ldots,x_{1+m})}(\mathrm{x}) &= \mathrm{v}_1, \\ N_{\tilde{h}_2(x_1,x_2,\ldots,x_{1+m})}(\mathrm{v}_1) &= \mathrm{v}_2, \\ N_{\tilde{h}_3(x_1,x_2,\ldots,x_{1+m})}(\mathrm{v}_2) &= \mathrm{v}_3, \\ &\ldots, \\ N_{\tilde{h}_k(x_1,x_2,\ldots,x_{1+m})}(\mathrm{v}_{k-1}) &= \mathrm{v}_k \end{aligned}$$

via symbolic computations. We refer to $F = \mathrm{v}_k$ as a result of symbolic computation with the given symbolic key and refer to a chain (x), v_j, $j = 1, 2, \ldots, k$ as decomposition of v_k into transition function of symbolic automaton $SV(I_m)$. We identify v_k with the corresponding multivariate map from $S(K^{m+1})$.

We refer to the deformation rule $G = \tau_L \mathrm{v}_k \tau_R$ and the chain v_i, $i = 1, 2, \ldots, k$ as decomposition of G of rank k into transition function of symbolic vertex automaton of the graph I_m. We say that G is symbolically decomposed via linguistic graph I_m.

Notice that for $F = (f_1, f_2, \ldots, f_{m+1})$ polynomial f_1 coincides with $h_k(x_1, d_1(\mathrm{x}), d_2(\mathrm{x}), \ldots, d_t(\mathrm{x})$. Let us investigate the equation

$$F(p_1, p_2, \ldots, p_{m+1}) = (b_1, b_2, \ldots, b_{m+1}).$$

Assume that $(b_1, b_2, \ldots, b_{m+1})$ is an element of image of F and p_i are variables. Then $h_k(p_1, d_1(\mathrm{p}), d_2(\mathrm{p}), \ldots, d_t(\mathrm{p})) = b_1$. We can rewrite it as $h_k(p_1, d_1(\mathrm{b}), d_2(\mathrm{b}), \ldots, d_t(\mathrm{b})) = b_1$. Notice that here we use the fact that vertices $(p_1, p_2, \ldots, p_{m+1})$ and $(b_1, b_2, \ldots, b_{m+1})$ (points or lines) are in the same connected component of the graph. Let us assume that for the subset Ω of K the equation $h_k(p_1, d_1(\mathrm{b}), d_2(\mathrm{b}), \ldots, d_t(\mathrm{b})) = b_1$, $p_1 \in \Omega$ has at most one solution. If $\mathrm{b} \in F(\Omega \times K^m)$ then we can find the solution $p_1 = p_1^*$. After that we can compute

$$\begin{aligned}
\beta_{k-1} &= h_{k-1}(p_1^*, d_1(\mathrm{b}), d_2(\mathrm{b}), \ldots, d_t(\mathrm{b})),\\
\beta_{k-2} &= h_{k-2}(p_1^*, d_1(\mathrm{b}), d_2(\mathrm{b}), \ldots, d_t(\mathrm{b})),\\
&\ldots\\
\beta_1 &= h_{k-2}(p_1^*, d_1(\mathrm{b}), d_2(\mathrm{b}), \ldots, d_t(\mathrm{b})).
\end{aligned}$$

It allows us to compute

$$\begin{aligned}
\mathrm{u}_{k-1} &= N_{\beta_{k-1}}(b_1, b_2, \ldots, b_{m+1}),\\
\mathrm{u}_{k-2} &= N_{\beta_{k-2}}(\mathrm{u}_{k-1}),\\
&\ldots\\
\mathrm{u}_1 &= N_{\beta_{k-2}}(\mathrm{u}_2),\\
(p_1^*, p_2^*, \ldots, p_{m+1}^*) &= N_{p_1^*}(\mathrm{u}_1).
\end{aligned}$$

So the restriction of the map F on $\Omega \times K^m$ is injective. The equation $F(\mathrm{x}) = \mathrm{b}$, where $\mathrm{x} \in \Omega \times K^m$, $\mathrm{b} \in F(\Omega \times K^m)$ has a unique solution.

Let $F' = \tau_L F \tau_R$ be the deformation of F and $T = {\tau_L}^{-1}(\Omega \times K^m)$. Then the equation $F'(\mathrm{x}) = \mathrm{b}$ for $\mathrm{x} \in T$ and $\mathrm{b} \in F'(T)$ has a unique solution. We say that the multivariate transformation F' of K^{m+1} is partially invertible on T. Such maps F' together with deformation rule $\tau_L F \tau_R$ and decomposition of F via transition functions of symbolic vertex automaton of linguistic graph can be used in symmetric cryptography. Let us consider two general examples in case $K = Z_l$, $l \geqslant 2$.

Example. Correspondents (Alice and Bob) take a linguistic graph I_m in cases $r = s = 1$ as above. Assume that they know the list of connectivity invariants $d_i(x_1, x_2, \dots, x_{m+1})$, $i = 1, 2, \dots, t$. They choose the type of an initial state. Without loss of generality we can take point $(x_1, x_2, \dots, x_{m+1})$. They set the length of computation of vertex symbolic automaton k and symbolic key

$$h_1(z_1, z_2, \dots, z_{t+1}), h_2(z_1, z_2, \dots, z_{t+1}), \dots h_k(z_1, z_2, \dots, z_{t+1}),$$

where $h_k = ax^r + f(z_2, z_3, \dots, z_{t+1})$, $a \in {Z_l}^*$, $(r, \phi(l)) = 1$. They choose affine transformation τ_L of kind

$$x_1 \to x_1 + x_2 + \dots x_{m+1}, x_j \to l_j(x_1, x_2, \dots, x_{m+1}),$$

where $l_j(x_1, x_2, \dots, x_{m+1})$ are general linear transformation of ${Z_l}^{m+1}$ into Z_l for $j = 2, 3, \dots, m+1$, and general bijective affine transformation τ_R.

We assume that the graph I_m, its connectivity invariants, and the plainspace $T = \{(x_1, x_2, \dots, x_{m+1}) \in {Z_l}^{m+1} | x_1 + x_2 + \dots + x_n \in {Z_l}^*\}$ are known to public. Cryptanalytic knows the general algorithm which depends on some unknown τ_L, τ_R and some symbolic key. Correspondents share the symbolic key $h_i(x_1, x_2, \dots, x_{t+1})$, $i = 1, 2, \dots, k$ and affine transformations τ_L and τ_R as above. Alice writes her plaintext $\mathrm{p} = (p_1, p_2, \dots, p_{m+1})$. She computes the tuple $\tau_L(\mathrm{p}) = (u_1, u_2, \dots, u_{m+1}) = \mathrm{u}$. She computes values of connectivity invariants $\beta_i = d_i(u_1, u_2, \dots, u_{m+1})$, $i = 1, 2, \dots, t$. After that Alice gets the values of symbolic keys

$$\begin{aligned}
\gamma_1 &= h_1(u_1, \beta_1, \beta_2, \dots, \beta_t), \\
\gamma_2 &= h_2(u_1, \beta_1, \beta_2, \dots, \beta_t), \\
&\dots, \\
\gamma_k &= h_k(u_1, \beta_1, \beta_2, \dots, \beta_t).
\end{aligned}$$

If chosen k is odd she takes the chain (u), $N_{\gamma_1}(\mathrm{u}) = [\mathrm{u}^1]$, $N_{\gamma_2}([\mathrm{u}^1]) = (\mathrm{u}^2)$, $\dots$, $N_{\gamma_k}((\mathrm{u}^{k-1})) = [\mathrm{u}^k]$. She takes $\tau_R(\mathrm{u}^k) = \mathrm{c}$ as ciphertext. Notice that in case of even K Alice gets $N_{\gamma_k}([\mathrm{u}^{k-1}]) = (\mathrm{u}^k)$.

Let us consider the decryption process. For simplicity we take the case when k is odd. Bob takes c. He computes ${\tau_R}^{-1}(\mathrm{c}) = \mathrm{u}^k$. He takes $[\mathrm{u}^k] = [b_1, b_2, \dots, b_n]$. Bob computes parameters β_i as $d_i([b_1, b_2, \dots, b_n])$ for $i = 1, 2, \dots, t$.

Bob looks at expression $ax^r + f(z_2, z_3, \dots, z_{t+1})$ and writes the equation $ax^r + f(\beta_1, \beta_2, \dots, \beta_t) = b_1$. So he computes $x^r =$

$(b_1 - f(\beta_1, \beta_2, \ldots, \beta_t))a^{-1} = \alpha$. So Bob gets u_1 as $\alpha^{r'}$ where r' is a multiplicative inverse in $Z_{\phi(l)}$.

Now, Bob computes $\gamma_i = h_i(u_1, \beta_1, \beta_2, \ldots, \beta_t)$, $i = 1, 2, \ldots, k-1$. So he gets

$$N_{\gamma_{k-1}}([\mathrm{u}^k]) = (\mathrm{u}^{k-1}), \quad N_{\gamma_{k-2}}((\mathrm{u}^{k-1})) = [\mathrm{u}^{k-2}], \ldots,$$
$$N_{\gamma_1}((\mathrm{u}^2)) = [\mathrm{u}^1], \quad N_{u_1}([\mathrm{u}^1]) = (\mathrm{u}).$$

Finally Bob obtains ${\tau_L}^{-1}(\mathrm{u}) = (p_1, p_2, \ldots, p_s)$.

Remark 1. It is easy to see that the scheme above can be easily modified in various ways. For instance, correspondents can use $T = {Z_l}^{*m+1}$ and take τ_L as linear monomial transformation $(x_1, x_2, \ldots, x_{m+1}) \to (\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_{m+1} x_{m+1})$, where $(\lambda_1, \lambda_2, \ldots, \lambda_{m+1}) \in {Z_l}^{*m+1}$.

Remark 2. The above scheme can produce rather fast symmetric encryption algorithm in case of various linguistic graphs. It is easy to define linguistic graph I_m such that the neighbour of each vertex can be computed in time $O(m)$. We can take an empty list of connectivity invariants (parameter t is zero). Assume that we work with sparse affine transformation τ_L and τ_R which can be completed in $O(m)$ elementary steps. Then the encryption algorithm above takes $O(m)$ operations.

Towards public key algorithm. Alice can take a linguistic graph I_m in case $r = s = 1$ as above. She knows the list of connectivity invariants $d_i(x_1, x_2, \ldots, x_{m+1})$, $i = 1, 2, \ldots, t$. She chooses the type of initial state. Without loss of generality we can take point $(x_1, x_2, \ldots, x_{m+1})$. Alice chooses the length k of computation of vertex symbolic automaton k and symbolic key

$$h_1(z_1, z_2, \ldots, z_{t+1}), h_2(z_1, z_2, \ldots, z_{t+1}), \ldots, h_k(z_1, z_2, \ldots, z_{t+1}),$$

where $h_k = ax^r + f(z_2, z_3, \ldots, z_{t+1})$, $a \in {Z_l}^*$, $(r, \phi(l)) = 1$. She chooses affine transformation τ_L of kind

$$x_1 \to x_1 + x_2 + \ldots x_{m+1}, x_j \to l_j(x_1, x_2, \ldots, x_{m+1}),$$

where $l_j(x_1, x_2, \ldots, x_{m+1})$ are general linear transformation of ${Z_l}^{m+1}$ into Z_l for $j = 2, 3, \ldots, m+1$, and general bijective affine transformation τ_R.

Alice takes the initial state $\mathrm{x} = (x_1, x_2, \ldots, x_{m+1})$. She computes the tuple $\tau_L(\mathrm{x}) = (v_1, v_2, \ldots, v_{m+1}) = \mathrm{v}$, where v_i are linear expressions in variables $x_1, x_2, \ldots, x_{m+1}$. Notice that $v_1 = x_1 + x_2 + \cdots + x_{m+1}$. After that

Alice takes computation of symbolic vertex automaton with symbolic key $h_i, i = 1, 2, \ldots, k$ starting in a new initial state $(v_1, v_2, \ldots, v_{m+1})$. It means that Alice uses symbolic computations for the constructions of multivariate invariants $d_t(v_1, v_2, \ldots, v_{m+1}) = d'_t(x_1, x_2, \ldots, x_{m+1})$, $i = 1, 2, \ldots, t$.

She computes $\tilde{h_1} = h_1(v_1, d'_2, \ldots, d'_{t+1})$, $\tilde{h_2} = h_2(v_1, d'_2, \ldots, d'_{t+1})$, $\ldots$ $\tilde{h_k} = h_k(v_1, d'_2, \ldots, d'_{t+1})$.

Alice computes the chain of elements from $Z_l[x_1, x_2, \ldots, x_{m+1}]^{m+1}$ (vertices of symbolic automaton, points and lines). The point $\mathrm{v} = (v_1, v_2, \ldots, v_{m+1})$, line $[\mathrm{v}_1] = N_{\tilde{h_1}}(\mathrm{v})$, point $(\mathrm{v}_2) = N_{\tilde{h_2}}([\mathrm{v}_1])$, $\ldots$, $(\mathrm{v}_{k-1}) = N_{\tilde{h_{k-1}}}((\mathrm{v}_{k-2}))$, $[\mathrm{v}_k] = N_{\tilde{h_k}}(\mathrm{v}_{k-1})$. For simplicity we take odd k. Alice treats $F = \mathrm{v}_k$ as multivariate map and computes $G = F\tau_R$ (composition of two maps).

Assume that Alice can complete all steps as above in polynomial time and get a resulting map G of finite degree. Then she can write the standard form of G: $x_1 \to g_1(x_1, x_2, \ldots, x_{m+1})$, $x_2 \to g_2(x_1, x_2, \ldots, x_{m+1})$, $\ldots$, $x_{m+1} \to g_{m+1}(x_1, x_2, \ldots, x_{m+1})$, where g_i, $i = 1, 2, \ldots, m+1$ are given by the lists of their monomial terms with respect to some standard order.

Then Alice can announce the public rules $g_i \in Z_l[x_1, x_2, \ldots, x_{m+1}]$, $i = 1, 2, \ldots, m+1$ to all of her correspondents together with the plainspace $\Omega_{m+1} = \{\mathrm{x} \in {Z_l}^{m+1} | x_1 + x_2 + \cdots + x_m \in {Z_l}^*\}$.

Public user (Bob) writes a message $(p_1, p_2, \ldots, p_{m+1}) \in \Omega_{m+1}$ and computes the ciphertext $(c_1, c_2, \ldots, c_{m+1})$ where $c_i = g_i(p_1, p_2, \ldots, p_{m+1})$, $i = 1, 2, \ldots, m+1$ and sends it to Alice.

Alice knows the deformation rule $G = \tau_L F \tau_R$ and the symbolic key which gives the decomposition of F into transition functions of the symbolic vertex automaton of the graph. So she can use the decryption process of symmetric encryption algorithm above and restore the plaintext $(p_1, p_2, \ldots, p_{m+1})$.

Remark 3. Similarly to symmetric algorithm Alice can change Ω_{m+1} for $T = {Z_l}^{*m+1}$ and take τ_L as linear monomial transformation

$$(x_1, x_2, \ldots, x_{m+1}) \to (\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_{m+1} x_{m+1}),$$

where $(\lambda_1, \lambda_2, \ldots, \lambda_{m+1}) \in {Z_l}^{*m+1}$.

Remark 4. One can assume that cryptanalytic knows the family of graphs I_m defined over Z_l, where l is known composite number.

We introduce free symbolic computation of odd case k for the general linguistic graph I_m over commutative ring K in case $r = s = 1$ as the

sequence $\mathrm{x} = (x_1, x_2, \ldots, x_{m+1})$ (initial state), line

$$\begin{aligned} N_{z_1}(\mathrm{x}) &= [\mathrm{u}_1], \quad \mathrm{u}_1 \in K[z_1, x_1, x_2, \ldots, x_{m+1}]^{m+1}, \\ N_{z_2}([\mathrm{u}_1]) &= (\mathrm{u}_2), \quad \mathrm{u}_2 \in K[z_1, z_2, x_1, x_2, \ldots, x_{m+1}]^{m+1}, \\ &\ldots \\ N_{z_k}((\mathrm{u}_{k-1})) &= [\mathrm{u}_k)], \quad \mathrm{u}_k \in K[z_1, z_2, \ldots, z_k, x_1, x_2, \ldots, x_{m+1}]^{m+1}. \end{aligned}$$

3. On some extremal algebraic graphs

Recall that the girth is the length of minimal cycle in the simple graph. Studies of maximal size $ex(C_3, C_4, \ldots, C_{2m}, v)$ of the simple graph on v vertices without cycles of length $3, 4, \ldots, 2m$, i. e. graphs of girth $> 2m$, form an important direction of Extremal Graph Theory.

As it follows from the famous Even Circuit Theorem by P. Erdős' we have inequality

$$ex(C_3, C_4, \ldots, C_{2n}, v) \leqslant cv^{1+1/n},$$

where c is a certain constant. The bound is known to be sharp only for $2n = 4, 6, 8$. The first general lower bounds of kind $ex(v, C_3, C_4, \ldots C_n) = \Omega(v^{1+c/n})$, where c is some constant $< 1/2$ were obtained in the 50th by Erdős' via studies of *families of graphs of a large girth*, i.e. infinite families of simple regular graphs Γ_i of degree k_i and order v_i such that $g(\Gamma_i) \geqslant c\log_{k_i} v_i$, where c is the independent of i constant. Erdős' proved the existence of such a family with arbitrary large but bounded degree $k_i = k$ with $c = 1/4$ by his famous probabilistic method.

One of the first examples of the family of graphs of large girth is the family of algebraic graphs $CD(n, q)$ (see [15] and further references). Graphs $CD(n, q)$ appear as connected components of graphs $D(n, q)$ defined via system of quadratic equations [16].

Graphs $D(n, q)$ and $CD(n, q)$ have been used in symmetric cryptography together with their natural analogs $D(n, K)$ and $CD(n, K)$ over general finite commutative rings K since 1998 (see [17]). The theory of directed graphs and language of dynamical system were very useful for studies of public key and private key algorithms based on graphs $D(n, K)$, $CD(n, K)$ (see [18–25] and further references).

There are several implementations of symmetric algorithms for cases of fields ([26], [27], [30]) and arithmetical rings ([28], [29]). Some comparison of bijective multivariate maps based on $D(n, K)$ and other graphs $A(n, K)$ are considered in [31].

4. Graphs $D(n, K)$ and new algorithms related to them

Let P and L be two copies of Cartesian power $\mathbb{K}^{\mathbb{N}}$, where $\mathbb{K}$ is the commutative ring and $\mathbb{N}$ is the set of positive integer numbers. Elements of P will be called *points* and these of L *lines*.

To distinguish points from lines we use parentheses and brackets. If $x \in V$, then $(x) \in P$ and $[x] \in L$. It will also be advantageous to adopt the notation for co-ordinates of points and lines introduced in [16] for the case of general commutative ring $\mathbb{K}$:

$$(p) = (p_{0,1}, p_{1,1}, p_{1,2}, p_{2,1}, p_{2,2}, p'_{2,2}, p_{2,3}, \ldots, p_{i,i}, p'_{i,i}, p_{i,i+1}, p_{i+1,i}, \ldots),$$

$$[l] = [l_{1,0}, l_{1,1}, l_{1,2}, l_{2,1}, l_{2,2}, l'_{2,2}, l_{2,3}, \ldots, l_{i,i}, l'_{i,i}, l_{i,i+1}, l_{i+1,i}, \ldots].$$

The elements of P and L can be thought as infinite ordered tuples of elements from $\mathbb{K}$, such that only finite number of components are different from zero.

We now define a linguistic incidence structure (P, L, I) defined by infinite system of equations as follows. We say the point (p) is incident with the line $[l]$, and we write $(p)I[l]$, if the following relations between their co-ordinates hold:

$$\begin{aligned} l_{i,i} - p_{i,i} &= l_{1,0}p_{i-1,i}, \\ l'_{i,i} - p'_{i,i} &= l_{i,i-1}p_{0,1}, \\ l_{i,i+1} - p_{i,i+1} &= l_{i,i}p_{0,1}, \\ l_{i+1,i} - p_{i+1,i} &= l_{1,0}p'_{i,i}. \end{aligned} \tag{1}$$

(These four relations are defined for $i \geqslant 1$, $p'_{1,1} = p_{1,1}$, $l'_{1,1} = l_{1,1}$). The incidence structure (P, L, I) we denote as $D(\mathbb{K})$. We speak now of the *incidence graph* of (P, L, I), which has the vertex set $P \cup L$ and edge set consisting of all pairs $\{(p), [l]\}$ for which $(p)I[l]$.

For each positive integer $k \geqslant 2$ we obtain a symplectic quotient (P_k, L_k, I_k) as follows. Firstly, P_k and L_k are obtained from P and L, respectively, by simply projecting each vector into its k initial coordinates. The incidence I_k is then defined by imposing the first $k-1$ incidence relations and ignoring all others. The incidence graph corresponding to the structure (P_k, L_k, I_k) is denoted by $D(k, \mathbb{K})$ (see [17]).

To facilitate notation in the future results on "connectivity invariants", it will be convenient for us to define $p_{-1,0} = l_{0,-1} = p_{1,0} = l_{0,1} = 0$, $p_{0,0} = l_{0,0} = -1$, $p'_{0,0} = l'_{0,0} = -1$, $p'_{1,1} = p_{1,1}, l'_{1,1} = l_{1,1}$ and to assume that (1) are defined for $i \geqslant 0$.

Notice, that for $i = 0$, the four conditions (6) are satisfied by every point and line, and, for $i = 1$, the first two equations coincide and give $l_{1,1} - p_{1,1} = l_{1,0}p_{0,1}$.

Let $k \geqslant 6$, $t = [(k+2)/4]$, and let $u = (u_\alpha, u_{11}, \cdots, u_{tt}, u'_{tt}, u_{t,t+1}, u_{t+1,t}, \cdots)$ be a vertex of $\mathrm{D}(k, \mathbb{K})$ ($\alpha \in \{(1,0),(0,1)\}$, it does not matter whether u is a point or a line). For every r, $2 \leqslant r \leqslant t$, let

$$a_r = a_r(u) = \sum_{i=0,r} (u_{ii}u'_{r-i,r-i} - u_{i,i+1}u_{r-i,r-i-1}),$$

and $a = a(u) = (a_2, a_3, \cdots, a_t)$. Similarly, we assume $a = a(u) = (a_2, a_3, \cdots, a_t, \dots)$ for the vertex u of infinite graph $D(\mathbb{K})$.

Let η_n (η) be the equivalence relation:

$$u\eta_n v \Leftrightarrow a(u) = a(v) \quad (u\tau v \Leftrightarrow a(u) = a(v))$$

on the vertex set of graph $D(k, \mathbb{K})$ ($D(\mathbb{K})$), respectively.

Proposition 1 (see [19] and further references)**.**

(i) *For any $t-1$ ring elements $x_t \in \mathbb{K}$, $2 \leqslant t \leqslant [(k+2)/4]$, there exists a vertex v of $\mathrm{D}(n, \mathbb{K})$ for which $a(v) = (x_2, \dots, x_t) = (x)$.*

(ii) *The equivalence class C_n for the equivalence relation τ on the set $\mathbb{K}^n \cup \mathbb{K}^n$ is an isomorphic to the affine variety $\mathbb{K}^t \cup \mathbb{K}^t$, $t = [4/3n]+1$ for $n = 0, 2, 3 \bmod 4$, $t = [4/3n] + 2$ for $n = 1 \bmod 4$.*

(iii) *the vertex set $\mathrm{C_n}$ is the union of several connected components of $\mathrm{D}(n, \mathbb{K})$.*

Let C be the equivalence class on τ on the vertex set $\mathrm{D}(\mathbb{K})$, then the induced subgraph with the vertex set C is the union of several connected components of $\mathrm{D}(\mathbb{K})$.

We shall use notation $C(t, \mathbb{K})$ ($C(\mathbb{K})$) for the induced subgraph of $D(n, \mathbb{K})$ ($D(\mathbb{K})$) with the vertex set C_n (vertex set C respectively).

The graph $C(t, \mathbb{K})$ in the case of $\mathbb{K} = \mathbb{F}_q$ coincides with $CD(n, q)$ which was introduced in [17].

The following statement was proven in [32].

Theorem 1. *Let $\mathbb{K}$ be commutative ring with unity of characteristic d, $d \neq 2$. Then graphs $C(t, \mathbb{K})$, $t \geqslant 6$ and $C(\mathbb{K})$ are connected.*

If $\mathbb{K} = \mathbb{F}_q$, q is odd, then graph $C(\mathbb{F}_q)$ is a q-regular tree. In cases $\mathrm{char}(\mathbb{K}) = 2$ the questions of the description of connected components of $C(t, \mathbb{K})$ and $C(\mathbb{K})$ are open.

5. The cryptosystem

We can rewrite result of [33] in the following form.

Proposition 2. *Let F_n be a regular computation of free symbolic automaton of linguistic graph $D(n, Z_l)$ and $\alpha_1, \alpha_2, \ldots, \alpha_k$, where k is even, are fixed elements of Z_l. Then the map $\tilde{F}_n$ corresponding to a specialisation of $z_2 = y+\alpha_1$, $z_3 = z_1+\alpha_1$, $z_4 = y+\alpha_3$, $z_5 = z_1+\alpha_5$, $\ldots$, $z_{k-1} = z_1+\alpha_{k-1}$, $z_k = y + \alpha_k$ is cubical multivariate map from $K[z_1, y, x_1, x_2, \ldots, x_n]^{m+1}$.*

Remark 5. Similar proposition is true for odd k. The map $\tilde{F}_n$ corresponding to a specialisation of $z_2 = y + \alpha_1$, $z_3 = z_1 + \alpha_1$, $z_4 = y + \alpha_3$, $z_5 = z_1 + \alpha_5$, $\ldots$, $z_{k-1} = y + \alpha_{k-1}$, $z_k = z_1 + \alpha_k$ is cubical transformation of ${Z_l}^n$.

Proposition 3. *Let F_n be a regular computation of an odd length s of a symbolic vertex automaton of $D(n.K)$ corresponding to symbolic key $h(z_1, z_2, \ldots, z_t)+\alpha_1$, $z_1+\alpha_2$, $h(z_1, z_2, \ldots, z_t)+\alpha_3$, $z_1+\alpha_4$. $\ldots$, $z_1+\alpha_{s-1}$, $h(z_1, z_2, \ldots, z_t) + \alpha_s$, where $h \in K[z_1, z_2, \ldots, z_t]$ has finite degree and α_i, $i = 1, 2, \ldots, s$ are constants from K. Then the degree of F_n is bounded by $3 \deg h(x_{01}, a_2(\mathrm{x}), a_3(\mathrm{x}), \ldots, a_t(\mathrm{x}))$.*

We say that the map F_n of ${Z_l}^n$ to itself is Eulerian partially invertible map on the domain $\Omega_n = \{\mathrm{x} | \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n + \alpha_{n+1} \in {Z_l}^*\}$ if it is partially invertible on Ω_n and solution of equation $F_n(\mathrm{x}) = \mathrm{b}$, $\mathrm{x} \in \Omega$ and $\mathrm{b} \in F_n(\Omega_n)$ can be reduced to a solution of $z^r = a$, $z \in {Z_l}^*$, $r \neq 1$, $(r, \phi(l)) = 1$.

Theorem 2. *Let $K = Z_l$, n be a natural number $\geqslant 2$, s is an odd number $\geqslant 3$. For each domain of kind $\Omega_n = \{\mathrm{x} | \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n + \lambda_{n+1} \in {Z_l}^*\}$ in ${Z_l}^n$, where $\lambda_i \neq 0$, $i = 1, 2, \ldots, n$ there is Eulerian map F_n of finite degree which has a symbolic decomposition of rank s. If l is a prime number, then Eulerian map F_n is a bijection.*

Proof. Let us consider a symbolic vertex automaton constructed for the family of graphs $D(n, Z_l)$. Let $a_2(\mathrm{x})$, $a_3(\mathrm{x})$, $\ldots$, $a_t(\mathrm{x})$, $t = [(n + 2)/4]$ be the list of quadratic connectivity invariants of the graph. We shall use polynomials from $Z_l[u_1, u_2, \ldots, u_t]$ to form special symbolic key. For $f \in Z_l[u_1, u_2, \ldots, u_t]$ we define $\tilde{f}$ as $f(z_1, a_2(\mathrm{z}), a_3(\mathrm{z}), \ldots, a_t(\mathrm{z}))$, where $(\mathrm{z}) = (z_1, z_2, \ldots, z_n)$ is initial point of the symbolic vertex automaton of graph $D(n, Z_l)$. We avoid double indexes for points and lines here. We have a free choice to take $H \in Z_l[u_1, u_2, \ldots, u_t]$ to form a sequence of

weights $\alpha_1(\mathrm{z}) = \tilde{H} + \beta_1$, $\alpha_2(\mathrm{z}) = z_1 + \beta_2$, $\alpha_3(\mathrm{z}) = \tilde{H} + \beta_3$, $\alpha_4(\mathrm{z}) = z_1 + \beta_4$, $\ldots$, $\alpha_{s-1}(\mathrm{z}) = z_1 + \beta_{s-1}$, $\alpha_s(\mathrm{z}) = \tilde{H} + \beta_s$, where β_i, $i = 1, 2, \ldots, s$ are fixed elements of Z_l. Let $F = F_n : {Z_l}^n \to {Z_l}^n$ be the multivariate map generated by symbolic computation above. We assume that $H(u_1, u_2, \ldots, u_t)$ is written in the form ${u_1}^r + S(u_2, u_3, \ldots, u_t)$, where S is arbitrary element of $Z_l[u_2, u_3, \ldots, u_t]$ and $r, r \neq 1$ is a parameter such that $(r, \phi(m)) = 1$. Symbol ϕ standardly stands for Euler function. Let us consider non-singular linear transformation $\tau_L : {Z_l}^n \to {Z_l}^n$ of kind

$$\begin{aligned} z_1 &\to \lambda_1 z_1 + \lambda_2 z_2 + \cdots + \lambda_n z_n + \lambda_{n+1}, \\ z_2 &\to l_2(z_1, z_2, \ldots, z_n), \\ z_3 &\to l_3(z_1, z_2, \ldots, z_n), \\ &\cdots \\ z_n &\to l_n(z_1, z_2, \ldots, z_n), \end{aligned}$$

where l_i are linear expressions from $Z_l[z_1, z_2, \ldots, z_n]$ of general kind. We form a composition $G_n = \tau_L F_n$.

Assume that $\mathrm{z} = (z_1, z_2, \ldots, z_n)$ is an element of Ω_n. Let us identify $\tau_L(\mathrm{z}) = (y_1, y_2, \ldots, y_n)$ with the point of the graph $D(n, Z_l)$. Notice that $y_1 \in {Z_l}^*$. Let us show that the reimage of $G_n(\mathrm{z})$ is uniquely determined. We write the equation $G_n(\mathrm{z}) = (b_1, b_2, \ldots, b_n)$. It is clear that $b_1 = {y_1}^r + S(u_2, u_3, \ldots, u_t) + \beta_s$. Notice that tuples y (point) and b (line) are located in the same connected component of the graph. So we have $a_i(\mathrm{y}) = a_i(\mathrm{b}) = \gamma_i$, $i = 2, 3, \ldots, t$. Thus ${y_1}^r + S(\gamma_2, \gamma_3, \ldots, \gamma_t) + \beta_s = b_1$.

Ler r' be the multiplicative inverse of r in $Z_{\phi(l)}$. We have $y_1 = (b_1 - S(\gamma_2, \gamma_3, \ldots, \gamma_t) - \beta_s)^{r'} = \alpha$.

The knowledge of parameter α allows us to compute all coordinates of tuple y. Really, we can compute values $\alpha_{s-1} = \alpha + \beta_{s-1}$, $\alpha_{s-2} = H(\alpha, \gamma_2, \gamma_3, \ldots, \gamma_t) + \beta_{s-2}, \alpha_{s-3} = \alpha + \beta_{s-3}, \ldots, \alpha_1 = H(\alpha, \gamma_2, \gamma_3, \ldots, \gamma_t) + \beta_1$, $\alpha_0 = \alpha$.

The value of y can be computed recursively $y^{s-1} = N_{\alpha_{s-1}}([\mathrm{b}])$, $y^{s-2} = N_{\alpha_{s-2}}((y^{s-1}))$, $\ldots$ $y^1 = N_{\alpha_1}((y^2))$, $y^0 = N_\alpha((y^1)) = (y_1, y_2, \ldots, y_n)$. The tuple z equals ${\tau_l}^{-1}(y^0)$.

The Proposition 3 establishes that the degree of G_n or F_n is bounded by $3\deg(\tilde{H}(\mathrm{z}))$. If $d = \deg(\tilde{S}) > r$ then the degree of G_n is bounded by $3d$. Notice, that in case of prime l the equation ${y_1}^r + S(\gamma_2, \gamma_3, \ldots, \gamma_t) + \beta_s = b_1$, $r \neq 0 \bmod p-1$ is always solvable for y_1. So maps F_n and G_n are bijections. □

Remark 6. In the theorem above we can change domain Ω_n for ${Z_l}^{*n}$.

Really we have to change a transformation τ_L in the proof for a linear monomial map $(x_1, x_2, \ldots, x_n) \to (\lambda_1 x_{\pi(1)}, \lambda_2 x_{\pi(2)}, \ldots, \lambda_n x_{\pi(n)})$, where λ_i, $i = 1, 2, \ldots, n$ are elements of $Z_l{}^*$ and π is a permutation from S_n.

The cryptosystem. Assume that Alice is the holder of a public key based on the family of maps used in the constructive proof of the previous theorem. So she takes l, $l \geqslant 2$ and parameter r, such that $(r, \phi(l)) = 1$. She chooses the odd length s, $s \geqslant 3$ of symbolic key for practical use we set size $O(n)$ for value of s. For example, Alice chooses the area $\Omega_n = \{\mathrm{x}|\mathrm{x}_1 + \mathrm{x}_2 + \cdots + \mathrm{x}_\mathrm{n} \in \mathrm{Z}_\mathrm{l}{}^*\}$ which will be a domain for Eulerian map of $G = Z_l{}^n$. Alice has a rather wide choice to pick the function $S \in Z_l[u_2, u_3, \ldots, u_t]$, $t = [(n+2)/4]$ and parameters $\beta_1, \beta_2, \ldots, \beta_s$ to form the symbolic key. She has set $l_1 = x_1 + x_2 + \ldots, x_n$ and may choose various linear functions $l_i \in Z_l[x_1, x_3, \ldots, x_n]$, $i = 2, 3, \ldots, n$ to form bijective affine map τ_l of $Z_l{}^n$ to itself. Finally, she has a free choice for another affine map τ_R.

So in polynomial time Alice generates map F_n via computation of symbolic vertex automaton of linguistic graph $D(n, Z_l)$ with the symbolic key: $\alpha_1(\mathrm{z}) = \tilde{H} + \beta_1$, $\alpha_2(\mathrm{z}) = z_1 + \beta_2$, $\alpha_3(\mathrm{z}) = \tilde{H} + \beta_3$, $\alpha_4(\mathrm{z}) = z_1 + \beta_4$, $\ldots$, $\alpha_{s-1}(\mathrm{z}) = z_1 + \beta_{s-1}$, $\alpha_s(\mathrm{z}) = \tilde{H} + \beta_s$, where β_i, $i = 1, 2, \ldots, s$ are fixed elements of Z_l. She computes the deformation $G_n = \tau_L F_n \tau_R$ in standard form $x_1 \to g_1(x_1, x_2, \ldots, x_n)$, $x_2 \to g_2(x_1, x_2, \ldots, x_n)$, $\ldots$, $x_n \to g_n(x_1, x_2, \ldots, x_n)$, where g_i, $i = 1, 2, \ldots, n$ are given by list of their monomial terms in some chosen order. Notice, that the degree of G_n is bounded by constant.

Alice announces the public the standard form of G_n and keeps data described above in secret. Cryptanalytic knows used graph and general form of a symbolic key.

Assume that a public user (Bob) creates an open text $\mathrm{p} = (p_1, p_2, \ldots, p_n)$. He computes $G_n((p_1, p_2, \ldots, p_n)) = (c_1, c_2, \ldots, c_n)$. Bounded degree of G_n insures that the computation of ciphertext can be computed in a polynomial time $O(n^c)$ for some positive constant c.

The knowledge of deformation rule $G_n = \tau_L F_n \tau_R$ and the docomposition of F_n into transition functions of symbolic vertex automaton of $D(n, Z_l)$ allows her to decrypt in polynomial time with the algorithm described in a previous section.

Remark 7. Alice can use $Z_l{}^{*n}$ instead of $\Omega_n = \{\mathrm{x}|\mathrm{x}_1 + \mathrm{x}_2 + \cdots + \mathrm{x}_\mathrm{n} \in \mathrm{Z}_\mathrm{l}{}^*\}$. In this case τ_L has to be chosen as monomial transformation.

Remark 8. In case of prime l we can change function $H + b_s$ for much more sophisticated expression. For instance $Z(x_2, x_3, \ldots, x_t) f(x_1) +$

$S(x_2, x_3, \ldots, x_t)$ where $Z(x_2, x_3, \ldots, x_t) = 0$ has no solution but $f(x_1) = d$ has exactly one solution in variable x_1 for each d.

Let $h(x) \in Z_p[x]$ has no linear divisors. Then $Z(x_2, x_3, \ldots, x_t) = h(M(x_2, x_3, \ldots, x_t)))$ is always different from zero for each $M \in K[x_2, x_3, \ldots, x_t]$.

The simplest case where we can use $M(x_2, x_3, \ldots, x_t)({x_1}^r) + S(x_2, x_3, \ldots, x_t)$, where $(r, p-1) = 1$ and the equation $M(x_2, x_3, \ldots, x_t) = 0$ has no solution. We say that such a cryptosystem is based on *hidden discrete logarithm problem*. For general parameter l we use the term hidden Eulerian equation. We can use recurrent expressions

$$M_k(\ldots(M_2(M_1(x_2, x_3, \ldots, x_t)({x_1}^{r_1}) + S_1(x_2, x_3, \ldots, x_t))^{r_2} + S_2(x_2, x_3, \ldots, x_t)) + \ldots M_{k-1}(x_2, x_3, \ldots, x_t)({x_1}^{r_{k-1}}) + S_{k-1}(x_2, x_3, \ldots, x_t))^{r_k} + S_k(x_2, x_3, \ldots, x_t)),$$

where $M_i(x_2, x_3, \ldots, x_t) = 0$ have no solutions for each $i = 1, 2, \ldots, k$.

References

[1] J. Ding, J. E. Gower, D. S. Schmidt, *Multivariate Public Key Cryptosystems*, 260. Springer, Advances in Information Security, v. 25, (2006).

[2] V. A. Ustimenko, *Explicit constructions of extremal graphs and new multivariate cryptosystems*, Studia Scientiarum Mathematicarum Hungarica, Special issue "Proceedings of The Central European Conference, 2014, Budapest",volume 52, issue, June 2015, pp. 185-204.

[3] V. A. Ustimenko, *Graphs with Special Arcs and Cryptography*, Acta Applicandae Mathematicae, vol. 71, N2, November 2002, 117-153.

[4] V. Ustimenko, *On the extremal graph theory for directed graphs and its cryptographical applications*, In: T. Shaska, W.C. Huffman, D. Joener and V. Ustimenko. Advances in Coding Theory and Cryptography. Series on Coding and Cryptology, V. 3, 2007, P. 181–200.

[5] V. A. Ustimenko, *On the flag geometry of simple group of Lie type and Multivariate Cryptography*, Algebra and Discrete Mathematics. V. 19. No 1. 2015. P. 130-144.

[6] V. Ustimenko, *On walks of variable length in Schubert incidence systems and multivariate flow ciphers*, Dopovidi of Nathional Acad. Sci. of Ukraine, 2014, No 3, P. 55 - 150.

[7] N. Koblitz, Algebraic aspects of cryptography, Springer (1998).

[8] V. Ustimenko, *On Multivariate Cryptosystems Based on Computable Maps with Invertible Decompositions*, Annales of UMCS, Informatica, volume 14 (2014) , Special issue "Proceedings of International Conference Cryptography and Security Systems", pp. 7-18.

[9] J. Patarin, *The Oil and Vinegar digital signatures*, Dagstuhl Workshop on Cryptography. 1997.

[10] Kipnis A., Shamir A., *Cryptanalisis of the Oil and Vinegar Signature Scheme*, Advances in Cryptology - Crypto 96, Lecture Notes in Computer Science, V. 1462, 1996, P. 257–266.

[11] S. Bulygin, A. Petzoldt, and J. Buchmann, *Towards provable security of the unbalanced oil and vinegar signature scheme under direct attacks*, In Guang Gong and Kishan Chand Gupta, editors, "Progress in Cryptology - INDOCRYPT", Guang Gong and Kishan Chand Gupta, editors, Lecture notes in Computer Science, V. 6498, 2010. P. 17–32.

[12] U. Romanczuk-Polubiec, V. Ustimenko, *On two windows multivariate cryptosystem depending on random parameters*, Algebra and Discrete Mathematics, 2015, V. 19. No. 1. P. 101–129.

[13] F. Harary, *Graph Theory*, Addison-Wesley Publishing Co, Reading, MA (1966).

[14] V. Ustimenko, *Maximality of affine group, hidden graph cryptosystem and graph's stream ciphers*, Journal of Algebra and Discrete Mathematics, 2005, v.1, pp 51-65.

[15] F.Lazebnik , V. Ustimenko and A.J.Woldar, *A new series of dense graphs of high girth*, Bulletin of the AMS 32 (1) (1995), 73-79.

[16] F. Lazebnik, V. Ustimenko, *Explicit construction of graphs with arbitrary large girth and of large size*, Discrete Applied Mathematics 60 (1995), 275-284.

[17] V. Ustimenko, *Coordinatisation of Trees and their Quotients*, in the Voronoj's Impact on Modern Science, Kiev, Institute of Mathematics, 1998, vol. 2, 125-152.

[18] V. Ustimenko, *CRYPTIM: Graphs as Tools for Symmetric Encryption*, Lecture Notes in Computer Science, Springer, LNCS 2227, Proceedings of AAECC-14 Symposium on Applied Algebra, Algebraic Algorithms and Error Correction Codes, November 2001, p. 278-286.

[19] V. Ustimenko, *Linguistic Dynamical Systems, Graphs of Large Girth and Cryptography*, Journal of Mathematical Sciences, Springer, vol.140, N3 (2007) pp. 412-434.

[20] V. Ustimenko, *On the graph based cryptography and symbolic computations*, Serdica Journal of Computing, Proceedings of International Conference on Application of Computer Algebra, ACA-2006, Varna, N1 (2007).

[21] V. Ustimenko, U. Romanczuk, *On Extremal Graph Theory, Explicit Algebraic Constructions of Extremal Graphs and Corresponding Turing Encryption Machines*, in "Artificial Intelligence, Evolutionary Computing and Metaheuristics ", In the footsteps of Alan Turing Series: Studies in Computational Intelligence, Vol. 427, Springer, January, 2013, 257-285.

[22] V. Ustimenko, U. Romanczuk, *On Dynamical Systems of Large Girth or Cycle Indicator and their applications to Multivariate Cryptography*, in "Artificial Intelligence, Evolutionary Computing and Metaheuristics ", In the footsteps of Alan Turing Series: Studies in Computational Intelligence, Volume 427/2012, 257-285.

[23] V. Ustimenko, *On the cryptographical properties of extreme algebraic graphs*, in "Algebraic Aspects of Digital Communications" IOS Press (Lectures of Advanced NATO Institute, NATO Science for Peace and Security Series - D: Information and Communication Security, Volume 24, July 2009, 296 pp.

[24] U. Romanczuk-Polubiec, V. Ustimenko, *On Multivariate Cryptosystems Based on Polynomially Compressed Maps with Invertible Decompositions*, Cryptography and Security Systems, Third International Conference, CSS 2014, Lublin,

Poland, September 22-24, 2014. Proceedings, Communications in Computer and Information Science, 448, p. 23-37.

[25] M. Klisowski, V. Ustimenko, *Graph based cubical multivariate maps and their cryptographical applications*, in "Advances on Superelliptic curves and their Applicationsions", IOS Press, NATO Science for Peace and Security series –D: Information and Communication Security, 2015, v. 41, 201, pp. 305-327.

[26] A. Tousene, V. Ustimenko, *CRYPTALL - a System to Encrypt All types of Data*, Notices of Kiev-Mohyla Academy, v. 23, 2004, pp. 12-15.

[27] A. Touzene, V. Ustimenko, *Graph Based Private Key Crypto System*, International Journal on Computer Research, Nova Science Publisher, v. 13 (2006), issue 4, 12pp.

[28] J. Kotorowicz, V. Ustimenko, *On the implementation of cryptoalgorithms based on algebraic graphs over some commutative rings*, Condenced Matters Physics, Special Issue: Proceedings of the international conferences "Infinite particle systems, Complex systems theory and its application", Kazimerz Dolny, Poland, 2006, 11 (no. 2(54)) (2008) 347–360.

[29] V. Ustimenko, S. Kotorowicz, *On the properties of Stream Ciphers Based on Extremal Directed graphs*, In "Cryptography Research Perspectives", Nova Publishers, Ronald E. Chen (the editor), 2008.

[30] , A. Touzene, V. Ustimenko, Marwa AlRaisi, Imene Boudelioua, *Performance of Algebraic Graphs Based Stream-Ciphers Using Large Finite Fields*, Annalles UMCS Informatica AI X1, 2 (2011), 81-93.

[31] V. A. Ustimenko, M. Klisowski, *On the Comparison of Cryptographical Properties of Two Different Families of Graphs with Large Cycle Indicator*, Mathematics in Computer Science, 2012, V. 6, Number 2, pp. 181-198.

[32] V. A. Ustimenko, *Algebraic groups and small world graphs of high girth*, Albanian J. Math, vol3, N1 (2009), 25-33.

[33] V. A. Ustimenko, A. Wroblevska, *On the key exchange with nonlinear polynomial map of stable degree*, arXiv:1304, 2920, v.1.

Contact information

V. Ustimenko University of Maria Curie Skłodowska in Lublin
E-Mail(*s*): vasyl@hektor.umcs.lublin.pl

Received by the editors: 30.09.2015
and in final form 30.09.2015.

Contents

Наукове видання

Алгебра та дискретна математика
Том 20, Номер 1, 2015

Заснований у 2002 році.
Свідоцтво про державну
реєстрацію
КВ № 14443-3414ПР від 14.08.2008.
Виходить чотири рази на рік
англійською мовою.

Журнал внесений
до переліку наукових
фахових видань України
(фізико-математичні науки)
Постанова президії ВАК України
від 14 жовтня 2009 р. № 1-05/4.

Засновник і видавець:

"Луганський національний університет імені Тараса Шевченка"

Підписано до друку
рішенням Вченої ради механіко-математичного факультету
Київського національного університету імені Тараса Шевченка
(протокол № 2 від 14 вересня 2015 р.)

Технічний редактор: А. Б. Попов

Здано до складання 01.07.2015р. Підписано до друку 14.09.2015р.
Формат 60x84 1/16. Папір офсетний. Гарнітура Times New Roman.
Друк лазерний. Умов. друк. арк. 10,46.
Тираж 125 екз.

Видавництво Державного закладу
"Луганський національний університет імені Тараса Шевченка"
пл. Гоголя, 1, м. Старобільськ, 92703. Тел.: (06461) 2-26-70

Надруковано у типографії ТОВ "Цифра принт".
Свідоцтво про реєстрацію Серія А01 N 432705 від 03.08.2009р.
61058 м. Харків, вул. Данилевського, 30.