Algebraic graphs and security of digital communications

KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

UMCS

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

Projekt „Programowa i strukturalna reforma systemu kształcenia na Wydziale Mat-Fiz-Inf”.
Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego.

Człowiek-najlepsza inwestycja

Maria Curie-Skłodowska University
Faculty of Mathematics, Physics and Computer Science
Institute of Computer Science

Algebraic graphs and security of digital communications

Vasyl Ustimenko

Lublin 2011

Institute of Computer Science UMCS
Lublin 2011

Vasyl Usitmenko (Institute of Mathematics UMCS)
ALGEBRAIC GRAPHS AND SECURITY OF DIGITAL COMMUNICATIONS

Reviewer: Yuri Kondratiev

Technical Editor: Marcin Denkowski
Cover Designer: Agnieszka Kuśmierska

Praca współfinansowana ze środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego

A free online edition of this book is available at informatyka.umcs.lublin.pl.

Publisher
Maria Curie-Skłodowska University
Institute of Computer Science
pl. Marii Curie-Skłodowskiej 1, 20-031 Lublin
Series Editor: prof. dr hab. Paweł Mikołajczak
www: informatyka.umcs.lublin.pl
email: dyrii@hektor.umcs.lublin.pl

Printer
ESUS Agencja Reklamowo-Wydawnicza Tomasz Przybylak
ul. Ratajczaka 26/8
61-815 Poznań
www: www.esus.pl

ISBN: 978-83-62773-17-6

Contents

Preface

The term *graph* becomes the common word of Modern Mathematics and Theoretical Computer Science. Recall, that the abstract model of a computer, if we ignore the memory, is a finite automaton, roughly a directed graph with colours on arrows taken from some finite alphabet. To make a graph theoretical model of a computer with memory working with potentially infinite data, one can use alternatively a concept of Turing Machine or definition of an infinite family of directed graphs of increasing order. Studies of families of graphs (not an individual graph) satisfying a special requirements are highly motivated by applications in Economics, Natural Sciences, Computer Science, Networking and in Mathematics itself. For instance, the problem of constructing infinite families of small world graphs has many remarkable applications in all above mentioned areas and in sociology. For instance, the "small world graph" of binary relation "two persons shake hands" on the set of people in the world has small diameter.

Other impotant direction in studies of infinite families of simple graphs is Extremal Graph Theory. The girth of the graph is minimal length of its size. Some impotant results in this direction had been obtained in 50th by Paul Erdös' via studies of *families of graphs of large girth* i.e. infinite families of simple regular graphs of fixed degree and increasing order such that the girth of the member is growing logarithmically with the growth of the order. the existence such a family with arbitrary large degree had been proven by Erdös' famous probabilistic method.

Basically, just 3 explicit constructions of families of graphs of large girth with unbounded girth and arbitrarily large degree are known: the family of Cayley graphs had been constructed by Field's award holder G. Margulis approximately 40 years after appearance of Erdös' probabilistic construction, the family of algebraic graphs $D(n, q)$ defined over the arbitrary finite field F_q and regular polarity graphs for $D(n, q)$. Families of graphs of large girth are traditionally used in Networking.

The above two families of simple graphs of large girth can be easily converted in special finite automata and used for cryptographical purposes. Theoretical studies and software implementations of related numerical and

symbolic algorithm were condacted during last 10 years. Depending on time dynamical system defined on the vector space ${F_q}^n$ in terms of finite automaton related to mentioned above polarity graphs turns appropriate tool for the construction of stream ciphers and polynomial public key algorithms.

By definition finite automata are directed graphs (or digraphs), for which the concepts of families of small world graphs and graphs of large girth can be easily reformulated. The first results on Extremal Digraph Theory were obtained very recently. Instead of prohibition of cycles of small length there used requirements of absence of commutative diagrams. The analog of Erdös' upper bound for the graphs on v vertices of girth $> 2n$ and some other bounds had been obtained. New theory is principally different from the case of simple graph: the Erdös' bound is known to be sharp only in exaptional case of $n = 2, 3$ and 5 but its analog for the digraphs is always sharp.

The framework of Extremal Digraph Theory allows construction an infinite family of algebraic directed graphs of large girth for each finite commutative ring with more than 3 regular elements and define depending on time dynamical systems over free modulae. Change of finite fields on arithmetical rings Z_{2^8}, $Z_{2^{16}}$ and $Z_{2^{32}}$ usually used in computers for arithmetical computations allow to speed up the computations in encryption algorithms.

Infinite families of graphs are are traditionally used in classical Coding Theory. Foundations of this theory are based on the concept of finite distance -transitive or distance-regular metrics (distance regular and distance transitive graphs in other terminology). Recall that according to famous Hilbert's approach to Mathematical Concept of Geometry it is a special incidence system (or multipartite graph) Great deal of known families of distance transitive graphs are constructed in terms of the incidence geometry of group of Lie type or geometry of its Weyl group. Known construction of families of distance - regular but not distance transitive graphs are also based on the properties of such geometries. Many important classification results of the Theory of Finite Geometries were obtained quite recently. The leading researcher in that area J. Tits was awarded by prestigious Abel Prize in 2008. In fact, some new nonclassical areas of Coding Theory like LDPC codes and turbocodes use objects constructed via finite geometries: for the first constructions of LDPC codes Tanner used finite geometries of rank 2, the infinite algebraic family of graphs of large girth is related to infinite rank 2 geometry over finite field have been applied to constructions of new families of LDPC codes. Quite recent development gives an application of linear codes (elements of finite projective geometry) and their lattices to cryptography. Incidence geometries becoming tools for the development of cryptographical algorithms.

The course is devoted to applications of families of graphs and digraphs

of large girth, small world graphs and finite geometries to Information Security Theory. The main direction here is Cryptography, we only give references on some applications of graphs of large girth to Coding Theory.

The material of the course were used for Monographical and Special courses for seniour and graduate UMCS students majoring in Informatics and Applied Mathematics. Formally prerequisites of the course in full amount have to be Finite Fields Theory (some chapters of "Galois Theory" by E. Artin), Introduction to Combinatorial Group Theory (some chapters from "Combinatorial Group Theory" by Magnus, Karas and Soliter), "Introduction to Simple Groups of Lie type" (for instance, the famous paper by C. Shevalley), some elements of Ring Theory can be useful for sections on dynamical systems and directed graphs. Anyway there is an option to make a natural shortcut. For the simplicity we can assume that

(1) the finite field F_p always contains prime number p of its elements and we do mod$p$ computations.
(2) consider only description of Weyl groups of type A_n (symmetric group of order $n!$) and B_n (hyperoctahedrial group of order $n!2^n$) as groups given by Coxeter generators.
(3) finite simple group is always a group $PSL_n(p)$, which is a factor group of commutant for $GL_n(p)$ by center, where the group $GL_n(p)$ is a group of $n \times n$ matrices $A = (a_{i,j})$, $a_{i,j} \in F_p$ satisfying $\det(A) \neq 0$.
(4) commutative ring is always Z_n and we deal with calculus mod p.

After such assumptions the reader can understand the content of most sections of the manuscrypt.

There is an option to add some information on the "Galois" package for Java and generate some examples of finite fields of order p^s, $s \geq 2$ (additionally to very basic class given in (1)).

If the teacher has some knowledge on world famous "GAP" package some additional examples of groups given by generators and relations can be introduced (see item (2)).

Additionally to information on $PSL_n(p)$ and related projective geometry of (3)we can introduce symplectic, orthogonal or hermitian classical groups and corresponding geometries over finite fields.

Then the teacher can recall the structure of some commutative rings studied within the standard UMCS course on Modern Algebra.

As we all know that a real life course and the supporting hanbook on the course are always slightly different, so I hope, that this handbook will be helpfull for students during their work on Bachelor Papers, Masters' and Doctoral Thesis.

Let us give the short overview of chapters content.

The first Chapter is a very short introduction to Cryptography and Multivariable Polynomials Theory. It starts with elements of Classical Cryptography. We introduce the language of symmetric cryptography, properties of one time pad private key, linear and affine encryption together with its cryptanalysis, encryption based on Little Fermat Theorem. Other part of this chapter is an Introduction of main ideas of Complexity theory and Modern Cryptography. We define one way functions, present the idea of the public key algorithm, trapdoors, digital signatures. This section contains description of famous RSA algorithm based on Euler Theorem and Diffie - Hellman protocol for the key exchange. We give also Imai Matsumota algorithm for digital signatures based on quadratic multivariable invertible map on the vector space over Galois field. This is a very natural bridge to the last section devoted to special nonlinear polynomials over fields and rings. Readers can implement Imai-Matsumoto encryption with usage of Galois package for Java language. They can study effective cryptanalysis for Imai Matsumoto given by J. Patarin (Paris). The description of Patarins method the reader can find in well known book Algebraic Cryptography by Neal Koblitz. The last part of first chapter contains the following interesting facts on polynomials. Each permutation on finite vector space can be written in the form of polynomial map, there is a formulae of prime number written in the form of multivariable polynomial with integer coefficients (Matijasevich statement), there exist depending on time dynamical system for the finite dimensional vector space.

The second chapter devoted to algebraic aspects of Extremal Graph Theory. We present well known upper bounds for the size of graphs without prescribed cycles. Even Circuit Theorem by P. Erdös' and some of its modifications are presented. The concepts of a family of small world graphs and family of graphs of large girth are introduced. The explicit algebraic constructions of such graphs are given in this Chapter. They could be used in Coding Theory (references are given) and in Cryptography (algorithms are introduced). Readers can implement various modifications of such algorithms.

Third chapter contains other examples of applied algebraic graphs. The constructions of graphs are given in the language of Group Theory and Finite Geometries Theory. New cryptographical algorithms corresponding to infinite families of graphs are given. For instance, author introduces key exchange protocol given in terms of Tits and Schubert automata. J. Tits is one of the founders of Finite Geometries Theory. This part of the book requires serious prerequisites on Combinatorial Group Theory. The last chapter is devoted to directed graphs (shortly digraphs) and their applications to Information security. It is important because of finite automaton is a directed graph with special coloring. The chapter contains the overview

of Extremal Digraph Theory and various digraph based cryptographical algorithms. Some of such algorithms use arithmetical rings instead of one fiels.

Finally, I would like to say how it all at started.

My teacher of Mathematics, Lev Kaluznin, was a sun of Russian emigrants of October Revolution 1917. He got higher education at the best French and German Universities. He became a prominent mathematician. After Stalin's death he and his mother got a permission to return to the USSR. During his life he kept continious mail correspondence with his close friends A. Weil, M. Lazar, C. Shevalley, M. Krasner, H. Cassenhouse, with J. Tits who was his youngest classmate in the classes of E. Artin, as well as with A. Kerber and many of his former students at Berlin University (see [92]). As a Professor of Kiev State University he became one of the founders of Computer Science in the USSR (particularly, he provided theoretical support for the Lebedev's team during their work on the construction of the first

Soviet supercomputer, see [55]) and some interdisciplinary theoretical areas (Mathematical linguistics, Mathematical chemistry and some other areas).

Two sections of my PhD Thesis were devoted to studies of special maximal subgroups of symmetrical groups started by L. Kaluznin and M. Klin. Topic of the third section was proposed by J. Tits in his private letter to my supervisor. It was the problem of studies the overgroups of $PSL_n(q)$, acting on the totality of m-dimensional subspaces of finite projective geometry. All problems were formally from Permutation Group Theory, but they were connected with studies of Hamming, Johnson and Grassman graphs of Coding Theory their symmetries and new distance regular metrics.

For the postdoctoral research J.Tits proposed studies of overgroups of Finite Simple Groups of Lie type acting on the elements of their geometries of prescribed type. I started to work on that asighnment together with my first graduate students V, Zhdan-Poushkin and M. Muzychuk (currently Professor of Bar Ilan University (Israel)). One of the applied byproducts of my research in this direction was a discovery of distance regular but not distance transitive metric of Algebraic Coding Theory ("Ustimenko graphs" from subject index of "Distance regular graphs" by A. Brower, A. Cohen and A. Niemaer, Springer, 1989). The other technical result was the interpretation of geometries of simple group of Lie type and geometries of Kac-Moody group in terms of linear algebra. In case of Kac-Moody group over F_q over diagram $\tilde{A}_1$ the analitical descriptions of $q+1$ - regular and q-regular forests via infinite system of quadratic equations were obtained. It gave an option to present a q-regular forest as a projective limit of algebraic q-regular graphs.

In 1988 J. Hemmeter (USA) used my construction of distance regular graphs for creation "Hemmeter graphs". Naturally we started joint research. From the beginning F. Lazebnik (USA) and A. Woldar (USA) were participants of our project. In 1990 the National Science Foundation for the first time in its history organized an International Competition for the best USA-Soviet joint research project funded by NSF and Soviet Academy of Sciences. The document had been sighned by Presidents J.Bush and M.Gorbachev. We presented our project for the NSF competition. In 1991 the Soviet Union collapsed and NSF announced our project as the only winner (funding to University of Delaware came from the USA side only). At the beginning J. Hemmeter and I were Principal Investigators (PI). The project was prolongated several times (1991-1997) as standard NSF project (J. Hemmeter and A. Woldar moved from Delaware to other Institutions and F. Lazebnik became a PI. We used one of the deformations of Kac-Moody geometry for the definition of q-regular graphs $D(n,q)$, their projective limit was a q-regular forest $D(q)$.

Andrew Woldar was a very important contributor for the success of the project, for instance he formulated the conjecture on the descryption of trees of q-regular forest $D(q)$ during his General Membership at the Institute of Advanced Studies (Prinston)). We used one of the deformations of Kac-Moody geometry for the definition of q-regular graphs $D(n,q)$, their projective limit was a q-regular forest $D(q)$. We (F. Lazebnik, V. Ustimenko and A. Woldar) applied this family of graphs for the analitical descryption of q-regular tree $CD(q)$ and got several constructive results in Extremal Graph Theory based on properties of this family.

In 1997 the Guinnand and Lodge from Ottawa Communication Center found an interesting applications of $D(n,q)$. The adjacency matrix for $D(n,q)$ can be used succesfully as a Tanner graphs for the constructions of LDPC codes and turbocodes of Coding Theory for the protection of channels from noise. This idea were practically used by NASA and other companies in satelite communications. Studies of theoretical properties of related to $D(n,q)$ codes became a popular topic.

In 1996 the important for me event did happened. My teacher Lev Kaluznin told us, the participants of his seminar, that it would be the time when we could talk in person with European and American colleagues, whose results we were discussing at his seminars. At that time , we did not belive him. In 1996 I got an invitation from Professor Peter Slodowy to take part in the mini workshop on Buildings at the Oberwolvach Institute. It was an opportunity to meet J. Tits and F. Buekenhout in person. So my teacher was right. To my surprise J. Tits turns out to be a person very interested in real life applications and the reader of IEEE publications (software and hardware engeneering). His generalised polygons were used by

Tanner for the first constructions of LDPC codes. So J.Tits kept an eye on the development of expanding graphs and graphs of large girth applications.

Our NSF project was a success as a pure mathematical project, but J.Hemmeter (the key writer of the first proposal) and myself were planning applications to Computer Science. So I felt that the circle of the started research was not completed yet. In 1997 I got a General Membership at Mathematical Research Science Institute (Berkeley, USA) and used this opportunity to participate at H. Lenstra seminar on Cryptography and discussed the idea of usage families of graphs of large girth in Cryptography with E. Berlekamp. Next year I presented some encryption algorithms at AMS Meeting (Loisville, USA(March 1998)), Gary Ebert's Seminar at Delaware University(USA), International Memorial Voronoi Conference (Kiev,Ukraine), and seminars of University of Manchester, London University (joint seminar of Imperial College, Kings and Queen Mary college,UK) during my visit to the UK under the invitation of Royal Society. During my stay in Britain I got an offer from the University of the South Pacific where I could participate at the first implementations of graph based encryption algorithms (University intranet, Oracle based University banner system, GIS). I still work on these problems.

ACKNOLEGEMENTS.

I would like to thank all people I mentioned above, all my collaborators in direction of graph based security from different continents and islands, my Master and graduate students from University of South Pacific, Sultan Qaboos University (Oman), University of Maria Curie Sklodowska.

I'd like to express my gratitude to late prominent number theorist Peter Pleasants (UK, Australia, Fiji Island), who was one of the first readers of my papers on encryption algorithms for his very usefull remarks. I am very thankful to J. Seidel (Holland), John Hosack (USA, Fiji Islands), Georgy Giemelfarb (Center of Information Technology, Auckland University, New Zealand), Professors Josef Pieprzak and I. Sparlinsky (Sydney), Takashi Soma (Tokio), Thomas Bier (Sultan Qaboos University, Oman), Tony Shaska (USA and Albania), Cary Huffman (USA), Abdelhak Azhari (Morocco), Alex Borovik (Manchester), Alex Ivanov (Imperial College), I. Faradjev (California and Moskow), J. Kozicki (UMCS), Prof A. Kerber and A. Kohnert (Germany) for their friendly support and advices.

Special thanks to my UMCS collegues Piotr Pikuta and Aneta Wróblewska for their patiance and help during my work on the first handouts of that special course in Polish.

I am greatly indebted to Urszula Romańczuk for the constant technical support during the continuous work on the manuscript.

My profound thanks to my beloved wife who is always on my side.

Chapter 1

On Polynomial Maps, Dynamical Systems and Cryptography

1.1. Basics of Symmetric Cryptography

1.1.1. Introduction

Cryptography has a tremendous potential to enrich math education. In the first place, it puts mathematics in a dramatic setting. students are fascinated by intrigue and adventure. More is at stake than a grade on a test: if you make a mistake, your agent will be betrayed.

In the second place, cryptography provides a natural way to get students to discover certain key mathematical concepts and techniques on their own. Too often math teachers present everything on a silver platter, thereby depriving the students of the joy of discovery. In contrast, if after many hours the youngsters finally develop a method to break a cryptosystem, then they will be more likely to appreciate the power and beauty of the mathematics that they have uncovered.

In the third place, a central theme in cryptography is what we do not know or cannot do. The security of a cryptosystem often rests on our inability to efficiently solve a problem in algebra, number theory, or combinatorics. Thus, cryptography provides a way to counterbalance the impression that students often have that with the right formula and a good computer any math problem can be quickly solved.

Finally, cryptography provides an excellent opportunity for interdisciplinary projects completed by team of students.

1.1.2. On terminology of classical cryptography, linear algebra methods

Assume that an unencrypted message, *plaintext*, which can be image data, is a string of bytes. It is to be transformed into an encrypted string or *ciphertext*, by means of a cryptographic algorithm and a *key*: so that the recipient can read the message, encryption must be *invertible*.

An assumption first codified by Kerckhoffs in the nineteen century is that the algorithm is known and the security of algorithm rests entirely on the security of the key.

Conventional wisdom holds that in order to defy easy decryption, a cryptographic algorithm should produce seeming chaos: that is, ciphertext should look and test random. In theory an eavesdropper should not be able to determine any significant information from an intercepted ciphertext. Broadly speaking, attacks to a cryptosystem fall into 2 categories: *passive attacks*, in which adversary monitors the communication channel and *active attacks*, in which the adversary may transmit messages to obtain information (e.g. ciphertext of chosen plaintext).

Attackers hope to determine the plaintext from the ciphertext they capture; even more successful attacks will determine the key and thus comprise the whole set of messages.

Passive attacks are subdivided into two following major types:

(i) *ciphertext only*-the adversary has access to the encrypted communications.
(ii) *known plaintext* – the adversary has some plaintextx and corresponding ciphertexts.

In case of attacks of type (i) the adversary hopes to determine the plaintext from the captured ciphertext. The goal of attacks of type (ii) is getting the key.

The revolutionary classical result on private key algorithm was obtained by C. Shannon at the end of 40th (see [52], [53] and further references). He constructed so called *absolutely secure* algorithms, for which keys are rings of random bits at least as long as a message itself. They achieve the seeming impossibility: an eavesdropper is not able to determine any significant information from obtained ciphertext. So his or her only option is a brut force serch via entire keyspace. The simplest classical example is the following *one-time pad*: if p_i is the i-th bit of the plaintext, k_i is the i-th bit of the key, and c_i is the first bit of the ciphertext, then $c_i = p_i + k_i$, where $+$ is exclusive or, often written XOR, and is simply addition modulo 2. One time pads must be used exactly once: if a key is ever reused, the system becomes highly vulnerable.

It is clear that the encryption scheme as above is irresistible to attacks of type (ii) - you need just subtract p_i from c_i and get the key.

Let us consider more general case of affine transformation $x \to Ax+b$ of the vector space ${F_q}^n$, where q is a prime power and F_q is a finite field of order q. We assume here that each element of the F_q is a column vector and write the encryption transformation as $x \to Ax + b$, where x is the plaintext, the nonsingular matrix A and $b \in F_q$ form the key. Let the plaintexts $p_1, \ldots, p_{n+1}$ are in "general position" i.e. $p_2 - p_1, p_3 - p_1, \ldots, p_{n+1} - p_1$ are linearly independent vectors and the adversary (or cryptanalyst) get related to each p_i ciphertext c_i, then this information is sufficient for the computation of the key (A, b) by methods of elementary linear algebra.

Nowadays the security of the encryption private key algorithms rests on the chosen password (key), it has to be resistant to attacks of type (i) and (ii). So in case of the plainspace ${F_q}^n$ we need polynomial bijections of degree ≥ 2 as encryption maps.

We assume conventional definition of algorithm define via exexution of Turing machine working with "potentially infinite text". It means that encryption algorithm working with "potentially infinite" plainspace.

All algorithms for the symmetric encryption are divided on *block ciphers* and *stream ciphers*. In case of block ciphers the plainspace P is partited onto blocks B_i, $i = 1, 2, \ldots, n$ of equal size equal to some constant b The encryption map corresponding to chosen key maps each set B_i onto itself.

Stream cipher is the fast encryption algorithm which is not a block cipher. It means that the partition onto invariant blocks does not exist. Let us use the language of permutation group theory for studies of principle difference between block ciphers and stream ciphers. The encryption map is a bijection (permutation) on plainspace. We may consider the permutation group G_A generated by encryption maps for chosen algorithm A. For the construction of G_A we may use various combinations of keys from the keyspace. Two points p and p' belongs to the same orbit if there is a permutation $\pi \in G_A$ such that $\pi(p) = p'$. In case of block cipher each block is a union of some orbits. So the size of orbit does not grow with the growth of size of the plainspace, it is bounded by b. In case of reasonable stream cipher size of each orbit is growing. By definition a transitive permutation group is a subgroup of corresponding symmetric group with exactly one orbit. The algorithm A with transitive group G_A has the following property: for arbitrary pair p and p' there is π in G_A corresponding to some combination of keys such that $\pi(p) = p'$.

1.1.3. Little Fermat Theorem and discrete logarithm

Let us consider example of algorithm with good resistence to attacks of type (ii).

Acording to famous Little Fermat's Theorem for each prime number p and integer $x \neq 0 \bmod p$ the equality $x^{p-1} = 1 \bmod p$ holds. The proof is very easy: Let $1, 2, \ldots, p-1$ be the list of elements of multiplicative group ${F_p}^*$ for F_p. We may choose $x \in {F_p}^*$ and form the new list $1x, 2x, \ldots, x(p-1)$.

Element x is invertible so both lists containe same elements just written in different order. So the computation of the products $1 \times 2, \ldots, \times p-1$ and $1x \times 2x, \ldots, \times x(p-1)$ gives us same number modulo p. We get the equality

$$(p-1)! = (p-1)! x^{p-1} \bmod p.$$

We can multiply lefthandside and righthanside by the inverse for $(p-1)!$ and get the equality written in Little Fermat Theorem

This statement implies $x^p = x$ and $x^\alpha = x$ for $\alpha = 1 \bmod p-1$.

The algorithms based on the Little Fermat Theorem

We assume that totality of residues $\{x|x \neq 0, x \neq 1 \mod p$, is our plainspace. Integers α such that α is mutually prime with $p-1$ form the key-space. So the map

$$f_\alpha : x \to y = x^\alpha$$

is encryption map.

Little Fermats Theorem allows to decrypt with the map

$$g_\beta : y \to z = y^\beta,$$

where $\alpha\beta = 1 \mod p-1$.

Alice and Bob can easily compute β via expanded Euclead's algorithm checking that the graatest common divisor for (α and $p-1$ is 1 and presenting 1 in the form $\alpha M + (p-1)N$. It is clear that $M(\mathrm{mod}p-1) = \beta$

Let adversary get a pair d (plaintext) and c (ciphertext). For the computation of the key he (or she) has to solve the equation $d^\alpha = c \mod p-1$ for variable α.

Finding α is famous difficult *discrete logarithm problem*. If p is "sufficiently large" (for instance p contains 200 digital numbers nobody knows how to solve it. The equation is on the list of known NP-hard problem (see [42]). Conjecture that there is no polynomial in time algorithm for solving problems from the list is still open, but nobody knows how to create such algorithm. It means that the encryption method based on Little Fermat's Theorem has good resistence to attacks of type (ii).

1.2. On the concepts of Modern Cryptography

1.2.1. Ideas of assymetry

The paper [25] by Diffie and Hellman was published in 1976. This event change the shape of Cryptography, some new directions were developed.

The basic definitions of Modern Cryptography are below.

One way function is the one to one correspondence satisfying following requirements

(i) there exists a polynomial algorithm for the computation of the value $F(x)$.

(ii) the polynomial algorithm of finding inverse map F^{-1} does not exist.

The conjecture on existence of one way function is open. For practical use one may substitute requirement (ii) on weakercondition:

$(ii)'$ the complexity of polynomial algorithm of finding inverse map F^{-1} is equivalent to solving of one NP-hard problem from the list [42].

Trapdoor function with a secret parameter K is a one to one correspondence $F_K : X \to Y$ satisfying following 3 requirements

(i) there exists a polynomial algorithm for the computation of the value $F_K(x)$ for each K and x.

(ii) the polynomial algorithm of finding inverse map F_K^{-1} for unknown K does not exist.

(iii) there exists a polynomial algorithm for the computation of the inverse for $F_K(x)$ with known parameter K.

Again the statement on the existence of trapdoor function is not proven yet.

There are examples of functions satisfying (i) and (iii) and requirement (ii)'. The most famous is the encryption function for RSA cipher.

The definitions above are motivated by idea of public key or assymetryc cryptographical algorithm. Let us consider the way to use trap-door functions for solution of new cryptographical assighnments.

Alice (the holder of secret papameter K) wants safe delivery of secret messages via open communication channel. Bob (public user) does not have a parameter K. He get an encryption function $F_K(x)$ via open channel without option to compute K. If Alice (or somebody else) sends him encrypted plaintext $F_K(p)$ he can not decrypt and get p. Of course the holder of K may enjoy the property (iii) and decrypt Bob's messages within polynomial time. The adversary is in the same shoes with Bob, so he has no option can not decrypt Bob's messages.

Notice, that the adversary can make attacks of type (iii) because he can compute the corresponding ciphertext for any chosen plaintext. Encryption based on the trap-door function (of course, in the case of its existence) has a wounderfull resistence to attacks of type (iii).

The term public key is used, because Alice presents encryption function to public (printing in telephone book, sending by internet, etc)

In the same paper Diffie and Hellman proposed the key exchange algorithm. They used the encryption function based on Little Fermat's Theorem introduced in previous unit. Correspondents Alice and Bob establish a primitive element b of multiplicative group of prime finite field F_p via open communication channel. They choose positive integers n_A and n_B, respectively.

They exchange elements $h_A = b^{n_A}$ and $h_B = b^{n_B}$ via open channel. Finally, Alice and Bob compute common vector c as ${h_B}^{n_A}$ and ${h_A}^{n_B}$, respectively. So they can use c as a key in some symmetric encryption method

The security of "symbolic Diffie-Hellman algorithm" is based on the discrete log problem for the cyclic multiplicative group for F_p:

Really, the adversary (Catherina) has field elements b, $c_1 = b^{n_A}$, and $c_2 = b^{n_B}$.

She have to solve one of the equations $b^x = c_i$, $i = 1, 2$. Let the adversary gets n_A as a solution of first equation. Then she computes c as $c_2{}_A^n$.

The dicrete logarithm problem is on the list of NP-hard problems. So if p is "sufficiently large" then the protocol for key exchange is secure.

1.2.2. Euler Theorem and RSA

Let us consider the commutative ring $Z_n = \{0, 1, 2, \ldots, p-1\}$. Element i of Z_n is regular if the greatest common divisor of i and n is 1. Leonard Euler generalised Little Fermat's Theorem via studies of the multiplicative group of all regular (invertible) elements of the ring Z_n. Recall that classes of Z_n are $0, 1, \ldots, n-1$. If the greatest common divisor of i and n is 1. Then extended Eucleads algorithm allows to write presentation of 1 in the form of linear combination of i and n:

$$1 = Mn + Ni$$

If we consider the lefthandside and righthandside of the above equality modulo n we obtaine $1 = Ni \bmod n$.

Thus $N \bmod n$ is the multiplicative inverse for i.

If the greatest common divisor of i and n is $d > 1$ then $i \times n/d$ equals 0. So i as a zero divisor does not have an inverse element.

So the multiplicative group ${Z_n}^*$ of Z_n is $\{i|(i, n) = 1\}$. The order of this group is known as Euler function $\phi(n) = |\{i|(i, n) = 1\}|$. Euler proved that remarkable multiplicative property of ϕ : for each pair m, n function $\phi(mn)$ coinsides with $\phi(n) \times \phi(m)$.

From the Little Fermats Theorem we get $\phi(p) = p - 1$ for the prime p. It is clear that $\phi(2^m) = 2^{m-1}$, $m \geq 1$ because the totality of mutual primes with 2^m on the interval $(0, 2^m)$ is exactly the collection of odd numbers from the interval. We can combine this two observations and get

$$\phi(p^m) = p^m - p^{m-1}$$

for prime p and integer $m \geq 1$.

Finally, the Main Theorem of Arithmetics allow us to present each positive integer n as a product of its divisors

$$n = {p_1}^{\alpha_1} \times {p_2}^{\alpha_2} \ldots {p_s}^{\alpha_s}$$

where $p_1, p_2, \ldots p_s$ is a list of all prime divisors of n.

The above decomposition allows to write Euler function in the form

$$\phi(n) = ({p_1}^{\alpha_1}) - {p_1}^{\alpha_1 - 1}) \times ({p_2}^{\alpha_2} - {p_2}^{\alpha_2 - 1}) \times \cdots \times ({p_s}^{\alpha_s} - {p_s}^{\alpha_s - 1}).$$

Anyway the computation of the Euler function can be a very hard task if the decomposition of n into primes is unknown.

The Euler Theorem can be obtained as a corollary of the well known statement of Finite Group Theory: The order of group element is a divisor of the group order. So in case of group ${Z_n}^*$ we have

$$g^{\phi(n)} = 1 \bmod n \quad \text{for } g \in {Z_n}^*.$$

In case $n = p$, where p is prime we are getting Little Fermat Theorem.

We can write Euler Theorem in the form

$$g^{\alpha} = g \bmod n \quad \text{for } \alpha = 1 \bmod \phi(n).$$

Let us consider the posibility of *symmetric encryption based on Euler Theorem*.

We assume that totality of residues $\{x|(x,n) = 1\}$ is our plainspace. Integers α such that α is mutually prime with $\phi(n)$ form the key-space. So the map

$$f_\alpha : x \to y = x^\alpha$$

is the encryption map.

Euler Theorem allows to decrypt with the map

$$g_\beta : y \to z = y^\beta,$$

where $\alpha\beta = 1 \bmod \phi(n)$.

Alice and Bob share α, they can easily compute β via expanded Euclead's algorithm checking that the graates common divisor for (α and $\phi(n)$ is 1 and presenting 1 in the form $\alpha M + \phi(n)N$. It is clear that $M(\mathrm{mod}\phi(n)) = \beta$

Let adversary get a pair d (plaintext) and c (ciphertext). For the computation of the key he (or she) has to solve the equation $d^\alpha = c \bmod \phi(n)$ for variable α.

Finding α is famous difficult *discrete logarithm problem*. The complexity depends heavily from parameter n. If n is the products of big primes nobody knows how to solve this problem. So we are getting a symmetric algorithm with the good resistence to attacks of type (ii).

Assymetric algorithm RSA based on Euler Theorem.

The main idea of RSA based on the following facts

(a) the computation of the products of two numbers can be completed fast by modern computer
(b) nobody knows fast algorithm forthe prime factorization of big integer m.

Of course, one can list all primes $\leq \sqrt(m)$ and divide m on each of them. It allows us to factorize m. The problem is that assimptotically the number of primes from such a list is $2\sqrt(m)\ln(m)^{-1}$ (see [81]). If m consist on 100 decimal digits, then there are at least 4×10^{42} primes within the interval. It means that for the computer making million of operations per second prime decomposition will take at least 10^{35} years.

Nowadeys are known more efficient algorithms, but all of them are also rather slow.

Authors of RSA proposed to chose number n in the form of product of two distinct primes p and q of approximately same order. So $\phi(n) = (p-1)(q-1)$. The unique condition for the choice of α in the above algorithm is $(\alpha, p-1) = (\alpha, q-1) = 1$.

So Alicia (the holder of the key) choses p and q. She computes $n = pq$ and chooses α. She knows $\phi(n) = (p-1)(q-1)$ and computes β via extended Eucleads algorithm. So she can encrypt with the function

$$f_\alpha : x \to y = x^\alpha$$

and decrypt with

$$g_\beta : y \to z = y^\beta,$$

where $\alpha\beta = 1 \bmod \phi(n)$.

So Alice can print the pair (n, α) in the telephone book. So any public user (Bob) can use encryption function f_α. Of course Alice keeps primes p an q, secretely. If primes are "sufficiently large" Bob is able to encrypt but not able to decrypt.

For the illustration of their method Rivest, Samir and Adlemann encrypted some frase in English. During the first step they used standard method of converting the text to number $a = 01, b = 02, \ldots, z = 26$, emptyspace = 00, on the second step they used encryption map f_α for $n =$ 11438116257578888 67669325779976146201021829672124236256256184293570693524573389783059712356395870505898907514759929002687954354 1 and $\alpha = 9007$. Numbers n, α had been published. The information that $m = pq$ where p and q are primes written with 63 and 64 decimal digits.

The authors promised the 100 dollars as award for solution. This story comes to an end 17 years later, when D. Atkins, M. Graff, A. K. Lenstra and P.C. Layland announced on decryption of the frase (see [2]). The result was achieved due to application of new quadratic sieve method for the prime decomposition and usage of enourmous computational power of 1600 computers, the work of approximately 600 volonteers during 220 days. The 100 dollars aword had been sent to Free Software foundation.

1.2.3. Cryptoanalitical example, Imai - Matsumoto encryption

Let K be an extension of degree n of the finite field F_q, where q is a power of 2, and let $\beta_1, \beta_2, \ldots, \beta_n$ be a basis of K as an F_q-vector space.

Alice will be using the Imai- Matsymoto system in K. She regards each element of K as an n-tuple over q.

Alice may choose to keep her basis secret in which case we can not assume that a cryptoanalyst (whom we shall name "Catherine") knows what basis she is using.

Both plaintext message units and ciphertext message units will be n-tples over F_q. We will use the vector notation

$$\mathrm{x} = (x_1, x_2, \ldots, x_n) \in {F_q}^n$$

for plaintext and

$$\mathrm{y} = (y_1, y_2, \ldots, y_n) \in {F_q}^n$$

for ciphertext. When working with matrices, we shall consider vectors to be column vectors (although in the text we shall continue writing them as rows).

In transforming paintext into ciphertext, Alice will work two intermidiate vectors, denoted $\mathrm{u} = (u_1, \ldots, u_n) \in {F_q}^n$ and $\mathrm{v} = (v_1, \ldots, v_n) \in {F_q}^n$. Given a vector in ${F_q}^n$, we shall use boldface to denote the corresponding element of K with respect to the basis β_j.

Next, Alice chooses an exponent h, $0 < h < q^n$, that is of the form

$$h = q^\alpha + 1$$

and satisfies the condition g.c.d.$(h, q^n - 1) = 1$. (Recall that q was choosen to be a power of 2, if q were odd, then g.c.d.$(h, q^n - 1)$ is at least 2.)

The condition g.c.d.$(h, q^n - 1) = 1$ is equivalent to requiring that the map $\mathrm{u} \to \mathrm{u}^h$ on K is one to one, its inverse is the map $\mathrm{u} \to \mathrm{u}^{h'}$, where h' is the multiplicative inverse of h modulo $q^n - 1$.

Alice may choose o keep h secret. However, since there are relatively few possible values for h, she must asume that Catherine will be prepared

to run through all possibilities for h. That is, even if she keeps h secret, the security of her system must be elsewhere.

In addition, Alice chooses two secret affine transformations, i. e., two invertible $n \times n$ matrices $A = (a_{ij})$, $1 \leq i, j \leq n$ and $B = (b_{ij})$, $1 \leq i, j \leq n$ with entries in F_q, and two constant vectors $\mathrm{c} = (c_1, \ldots, c_n)$ and $\mathrm{d} = (d_1, \ldots, d_n)$.

The purpose of the two transformations is to "hide the monomial map" $u \to u^h$ - hence the name "hidden monomial cryptosystem".

We now describe how Alice gets her public rule for going from plaintext $\mathrm{x} \in {F_q}^n$ to ciphertext $\mathrm{y} \in {F_q}^n$.

First, she sets

$$\mathrm{u} = A\mathrm{x} + \mathrm{c}.$$

Next, she would like to have $v \in K$ simply equal to the h-th power of $\mathrm{u} \in K$ and then set

$$\mathrm{y} = B^{-1}(\mathrm{v} - \mathrm{d})$$

that is $\mathrm{v} = B\mathrm{y} + \mathrm{d}$.

However, her public encryption rule will go right from x to y, and will not directly involve exponentiation at all.

In order to get formulas going from x directly to y Alice notices that since $v = u^h$ and $h = q^\theta + 1$, she has

$$v = u^{q^\theta} u.$$

Recall that for any $k = 1, 2, ..., n$ the operation of raising to the q^k-th power in K is an F_q-linear transformation. Using linear algebra, she can get n-equations that express each y as a polynomial of total degree 2 in the $x_1, \ldots, x_n$.

Alice makes these n equations public. If Bob wants to send her a plaintext message x, he substitutes the x_i in these equations and finds y_i. On the other hand, Catherine, who knows only the ciphertext (and the public key), must solve a nonlinear system for the unknowns x_i.

When Alice receives the ciphertext y, she uses her knowledge of A, B, c and d and h to recover x, without having to solve the publicly known equations for the x_i. Namely, let h' be the multiplicative inverse of h modulo $q^n - 1$, so that the map $u = v^{\mathrm{h}'}$ inverts the map $v = u^{\mathrm{h}}$ on K. Alice first computes $\mathrm{v} = B\mathrm{y} + d$, then raises $\mathrm{v} = \sum v_i \beta_i \in K$ to the h'-th power (i.e., sets $u = v^{\mathrm{h}'}$, and finaly compute $\mathrm{x} = A^{-1}(\mathrm{u} - \mathrm{c})$.

The following summarises Alice's decryption:

$$(y_1, \ldots, y_n) \to \mathrm{y} = B\mathrm{y} + d \to \mathrm{u} = \mathrm{v}^{\mathrm{h}'} \to \mathrm{x} = A^{-1}(\mathrm{u} - \mathrm{c})$$

Remark. The cryptosystem described above is a simplified version of the one proposed in the original paper of authors. For details about breaking the original Imai-Matsumoto system see [59], [80] and further references.

1.3. Remarks on the power of bijective polynomial maps

1.3.1. On the group of bijective polynomials

Let $(F_q)^n$ be a vector space over the finite field F_q, where q is the prime power.

As it is usual in cryptography, we can apply a term *plaintext* to a string of characters $\mathrm{x} = (x_1, x_2, \ldots, x_n)$ over the alphabet F_q. When working with matrices, we shall consider vectors to be column vectors (although in the text we shall continue use them as rows). We can consider x as a message containing a certain information. If π is some bijective transformation of $(F_q)^n$, then $\pi(\mathrm{x})$ is an encrypted message or a *ciphertext*.

The natural choice for π is a combination of some affine transformations $a_i = A_i\mathrm{x} + b_i$, $i = 1, \ldots, k$, where A_i is a square matrix and $b_i \in (F_q)^n$, with some nonlinear transformation T of the vector space $(F_q)^n$.

Let us consider the case of F_p, where p is a prime number. Affine transformations $\mathrm{x} \to A\mathrm{x} + b$, where A be an invertible matrix and $b \in (F_q)^n$ form an affine group $AGL_n(F_p)$ of order $p^n(p^n-1)(p^n-p)\ldots(p^n-p^{n-1})$. This is a subgroup of the symmetric group S_{p^n} of order $(p^n)!$.

The following fact had been proven in [76].

Theorem 1. *Let G a proper subgroup of S_{p^n} containing $AGL_n(F_p)$. Then G coincides with $AGL_n(p)$ or S_{p^n}.*

Let us choose the nonlinear transformation T. The following statement follows directly from the theorem

Corollary 1. *Let T be a chosen nonaffine transformation of the vector space $V = {F_p}^n$. Then each bijective transformation T of the vector space $V = (F_p)^n$ can be presented as a product of " basic " transformations $Q(\alpha_1, \alpha_2) = \alpha_1 T \alpha_2$ where α_1 and α_2 are appropriate affine transformations of V.*

We recall the following well-known algebraic fact:

Theorem 2. *Each transformation T of the vector space $V = (F_p)^n$ can be treated as a polynomial map $P : \mathrm{x} \to \mathrm{y}$, where $\mathrm{x} = (x_1, \ldots, x_n)$, $\mathrm{y} = (y_1, \ldots, y_n)$, $y_i = P_i(x_1, \ldots, x_n)$, $i = 1, \ldots, n$ for some polynomial expressions P_i in variables x_i over the finite field F_p.*

It means that symmetric group S_{p^n} is the Cremona group for the vector space ${F_p}^n$ of all regular polynomial avtomorphisms. Notice that change open variety F_2^n onto ${F_2}^n - \{0\}$.

The following "public key" strategy can be derived naturally from the statements above:

(A) Choose polynomial transformation P, which you can invert fast (for polynomial number of steps $f(n)$), where $\deg f(n)$ is "small")
(B) select the family Ω of affine transformations $\alpha_i, i = 1, \ldots, m$ and quantum maps $Q_j = \beta_j P \gamma_j$, $j = 1, \ldots, l$, where $\beta_j, \gamma_j \in \Omega$
(C) compute the polynomial map $Q = Q_1 Q_2 \ldots Q_l$ (composition of Q_i), i.e. get the formula $y = Q(x) = (P_1(x_1, \ldots x_n), \ldots, P_n(x_1, \ldots, x_n))$, where polynomials P_i written in canonical form.
(D) of Q into quantum maps Q_i secret and give the list L of public equations P_i to you correspondent B.

Now, you correspondent can encrypt his/her messages to you by applying Q to the plaintext, but the problem of decryption , i.e. computation of inverse map Q^{-1} can reach any level of complexity: you may obtain any permutation from the symmetric group of S_{p^n} as your expression Q.

For you, the problem, of decryption can be feasible if length l is "reasonably moderate". You can invert each Q_i and apply them to the ciphertext in the reverse order with respect to the known decomposition of Q in Q_i.

Of course, we have no illusion to solve mathematically the great problem of cryptography on the existence of asymmetric function: corollary from the theorem 1 does not contain any restrictions on the number l of basic transformations.

But, it is reasonable to assume that even in case of polynomial length l we are able to produce practically secure public keys. In fact, well known Imai-Matsumoto encryption scheme, Small Dragon and its modifications by J. Patarin are realisation of A-D in case $l = 1$. They are examples of basic "transformations". The reader can find correspondent cryptanalysis in [59]. Multiple rounds of these algoritms (the case of $l \geq 2$) could be secure. It could be that basic transformation is more sophisticated than the composition of many basic transformations.

Remark 1. Substitution the field $GF(q)$, where $q = p^j$, p is a prime number, instead of F_p in the scheme (A)-(D) does not led to more general scheme, because of vector space $(F_q)^d$ over the ground field F_q, is a vector space of dimension jd over F_p, but such a substitution can be useful in practical applications. We can consider also a K^j, where K is a commutative field, instead of vector space ${F_p}^j$.

Remark 2.Size of the family Ω of step B can be bounded by polynomial expression in variable n, we may think that Ω consist of some elementary transvections $t_{i,j}(1)$, $i \neq j$ and diagonal matrices for which exactly one entry equals to fixed generator of of multiplicative group of F_p and other entries equal 1, regular translations $\mathrm{x} \to \mathrm{x} + \mathrm{e}_i$, where e_i is addition of 1 to x_i. One can consider even smaller set of generators of Affine group.

1.3.2. Towards polynomial formulae for the prime number

In time of Leonard Euler, who noticed that $x^2 + x + 1$ takes on prime values for $x = 0, 1, 2, \ldots, 39$, mathematicians were working on the search for the plynomial formulae for the prime number p, i.e element $q(x) \in Z[x]$ such that for each natural number n the number $q(n)$ is prime (positive or negative) and for each prime number p there is natural x such that $q(x) = n$. Nowadays we know that there is no a polynomial $q(x)$ with this property. Sadly, it is easy to show that a polynomial $(P(z_1, z_2, \ldots, z_k)$ with complex coefficients, which takes only prime values of nonnegative integers, must be constant.

Anyway this directions was not useless. Modified approach might be to ask if there is a non constant polynomial all of whose *positive values* (as the variables range in the set of non-negative integers) are all primes.

Yuri Matijasevic got a proof that this was possible for polynomial with 32 variables as a byproduct of his 10-th Hilbert problem solutions a byproduct of his 10-th Hilbert problem solution (see [72]).

Jones. Sato, Wada and Wiens [49] gave the following explicit example of such a polynomial with 26 variables (and degree 25):

$$
\begin{aligned}
&(k+2)\{1-[wz+h+jq]^2-[(gk+2g+k+1)(h+j)+hz]^2-\\
&-[2n+p+q+ze]^2-[16(k+1)^3(k+2)(n+1)^2+1-f^2]^2-\\
&-[e^3(e+2)(a+1)^2+1-o^2]^2-[(a^2-1)y^2+1-x^2]^2-\\
&-[16r^2y^4(a^2-1)+1-u^2]^2-\\
&-[((a+u^2(u^2-a))^2-1)(n+4dy)^2+1-(x+cu)^2]^2\\
&-[n+l+v-y]^2-[(a^2-1)l^2+1-m^2]^2-[ai+k+1-li]^2-\\
&-[p+l(a-n-1)+b(2an+2a-n^2-2n-2)-m]^2-\\
&-[q+y(a-p-1)++s(2ap+2a-p^2-2p-2)-x]^2-\\
&-[z+pl(a-p)+t(2ap-p^2-1)-pm]^2\}
\end{aligned}
$$

Notice that this polynomial factors! Look at the special form of the second part: it is one minus a sum of squares, so the only way for it to be

positive is for each of the squared terms to be zero.

Challenge:
Can you find a values $(a, b, c, d, \dots, z)$ (all non-negative) for which the polynomial above is positive?

The record for the lowest degree of such a polynomial is 5 (with 42 variables), and the record for fewest variables is 10 (with degree about $1.6.10^{45}$ [73]).

From other side there is no problem to increase the number of variables: we say that $Q(x_1, \dots, x_n)$ is a prime generating polynomial PGP if the range of Q for nonnegative values of variables is the totality of all primes. Of course PGP's exist for each $n \geq 10$, they have an interesting cryptographical properties, it is an interesting problem to study the complexity of their computation.

1.4. Arithmetical dynamical systems on a free module and hidden discrete logarithm

1.4.1. 0n the existence of arithmetical dynamical systems

It is well known that a continuous bijection of the interval $[a, b]$ has a fixed point. In case of open variety K^n, where K is commutative ring situation is different. We can define various nonlinear polynomial bijections on K^n which do not have a fixed point. Families of special nonlinear maps of this kind with additional cryptographical properties can be defined via *arithmetical dynamical systems* [118].

This section is devoted to the special key management block for the polynomial stream ciphers defined in [40] via such system defined on the free module K^n for each commutative ring. Security of the key based on the complexity of the discrete logarithm problem. It has additional heuristic security because of the "hidden base" and "hidden value" of the discrete logarithm function. Implemented software package has been used for the evaluation of mixing properties and speed of the private key encryption [46]. Let K be the commutative ring, $F(K) = K[t, x_1, x_2, \dots]$ is the ring of all polynomials in variables $t, x_1, x_2, \dots$. We use symbol $\mathrm{Reg}(K)$ for the totality of regular elements i.e not a zero divisors: $a \in \mathrm{Reg}(K)$ implies $a \times x \neq 0$ for each $x \neq 0$. Let $K^\infty = \{\mathrm{x} = (t, x_1, x_2, \dots) | x_i \in K, t \in K, \mathrm{supp(x)}, \infty\}$ and $K^n = \{(x_1, x_2, \dots, x_n) | x_i \in K\}$.

Let us consider two polynomial maps P and R of K^∞ into K^∞:

$$(t, x_1, x_2, \dots,) \to (t, P_1(t, x_1, x_2, \dots), P_2(t, x_1, x_2, \dots), \dots)$$

$$(t, x_1, x_2, \dots,) \to (t, R_1(t, x_1, x_2, \dots), R_2(t, x_1, x_2, \dots), \dots),$$

where $P_i(t, x_1, x_2, \dots)$ and $R_i(t, x_1, x_2, \dots)$, $i = 1, 2, \dots$ are elements of $F(K)$.

We consider two families:
f_t^n and ${g_t}^n$ of K^n onto K^n sending $(x_1, x_2, \dots, x_n)$ to

$$(P_1'(t, x_1, x_2, \dots),\ P_2'(t, x_1, x_2, \dots), \dots,\ P_n'(t, x_1, x_2, \dots x_n))$$

and

$$(R_1'(t, x_1, x_2, \dots),\ R_2'(t, x_1, x_2, \dots), \dots,\ R_n'(t, x_1, x_2, \dots x_n)),$$

where P_i' and R_i', $i = 1, 2, \dots, n$ correspond to the specialisations $x_{n+1} = 0, x_{n+2} = 0, \dots$ of P_i and R_i associated with the pair $(P,\ R)$. We identify f_t and g_t, $t \in K$ with the corresponding maps $K^n \to K^n$

Let $\mathrm{r} = (r_1, r_2, \dots r_t) \in \mathrm{Reg}(K)^t$ be the tuple of length $l(\mathrm{r})) = t$. We introduce F_r, as the composition of maps $f_{r_1}, g_{r_2}, \dots, f_{r_{2s-1}}, g_{r_{2s}}$ in case of $t = 2s$ and as composition of $f_{r_1}, g_{r_2}, \dots, f_{r_{2s-1}}, g_{r_{2s}}, f_{r_{2s+1}}$ for $t = 2s + 1$.

We say that the pair P and R form an arithmetical dynamical system depending on time t if the following conditions hold

(1) existence of $\mathrm{x} = (x_1, \dots, x_n) \in K^n$ such that

$$f_{t_1}(x_1, x_2, \dots, x_n) = f_{t_2}(x_1, x_2, \dots, x_n)$$

for some t_1 and t_2 implies the equality $t_1 = t_2$.

(2) maps f_t and g_t, $t \in K$ are bijections and f_{-t} and g_{-t} are inverse maps for them.

(3) There is a constant c , $c > o$ such that for each pair of tuples r, b of regular elements, conditions $l(\mathrm{r}) \le cn$, $l(\mathrm{r}) \le cn$ and $F_\mathrm{r}(\mathrm{x}) = F_\mathrm{b}(\mathrm{x})$ for some x implies $\mathrm{r} = \mathrm{b}$.

If (P, R) form an arithmetical dynamical system, then the inverse of F_r, $l(\mathrm{r}) = 2s+1$ is F_b, where $\mathrm{b} = (-r_{2s+1}, -r_{2s}, \dots, -r_1)$. If $l(\mathrm{r}) = 2s$ then F_r^{-1} is the composition of $g_{-r_{2s}}$ and F_d, where $\mathrm{d} = (-r_{2s-1}, -r_{2s-2}, \dots, -r_1)$.

We can treat K^n as the plainspace, refer to the union U of $\mathrm{Reg}(K)^t$, $1 < t < cn$ as the key space and treat

$$\mathrm{x} \to F_\mathrm{a}(\mathrm{x})$$

as the encryption map corresponding to the key a. The ciphertext

$$y = F_\mathrm{a}(\mathrm{x})$$

can be decrypted by the map $F_{\rm a}^{-1}$ written above. So the algorithm is symmetrical. The property 3 implies that different keys of length $< cn$ produce distinct ciphertexts.

We introduce the following directed graph $\phi = \phi(n)$ corresponding to maps ${f_t}^n$ and ${g_t}^n$ over K^n. Firstly we consider two copies of P (set of points) and L (set of lines) of the free module K^n. We connect point $p \in P$ with the line $l \in L$ by directed arrow if and only if there is $t \in {\rm Reg}(K)$ such that $f_t(p) = l$. Let t be the colour of such a directed arrow. Additionally we join $l \in L$ and $p \in P$ by directed arrow with the colour t if there is $t \in {\rm Reg}(K)$ such that $g_t(l) = p$. We can consider ϕ as finite automaton for which all states are accepting states. We have to chose point p (plaintext) as initial state. It is easy to see that f_t and g_t are the transition functions of our automaton. Let $t_1, \dots, t_s$ be the "program" i.e. sequence of colours from ${\rm Reg}(K)$. Then the computation is the directed pass p, $f_{t_1}(p) = p^1, g_{t_2}(p^1) = p^2, \dots$. If s is even then the last vertex is $f_{t_s}(p^{s-1})$, in case of odd s we get $g_{t_s}(p^{s-1}) = p^s$ as the result of the computation (encryption). The stop of the automata corresponds just to the absence of the next colour.

The inverse graph $\phi(n)^{-1}$ can be obtained by reversing of all arrows in ϕ. We assume that colours of arrow in ϕ and its reverse in ϕ^{-1} are the same. So we can consider $\phi(n)^{-1}$ as an automaton as well. Then the decryption procedure starting from the ciphertext p^s corresponds to the pass in ϕ^{-1} defined by sequence of colours $-t_s, -t_{s-1}, \dots, -t_1$.

Finally, we can consider well defined projective limit ϕ of automata ϕ^n, $n \to \infty$ with the transition function $P_t(x_1, x_2, \dots) = P(t, x_1, x_2, \dots)$ and $R_t(x_1, x_2, \dots) = R(t, x_1, x_2, \dots)$. In case of finite K we can use ϕ as a Turing machine working with the potentially infinite text in the alphabet K. Results of [114] allow to formulate the following statement.

Theorem 3. *For each commutative ring K there are a cubical polynomial maps P and R on K^∞ forming arithmetical dynamical system with the constant $c \geq 1/2$ such that for each string* r *of elements from* ${\rm Reg}(K)$ *the polynomial map $F_{\rm r}$ is cubical.*

The example as above has been defined explicitly in [118] in graph theoretical terms. The maps P and R will stand further for that particular example. Corresponding to (P, R) graphs $\phi(n)$ are strongly connected i.e. from the existence of directed pass from vertex v to w follows that w and v are connected by a directed pass. So connected components of $\phi(n)$ are well defined.

We combine the encryption process $F_{\rm r}$ corresponding to finite automaton $\phi(n)$ and string r of elements from ${\rm Reg}(K)$ with two invertible sparse affine transformation ${\rm Af}_1$ and ${\rm Af}_2$ and use the composition ${\rm Af}_1 \times F_{\rm a} \times {\rm Af}_2$ as encryption map. We refer to such a map as *deformation* of $F_{\rm r}$. In case of

$\mathrm{Af}_1 = \mathrm{Af}_2{}^{-1}$ we use term *desynchronization*. In case of desynchronization the ciphertext is always distinct from the plaintext. We assume that Af_1 and Af_2 are parts of the key. Deformated or desynchronised encryption is much more secure, because it prevents adversary to use group automorphisms and special ordering of variables during his/her attacks.

In the case of deformation with fixed Af_1 and Af_2 and flexible r the property that the different passwords of kind r lead to different ciphertexts is preserved, but the situation, where the plaintext and corresponding ciphertext are the same can happen. Anyway the probability of such event is $1/|V|$, where $V = K^n$ is the plainspace.

1.4.2. Watermarking equivalence and hidden discrete logarithm

The following statement is published in [118].

Theorem 4. *Let $\phi(n)$, $n \geq 6$ be the directed graph with the vertex set K^{n+1} defined above for the pair (P, R).*

(i) There are the tuple $a = a(\mathrm{x})$, $\mathrm{x} \in K^{n+1}$ of quadratic polynomials $a_2, a_3, \ldots, a_t$, $t = [(k+2)/4]$ in $K[x_0, x_1, \ldots, x_n]$ such that for each pair of vertices u and v from the same connected component we have $a(u) = a(v)$.

(ii) For any $t-1$ ring elements $x_i \in K$), $2 \leq t \leq [(n+2)/4]$, there exists a vertex v of $\phi(n)$ for which $a(v) = (x_2, \ldots, x_t) = (x)$. So classes of equivalence relation $\tau = \{(u, v) | a(u) = a(v)\}$ are in one to one correspondence with the tuples in K^t.

(iii) The equivalence class C for the equivalence relation τ on the set $K^{n+1} \cup K^{n+1}$ is isomorphic to the affine variety $K^t \cup K^t$, $t = [4/3n] + 1$ for $n = 0, 2, 3 \bmod 4$, $t = [4/3n] + 2$ for $n = 1 \bmod 4$.

We refer to τ as watermarking equivalence and call C as above generalised connected component of the graph,

Let $|K| = d$ and η numerating function i.e bijection between K and $\{0, 1, \ldots, d-1\}$. For each tuple $\mathrm{t} = (t_0, t_1, \ldots t_s) \in K^s$ we consider its number $\eta(\mathrm{t}) = \eta(t_0) + \eta(t_1)d + \cdots + \eta(t_s)d^s$. Let $\mathrm{Reg}(K) = b \geq 2$, μ be the bijection between $\mathrm{Reg}(K)$ and $\{0, 1, \ldots, b-1\}$. We obtain reg(t) by taking the string of digits for $\eta(\mathrm{t}) = l_0 + l_1 b + \cdots + l_{j-1} b^{j-1}$ base b and computing μ^{-1} for each digit. So $\mathrm{reg}(\mathrm{t}) = (\mu^{-1}(l_0), \mu^{-1}(l_1), \ldots, \mu^{-1}(l_{j-1})$ is a string of characters from the alphabet $\mathrm{Reg}(K)$.

The symmetrical algorithm with the key management (see [108] and further references).

Correspondents Alice and Bob are taking smallest prime number p from interval $(b^{[(n+5)/2c]}, b^{[(n+5)/2]})$, where c is some constant $> 3/2$ and some number m, $m < p$. Alice takes the plaintext x, then computes string $a(\mathrm{x})$ (see previous theorem), then $z = \eta(a(\mathrm{x}))$ and $u = z^m \bmod p$. She treats u as integer and takes string $d(\mathrm{x}) = \mathrm{reg}(u)$ of characters from $\mathrm{Reg}(K)$. Her encryption is $\mathrm{Af}_1 \times {F^n}_{d(\mathrm{x})} \times \mathrm{Af}_2$. We think that numbers m, c and fixed maps Af_i, $i = 1, 2$ are parts of the key.

Let $C_n(\mathrm{x}) = C(\mathrm{x})$ be the encryption function corresponding to deformation of dynamical system. The adversary may try to find the factorization $C_n(\mathrm{x}) = ((\mathrm{Af}_1)) \times F^n_{d(\mathrm{x})} \times \mathrm{Af}_2$, where $Af_i, i = 1, 2$ are unknown and the function $d(\mathrm{x})$ is $\mathrm{reg}((\eta(a(\mathrm{x})^m))$, where m is also unknown. During his active attack he can compute finite number of values $C(\mathrm{x}_i)$, $i \in J$ and use this information for finding the factorization. The following heuristic argument demonstrates that such a task is not easy.

Let as assume that affine transformation Af_i, $i = 1, 2$ are known for adversary. Notice that finding them can be very difficult. Then the adversary can compute $\mathrm{d}_i = F^n_{d(\mathrm{x}_i)} = (\mathrm{Af}_1)^{-1} C(\mathrm{x}_i)(\mathrm{Af}_2)^{-1}$. The pass between vertices of the graph is unique. The Dijkstra algorithm is not suitable for finding the pass because the vertex space of the graph is the plainspace. But may be large group of automorphisms (see [118] and further references) will allow to find the pass. Then the adversary computes number $b_i = \eta(a(\mathrm{x}))^m$ modula known big prime. Still he is not able to find number m because of the complexity of discrete logarithm problem. So he has to take for the set $\{\mathrm{x}_i | i \in J\}$ the totality of representatives from classes of watermarking equivalence (transversal). So $|J| > O(|K|^{[1/4]})$ because of the theorem 2.

We use term *hidden discrete logarithm* to name modified algorithm because affine transformations do not allow the adversary to compute the class of watermarking equivalence containing the plaintext (base of the logarithm) and pass in the finite automaton corresponding to the value of the logarithm.

Notice that the problem of breaking the key here is considered as a problem of finding the decomposition of given multivariable polynomial into the *composition* (not a product) of special polynomials. So the problem is different from the one considered in [65]. Anyway it could be very important and interesting to find the decompositions of cubic multivariable polynomial F_{r} into the product of irreducible terms for the case of general string r of characters from the $\mathrm{Reg}(K)$.

The public key algorithms associated with the above dynamical system have been introduced in [11],[113], [114], [115].

Let us consider first *generalisation* and corresponding *public key mode*.

Let K be general commutative ring and $(x_0, x_1, \ldots, x_n)$ be the vector from K^{n+1} (the plainspace). We will compute the string

$$\mathrm{a(x)} = (a_2(\mathrm{x}), a_3(\mathrm{x}), \ldots, a_t(\mathrm{x}))$$

which allow us to determine the class of watermarking equivalence containing x.

Let as consider the *symbolic key* which is chosen sequence of functions $k_i \in K[y_1, y_2, \ldots, y_{t-1}]$, $i = 1, 2, \ldots, s$. We say that the symbolic key is regular where $s \leq [(n+5)/2]$. Let $F_{(k_1,k_2,\ldots,k_s)}$ be the map sending x to its image under the composition of $F_{k_i(a_2(\mathrm{x}),\ldots,a_t(\mathrm{x}))}$, $i = 1, 2, \ldots, s$. The encryption map will be $E = Af_1 \times F_{(k_1,k_2,\ldots,k_s)} \times Af_2$. It is clear that the map is invertible and "hidden discrete logarithm" method is a particular case of the above general encryption with special regular symbolic key.

Example 1. Let $K = F_p$ for the prime p. Each nonlinear transformation $F_{(k_1,k_2,\ldots,k_s)}$, $s \geq 1$ together with affine transformations of $AGL_{n+1}(p)$ generates symmetric group S_p^{n+1}. Reasonable choice for the polynomial k_i in the case of public mode is a polynomial of bounded degree given by the list of its pseudorandom coefficients.

Example 2. An interesting case for theoretical studies is $K = Z$. For the private key algorithm we can chose all k_i as Prime Approximation Polynomials (see previous section). Notice that we have infinitely many classes of watermarking equivalence, they form equiprobable space isomorphic to Z^t.

Example 3. Let K be the ring of Gaussian numbers i.e. complex numbers of kind $a+bi$, where a and b are integers. Then we can think about the realisation of private key algorithms on Probabilistic Machine (Quantum Computer, in particular). The result of single probabilistic computation E will be a vector c from C^{n+1}. The nearest to c vector from the lattice K^{n+1} within the metric of Hilbert Space C^{n+1} will be treated as an approximation of the ciphertext by single computation. Of course we can change the above metric on some other natural metric for C^{n+1} (Eucledean metric of R^{2n+2}, in particular). The algorithm of example 2 also can be implemented on Probabilistic Machine (the result is the nearest to c $\in R^{n+1}$ vector from Z^{n+1}. One of the last results on Quantum Cryptography the reader can find in [1].

Remark. We can consider some union L of classes of watermarking equivalence and consider the restriction ${\phi_n}'$ of ϕ_n onto L. The private and

public key algorithms corresponding to ${\phi_n}'$ have been considered in [110]. In the section 6 we consider much more general encryption schemes in terms of special automata corresponding to families of directed algebraic graphs with special colouring.

Chapter 2

Simple graphs with special arcs and Cryptography

2.1. Graphs with special walks, definitions and motivations

2.1.1. Walks on simple graphs and cryptography

A combinatorial method of encryption with a certain similarity to the classical scheme of linear coding has been suggested in [107]. The general idea is to treat vertices of a graph as messages and arcs of a certain length as encryption tools. We will study the quality of such an encryption in case of graphs of high girth by comparing the probability to guess the message (vertex) at random with the probability to break the key, i.e. to guess the encoding arc. In fact the quality is good for graphs which are close to the Erdös bound, defined by the Even Cycle Theorem.

In the case of parallelotopic graphs there is a uniform way to match arcs with strings in the certain alphabet. Among parallelotopic graphs we distinguish linguistic graphs of affine type whose vertices (messages) and arcs (encoding tools) both could be tuples over certain finite commutative ring.

This unit presents the description of graph theoretical approach to symmetric encryption as well as to the construction of public key algorithm.

We will show that our approach allows us to construct absolutely secure algorithms to encrypt a potentially infinite text even with some resistance to attacks of type (ii) (complexity of the disclosing the key is estimated by the large constant). In our examples such schemes based on the incidence graph of generalized polygon.

In case of *linguistic graphs*, when vertices of graph are tuples over $GF(q)$ as well as walks on graphs (passwords), we are able to construct *absolutely optimal schemes of encryption* which are asymptotical one time pads when q is growing.

In more practical situation when length of plaintext and password form a constant ratio r we are able to present an *optimal encryption schemes* based on linguistic graphs. They have good resistance to attacks of both types, resistance to the attack of type (ii) is increasing with growth of r. Last feature allow us

(1) to avoid in some situations the partition of encryption tools into private key and public key algorithms
(2) to consider the modifications of our algorithm for the use in public key fashion.

The theoretical resistance of well-known RSA algorithms to attacks of type (ii) rests on our believe that nobody can factor numbers fast. In the case of our encryption schemes based on linguistic graphs the idea based on fact that finding a pass between 2 given vertices of infinite k- regular tree

require non polynomial expression $f(k,d)$ for the number of steps (natural branching process give us $k(k-1)(d-)$ steps). If the distance d is unknown the problem getting harder, the complexity $f(k,d)$ is growing, when d is increasing.

One of the popular mathematical models of the procedure for sending a message is the following:

(1) treat the information, to be sent, as a vector $x = (x_1, \ldots, x_n) \in GF(q)^n$
(2) "encode" our message by computing $Ax = y$, where A is (m,n)- matrix over $GF(q)$
(3) send the message y (via radio, telephone network, etc.)
(4) our receiver detects a message y', which may be different from y, due to transmission errors.

There are several well-known problems associated with the above method:

(i) error detection: warn us if y is not an acceptable message
(ii) error-correction: find an acceptable message y" which is probably y
(iii) investigation of the complexity of decoding the message, and other problems of cryptography.

For questions of cryptography, it is very convenient to have square matrix $(m = n)$.

One of the popular schemes of linear coding is the following:

We treat our message as a polynomial $f(x)$ over $GF(q)$ (our tuple is an array of coefficients of $f(x)$). The linear coding procedure is just a multiplication of our $f(x)$ of degree $n-1$ by a polynomial $g(x)$, $\deg(g(x)) = t$, $t > 0$. Thus, y is just an array of coefficients of the polynomial $F(x) = f(x)g(x)$, $m = \deg F(x) = n + t - 1$. It is easy to see that we could get the same y' by a multiplication of our y with a certain matrix B. Again, it is convenient to treat y' as an array of coefficients of the polynomial $F^{'}(x)$. An initial error detection check is just to take the remainder $r(x)$ of $F^{'}(x)$ by modulo $g(x)$. If $r(x)$ is not the zero polynomial, then we received a message with errors. This scheme is very convenient in many situations in Coding Theory, but it is not a case of $n = m$.

The general scheme of linear coding is very popular because

(A) our messages (vectors) are strings in an alphabet $GF(q)$ and we have a natural matching between information and these messages,
(B) encoding tools (matrices) are also strings in the same alphabet,
(C) the encoding procedure is computationally effective.

There are some unpleasant moments possible in linear coding. For instance

(D) if $m = n$, the initial and encoded messages could be the same!

It is clear, that in case (D) the initial message is an eigenvector with eigenvalue 1.

The encryption procedure of linear coding could be quite insecure. For instance, in case of encryption via multiplication by a polynomial computation of gcd's of consecutive messages will break the key fast.

Let us consider the following general idea of walks on graphs as coding tools.

Let Γ be a simple graph and $V(\Gamma)$ its set of vertices. Let us refer to the sequence $\rho = (v_1, v_2, \dots, v_n)$, where $v_i \in V(\Gamma)$, $v_i \neq v_{i+2}$, $i = 1, \dots, n-2$, and v_i is adjacent to v_{i+1}, $i = 1, \dots, n-1$ and $\rho(v_1) = v_n$ as *the encoding sequence* and *the encoded vertex* of v_1. We refer to $(v_n, v_{n-1}, \dots, v_1)$ as the decoding sequence for v_n.

Let us imagine that our message is the password to a computer account. We have the decoded message and s attempts to get into the account. Then suppose that we use our Γ graph in an "open algorithm" fashion. This means that information about the graph is available and the length n of the encoding sequence is known. Let $p_{\text{key}}(\Gamma, n)$ be the probability to break the key, i.e to guess the encoding sequence for one attempt and decode the message.

Then suppose, that we have no information about the graph. The only known object is the set of vertices, or partition set containing the message in the case of a bipartite graph. The only way to get into the account is to "guess" the message. Let $p_{\text{mes}}(\Gamma, n)$ be the probability of a success in this "dark" situation.

The purpose of this paper is to consider special cases of graphs, for which there is some similarity with linear coding, it satisfies some of the properties A, B, C and completely avoids situation D. For some of them the probability $p_{\text{key}}(\Gamma, n)$ can be computed.

We may assume without loss of generality that the edges of k-regular graph are marked by symbols from some alphabet A (set of colors) in such way that neighbors of each vertex are of different color. A graph with such marking is a Deterministic Finite Automaton in case when all states are accepting and all arrows are invertible. The arc of Γ is determined by its initial vertex and a string over the alphabet, which is a sequence of colors for edges from the arc.

In case of parallelotopic graphs, which we shall define below, the coloring of edges is induced by the special coloring of vertices, such that the neighborhood of each vertex is a set of vertices with different colors.

In general situation, we can consider the neighbor $w = N_a(v)$ of vertex v in the graph Γ such that the color of edge $\{w, v\}$ is $a \in A$.

Let

$$N(x_1, x_2, \dots, x_t) = N_{x_t}(N_{x_{t-1}}(\dots(N_{x_1}(v))\dots))$$

in variables $x_i \in A$, $v \in V(\Gamma)$. As previously we will treat an element v of $V(\Gamma)$ is a plaintext, and sequence $v = v_0, v_1, v_2, \ldots, v_t$, where $v_i \Gamma v_{i+1}$ is the encryption tool. If $N_{a_{i+1}}(v_i) = v_{i+1}$ then the pass is uniquely defined by string (or word) $x_1, x_2, \ldots, x_t$ over the alphabet of "colors" and the initial vertex. M and "inverse" string is $x_t, x_{t-1}, \ldots, x_1$. Thus we identify walks on Γ with strings over the M.

The RSA algorithm demonstrated that the information for encryption (number pq) can be just part of the information for decryption (at least numbers p and q).

Let us consider such a situation ("encryption with secret") in case of graph encryption.

Let ϕ_w be the binary relation $\phi_w = \{(u, v) | v = N(a_1, a_2, \ldots, a_t)(v)\}$, w is the string $a_1, a_2, \ldots, a_t$. It is clear that for the encryption with the key w we do not need the information about our graph Γ we need just graph Γ_w of the binary relation ϕ_w. Let $N^w(v)$ be the operator of taking neighbor of the vertex v in the graph Γ_v. The usual situation is that the complexity of computation N^w is much worse than N^w if you do not know the decomposition

$$N^w = N_{a_1}(N_{a_2}(\cdots(N_{a_t})\cdots)).$$

So you may present the function N_w in the form

$$N^w = N^{w_1}(N^{w_2}(\cdots(N^{w_s})\cdots)),$$

where word w is a product (concatenation) of words w_1, $w_2 \cdots, w_s$ to make computation of N_w faster.

It is clear that to find the decomposition above could be a hard task even in case when the graph Γ is known.

You can give your correspondent the "public key " $N^{w_1}, \ldots, N^{w_s}$. He can encrypt, but he can not decrypt if computation of superpositions of $(N^{w_s})^{-1}$, $(N^{w_{s-1}})^{-1}$, ..., $(N^{w_1})^{-1}$ is sufficiently hard.

Let us discuss this approach further in the special case of *linguistic graphs*.

2.2. Graphs with special walks, definitions and motivations

2.2.1. Families of graphic coding schemes, optimal and absolutely optimal families

Let us refer to a pair (Γ, n) where Γ is the graph and n is the length of encoding arcs, as a *graphic coding scheme*. Let $d = d(\Gamma, n)$ be given by $(p_s(\Gamma, n))^d = p_m(\Gamma, n)$. We will use the term *quality coefficient* for d.

We call a family (Γ_i, n_i) , $V(\Gamma_i) \subset V(\Gamma_{i+1})$, $n_i \leq n_{i+1}$ a *proper family of coding schemes* if $\lim p_s(\Gamma_i, n_i) = 0$, $i \to \infty$.

We call a family (Γ_i, n_i), $V(\Gamma_i) < V(\Gamma_{i+1})$, $n_i \leq n_{i+1}$ a *family of optimal schemes* if the *quality function* $d(i) = d(\Gamma_i, n_i)$ is bounded.

It is clear, that an optimal family is a proper family of coding schemes. For an optimal family such that $d(i)$ tends to 1, we will use the term *absolutely optimal*.

We can assume that the edges of any graph Γ_i are marked by elements of some alphabet A_i, such that $|A_i|$ is a maximal degree of Γ_i and the graph with this marking is a Deterministic Finite Automaton (all states are accepting). Thus the arc of Γ_i is determined by its initial vertex v and a string over the alphabet, which is a sequence of marks for edges from the arc.

Let us examine two special cases of proper families of schemes (Γ_i, n_i).

(i) The case of chosen length of encoding arcs $n_i = N$, and unbounded A_i.

We need to construct a password as a string of given length. There are no restrictions on the size of our alphabet. Increasing the size of the alphabet leads to a reduction in the probability of being able to "break the key".

(ii) Case of bounded $|A_i|$ and unbounded encoding length. We have a chosen alphabet $A = A_i$ and are ready to encode text of any length in this alphabet.

2.2.2. Graphs of large girth and their cryptographic properties

The girth $g = g(\Gamma)$ of a graph Γ is the length of the shortest cycle in the graph. If the length of the encoding arc ρ of the graph of girth g is less then g, then $\rho(v) \neq v$ for any vertex v, and we never have situation D. Another important feature of a regular or bipartite biregular graph Γ of high girth is the existence of a closed formula for the probability $p_{\text{key}}(\Gamma, n)$ when $n < g/2$.

Lemma 1. *Let Γ be a k-regular graph of girth g. Then in the case of length $n < g/2$ of the decoding sequence, the probability of generating the correct message by applying the encoding sequence r times at random is $r/(k(k-1)^{n-1})$.*

Proof. Let us imagine that our message is the password to a certain account. We have the decoded message and r attempts to get into the account. So we are trying r different encoding sequences to recover the correct message. All of them give us different results because of the inequality for n ensuring

that there are no C_{2n}. So, the probability to get into the account for r steps is $r/(k(k-1)^{n-1})$.

□

Analogously, we can check the following statement

Lemma 2. *Let Γ be a bipartite (a,b)-biregular graph of girth g. Then the probability of generating the correct message after r random attempts applying decoding sequence of the length $n < g/2$ to a vertex of valency b, would be $r/(b(a-1)^s(b-1)^{s-1}$ for $n = 2s$ and $r/(b(a-1)^s(b-1)^s$ for $n = 2s+1$.*

We will use term *scheme of high girth* for (Γ, n) when $n < (g-2)/2$.

Corollary 2. *Let Γ be a bipartite (a,b)-biregular graph of girth g, and (Γ, n) be a scheme of high girth, with inputs of valency b.*

Then $p_{\rm key}(\Gamma, n) = 1/(b(a-1)^{[n/2]}(b-1)^{[n/2]})$.

Corollary 3. *We are counting probabilities $p_{\rm mes}$ and $p_{\rm key}$ in a equiprobable space. Thus the condition $p_{\rm mes} = p_{\rm key}$ corresponds to one time pad encryption scheme.*

Let (Γ_i, n_i) be a family of regular or bipartite biregular schemes of high girth with non decreasing k_i and non decreasing bidegrees a_i and b_i such that $a_i + b_i + n_i$ is unbounded. Then (Γ_i, n_i) is a proper family of schemes.

For instance, any sequence of schemes of high girth of unbounded degree or unbounded length of encoding arcs contains a subsequence of proper schemes.

The constructions of *absolutely optimal* families of schemes of high girth of increasing degree are connected with studies of some well-known problems in Extremal Graph Theory (see [11], [91]). Let $ex(v,n)$ be, as usual, the greatest number of edges (size) in a graph on v vertices, which contains no cycles C_3, C_4, ..., C_n.

From Erdös' Even Cycle Theorem and its modifications [11], [91] it follows that

$$ex(v, 2k) \le Cv^{1+1/k} \tag{2.1}$$

where C is a positive constant.

It is easy to see that the magnitude of the extremal family of regular graphs of given girth and of unbounded degree have to be on the Erdös upper bound (4.1). This bound is known to be sharp precisely when $k = 2, 3,$ and 5. Thus the problem of constructing absolutely optimal families of high girth is a difficult one. A bound similar to (2.1) can be obtained for the bipartite biregular graphs with a given logarithmic ratio of valencies. Bipartite biregular families of schemes of high girth have to be on this bound.

In the case of the optimal monotonic scheme corresponding to graphs of degree l_i and unbounded girth g_i we have

$$g_i \geq \gamma \log_{l_i-1}(v_i) \tag{2.2}$$

The last formula means that Γ_i, $i = 1, \ldots$ form an infinite family of graphs of large girth in the sense of N. Biggs [8] (see, also [7], [9], [10], [48], [57],[59],[61], [64], [66], [69],[70]) for examples of such families).

We have $\gamma \leq 2$, because of (2.1), but no family has been found for which $\gamma = 2$. Bigger γs correspond to more secure coding schemes. A. Lubotzky conjectured that $\gamma \leq 4/3$.

2.2.3. Parallelotopic graphs

In this subsection we will consider the *parallelotopic graphs* for which arcs can be identified naturally and effectively with words in a certain alphabet M without marking of edges. We can just paint the vertices of our graph for this purpose.

We say [107], [110] that $\Gamma = (\Gamma, M, \pi)$ is a *parallelotopic graph* over a finite set M if we have a surjective function $\pi : V(\Gamma) \to M$ such that for every pair (v, m), $v \in V(\Gamma)$, $m \in M$, there is a unique neighbour u of v satisfying $\pi(u) = m$.

We refer to the function π in the definition above as a *labelling*. It is clear that a parallelotopic graph $\Gamma = (\Gamma, M, \pi)$ is an $|M|$-regular graph. We can consider M as the set of colors, and π as a coloring of the vertices of Γ such that for any given vertex v, and any color m, there exists exactly one neighbour u of v of color m.

Let Γ be a parallelotopic graph. Let $N(t, v)$ be the operator taking the neighbour u with colour t of a vertex v of a parallelotopic graph Γ. If $(t_1, t_2, \cdots, t_n)$, $t_i \in M$ is a tuple such that $t_i \neq t_{i+2}$, then

$$(v, v_1 = N(t_1, v), v_2 = N(t_2, v_1), \cdots, v_n = N(t_n, v_{n-1}))$$

is the arc of the graph Γ which we can consider as an *encoding arc* for any chosen vertex v.

Let us refer to this tuple $\rho = (t_1, \cdots, t_n)$ over M as an *encoding tuple*. It is clear that $\rho^{-1} = (t_n, t_{n-1}, \cdots, t_1)$ is the "decoding tuple", because it corresponds to the decoding arc.

It is reasonable to consider the following modification of parallelotopic graphs.

Let Γ be a bipartite graph with partition sets P_i, $i = 1, 2$. Let M be a disjoint union of finite sets M_1 and M_2. We say that Γ is a *bipartite parallelotopic graph* over (M_1, M_2) if there exists a function $\pi : V(\Gamma) \to M$

such that if $p \in P_i$, then $\pi(p) \in M_i$ and for every pair (p, j), $p \in P_i$, $j \in M_i$, there is a unique neighbour u with given $\pi(u) = j$.

It is clear that the bipartite parallelotopic graph Γ is a $(|M_1|, |M_2|)$ - biregular graph.

We refer also to the function π in the definition of bipartite parallelotopic graph also as a *labelling*. We will often omit the term "bipartite ", because all our graphs are bipartite.

A surjective homomorphism $\eta : \Gamma_1 \to \Gamma_2$ of bipartite parallelotopic graphs Γ_1, $V(\Gamma_1) = P_1 \cup P_2$ and Γ_2, $V(\Gamma_2) = Q_1 \cup Q_2$ over the same(M_1, M_2) with labellings π_1 and π_2 such that $\pi_2(\eta(v)) = \pi_1(v)$ and $\eta(v) \in Q_i$ iff $v \in P_i$ is referred to as a *parallelotopic morphism* of graphs.

In this situation, we refer to a graph Γ_1 as the *parallelotopic cover* of Γ_2 and a graph Γ_2 as a *parallelotopic quotient* of Γ_1. It is clear that parallelotopic morphism is local isomorphism.

Let $Spec(\Gamma)$ be the spectrum of the graph Γ, i.e., the set of eigenvalues of the adjacency matrix for Γ.

Lemma 3. (see [107]) *Let $\phi : \Gamma_1 \to \Gamma_2$ be a parallelotopic morphism of finite bipartite graphs Γ_1 and Γ_2. Then* $\mathrm{Spec}(\Gamma_2)$ *is a subset of* $\mathrm{Spec}(\Gamma_1)$.

Let M^t be the Cartesian product of t copies of the set M.

We say that the graph Γ is a *linguistic graph* over the set M with parameters m, k, r, s if

(i) Γ is a bipartite parallelotopic graph over (M_1, M_2), $M_1 = M^r$, $M_2 = M^s$ with the set of points $I = M^m$ (inputs)

(ii) set of lines $O = M^k$ (outputs).

(i.e. M^m and M^k are the partition sets of Γ). It is clear that $m + r = k + s$.

We use the term *linguistic coding scheme* for a pair (Γ, n), where Γ is linguistic graph and $n < g$ is the length of encoding sequences.

We choose a bipartite graph in the definition above because regular trees are infinite bipartite graphs and many biregular finite graphs of high girth can be obtained as their quotients (homomorphic images).

For linguistic graphs our messages and coding tools are words over the *alphabet* M and we can use the usual matching between real information and vertices of our graph.

We use the term *linguistic graph over $GF(q)$ of affine type* when we have a linguistic graph with alphabet $M = GF(q)$ and the set of neighbours of any vertex v is an affine manifold over $GF(q)$, i.e. is the totality of solutions of a certain system of linear equations. Here, of course, the similarity with the classical scheme, will be stronger: our messages and encoding tools are tuples over $GF(q)$ again and we have some linearity conditions.

Let a linguistic graph Γ of affine type satisfy the additional condition: operator $N(t,x)$, $t=(y_1,\dots,y_s)$, $x=(x_1,\dots,x_m)$, $x\in P$ ($x\in L$) can be given by polynomial expressions f_P (respectively f_L) in the variables $y_1,\cdots,y_r$, $x_1,\cdots,x_n$ depending only on the type of vertex.

Then the "complexity coefficient" $L(\Gamma)=(\deg f_P+\deg f_L)/2$ is a rough measure of the complexity of encoding procedure. The following section contains examples of linguistic graphs of both "good" quality and complexity coefficients.

2.3. Existence of graphs with special walks

In the previous section we gave several definitions of graphs with special walks. We also considered special morphisms and spectral properties of such graphs. We will be in trouble if no such objects exist. To show the existence of all objects defined above we need an example of a family of absolutely optimal linguistic graphs of affine type. We consider a several families of such graphs in this section.

Example 1. Let $P=\{(x_1,x_2)|x_i\in GF(q)\}$, $L=\{[y_1,y_2]|y_i\in GF(q)\}$. Let us define an incidence relation $I1$ as: $(a,b)I1[x,y]$ if and only if $y-b=xa$. Let us consider the function $\pi: P\cup L\to GF(q)$, such that $\pi((x_1,x_2))=x_1$, $\pi([y_1,y_2])=y_1$. It is easy to check that π is a labelling for the graph $I1=I1_q$. It defines a linguistic coding scheme $(I1,2)$ with parameters (1, 1, 2, 2) of affine type over $GF(q)$. We will show that the girth of $I1$ is at least 6. It is clear that the complexity $L(I1_q)=2$, $p_l(I1_q,2)=1/q(q-1)$ and $p_d=1/(q^2)$. Thus $(I1_q,2)$ is a family of absolutely optimal linguistic schemes of affine type.

Example 2. Let

$$P=\{(x_1,x_2,x_3)|x_i\in GF(q)\},$$
$$L=\{[y_1,y_2,y_3]|y_i\in GF(q)\}.$$

Let us define an incidence relation $I2$ as: $(a,b,c)I2[x,y,z]$ if and only if

$$y-b=xa$$
$$z-c=xb$$

Let us assume that $\pi((x_1,x_2,x_3))=x_1$ and $\pi([y_1,y_2,y_3])=y_1$. It is clear, that $I2$ defines a family of linguistic schemes of affine type over $GF(q)$ with parameters (1, 1, 3, 3). The complexity $L(I2_q)=5/2$ because $\deg f_P=2$ and $\deg f_L=3$. We will show that the girth of $I2$ is at least 8. It is clear that $p_s(I2_q,3)=1/q(q-1)^2$ and $p_m(I2_q,3)$ is q^3. Thus $(I2_q,3)$ is a family of absolutely optimal linguistic schemes of affine type.

Example 3. Let

$$P = \{(x_1, x_2, x_3, x_4, x_5) | x_i \in GF(q)\},$$
$$L = \{[y_1, y_2, y_3, y_4, y_5] | y_i \in GF(q)\}.$$

Let us define an incidence relation $I3 = I3_q$ as: $(a, b, c, d, e)I3[x, y, z, u, v]$ if and only if

$$y - b = xa$$
$$z - 2c = -2xb$$
$$u - 3d = -3xc$$
$$2v - 3e = 3zb - 3yc - ua$$

From the equations above, it follows that π: $\pi((x_1, x_2, x_3, x_4, x_5)) = x_1$ and $\pi([y_1, y_2, y_3, y_4, y_5]) = y_1$ is a labelling for $I3_q$.

We will show that (see Proposition 9.7), for char$GF(q) > 3$ the girth of this graph is at least 12. Directly from the equations above we can get that $I3$ defines the linguistic coding scheme with parameters (1, 1, 5, 5) of affine type over $GF(q)$. The complexity $L(I3_q)$ is 7/3, because of $\deg f_L = 3$ and $\deg f_P = 4$. It is clear that $p_s = q(q-1)^4$, $p_m = q^5$ and $(I3_q, 5)$ is an absolutely optimal family of schemes.

Example 4. Let $GF(q^2)$ be the quadratic extension of $GF(q)$ and $x \to x^q$ be the Frobenius automorphism of $GF(q^2)$. Let

$$P = \{(x_1, x_2, x_3) | x_1 \in GF(q), x_2 \in GF(q^2), x_3 \in GF(q)\},$$
$$L = \{[y_1, y_2, y_3] | y_1 \in GF(q^2), y_2 \in GF(q^2), y_3 \in GF(q)\}.$$

Let us define the incidence relation $I4 = I4_q$ as: $(a, b, c)I4[x, y, z]$ if and only if

$$y - b = xa$$
$$z - c = ay + ay^q.$$

It is clear that rules $\pi(x_1, x_2, x_3) = x_1$ and $\pi([y_1, y_2, y_3]) = y_1$ define the parallelotopic scheme of affine type over the $GF(q)$ (but not over the $GF(q^2)$). Its parameters are $(1, 2, 4, 5)$. Complexity of $L(I_4)$ for this scheme of affine type is 2 because $\{f_P, f_L\} = \{2, 2\}$. We will show that the girth of $I4$ is at least 8 (see Proposition 6.7). It is easy to check that $I4(q)$ is a family of absolutely optimal graphs.

Examples 1 – 3 give us families of graphs with sizes on the Erdös bound, and Example 4 gives examples of graphs with the sizes on the similar bound for biregular graphs of given degree. All the examples above are absolutely optimal families of a coding scheme.

2.4. Folders of graphs

The following proposition follows directly from the lemma 2.1.

Proposition 1. *Let Γ_i be a family of regular or bipartite biregular graphs of encreasing girth g_i and bounded degree. Then $(\Gamma_i, [(g_i - 1)/2])$ is a proper family of schemes.*

A natural way of constructing families of regular or bipartite biregular graphs of increasing girth is taking successive finite quotients (homomorphic images) of regular or biregular trees, which are graphs of infinite girth.

We need homomorphisms which are local isomorphisms, because in this case all quotients will have the same valency as the initial tree. If the initial tree is the projective limit of our quotients, then the quotients form a family of graphs of increasing girth.

So instead of using algebraic tools like polynomials over finite fields, we can just go to an infinite forest, choose a regular tree, cut it into proper pieces and relax. In fact, even doing the lumberjack's job in a fresh combinatorial air we can feel the smell of algebraic moonshine.

Folders of graphs (see [110])

For the purpose of convenient encoding by graphs of "potentially infinite" text over a finite alphabet (like the External alphabet of a Turing machine), we need an infinite family of parallelotopic graphs of increasing girth, with a hereditary property: we can add a new part of text, and encode the entire text in a larger graph, in such a way that the encoding of the initial part will be the same. This leads to the idea of a *folder* of parallelotopic graphs.

A folder F is a family Γ_j, $j = 1, 2, \ldots$ of graphs and homomorphisms $t_{i,j}$ satisfying the following properties:

(P_1) The Γ_i are parallelotopic (or bipartite parallelotopic) graphs over a finite set M with local labellings denoted by π.
(P_2) For any pair i, j of positive integers, $i > j$, there is a parallelotopic morphism $t_{i,j}$ from Γ_i to Γ_j.
(P_3) $t_{i,j} \circ t_{j,k} = t_{i,k}$ for $i > j > k$ (commutative properties)

Let us assume the existence of the projective limit Γ of Γ_i. We refer to Γ as the *cover of folder* Γ_i.

If Γ is a forest we refer to the folder as a free parallelotopic (bipartite parallelotopic) folder. It is clear that in this case the Γ_i, $i = 1, \ldots$ form an infinite family of graphs of unbounded girth. There is a canonical parallelotopic morphism $t_i : \Gamma \to \Gamma_i$. If T is a connected component of the forest

Γ then $t_i(T)$ is a connected component of $t_i(\Gamma)$ and family $t_i(T)$ is a free folder with the cover T.

Remark 1. Let Γ_i be a free folder over the $GF(q)$, where the cover Γ is a q-regular tree. We could construct the "Theory of Γ_i-codes" in which the distance in the graph Γ_i would play the role of a Hamming metric in the classical case of linear codes. Of course, the Hamming metric is distance-transitive, i.e., for each k the automorphism group acts transitively on pairs of vectors at a distance k. The distance in the graph Γ_i may not be distance transitive, but we have an "asymptotical" distance transitivity, because of the distance transitivity of the tree Γ and the fact that $\lim(\Gamma_i) = \Gamma$.

Remark 2. Let $\phi : \Gamma_1 \to \Gamma_2$ be a parallelotopic morphism of parallelotopic graphs over M and $s = (t_1, t_2, \cdots, t_k)$ be an encryption tuple, then the operators $enc_i : v \to v^s$, $v \in V(\Gamma_i)$, $i = 1, 2$ satisfy

$$enc_1\phi = enc_2$$

Let $\rho : \Gamma_1 \to \Gamma_2$ be a parallelotopic morphism of (bipartite parallelotopic) graphs over the same set M such that $V(\Gamma_1)$ is a Cartesian product $D \times V(\Gamma_2)$ for some set D and ρ is the canonical projection $\rho((d, v)) \to v$, $d \in D$, $v \in V(\Gamma_2)$. We will say that ρ is a *triangular parallelotopic morphism.* For such a morphism, hereditary properties of encryption schemes (Γ_1, k) and (Γ_2, k) are stronger. We will use the word triangular instead of parallelotopic to get *triangular cover* and *triangular quotient.* We will use the term *triangular folder* for a parallelotopic folder for which all morphisms are triangular.

Let ρ: $\Gamma_1 \to \Gamma_2$ be a triangular morphism from a graph Γ_1 to Γ_2. Let us consider the sets of vertices $V_1 = V(\Gamma_1)$ and $V_2 = V(\Gamma_2)$ and the vector spaces F_i, $i = 1, 2$ consisting of the sets of functions $\{f | V_i \to R\}$ over the field R of real numbers and subspace

$$F_\rho = \{f \in F_1 | (\rho(x) = \rho(y)) \to (f(x) = f(y))\}.$$

The vectors b_u , $u \in V_1$ such that $b_u(u) = 1$ and $b_u(x) = 0$, $x \in V_1 - \{u\}$ form a natural basis of F_1 and the elements $c_v = \sum_{u|f(u)=v} b_u$, $v \in V_2$, form a basis of F_η.

Let A_i , $i = 1, 2$ be the adjacency matrix for Γ_i. In the triangular case F_ρ is an invariant subspace of A_1 and the restriction of A_1 to F_ρ coincides with A_2 with respect to some basis c_v, $v \in V_2$. Let ρ be the natural projection of F_ρ on F_2.

We then have:

Lemma 4. *If Γ_1 is a finite triangular cover of Γ_2, then*

(i) Spec(Γ_2) is a subset of Spec(Γ_1).
(ii) If f_2 is an eigenvector of Γ_2, then there is an eigenvector f_1 of Γ_1 from F_η , such that $f_1 = \rho(f_2)$ and the eigenvectors f_i, $i = 1, 2$, have the same eigenvalue.

2.5. Existence of free triangular optimal folders

In this section we will use the graphs $D(k, q)$ ([57], [59]) for the encryption procedure.

Let q be a prime power, and let P and L be two countably infinite dimensional vector spaces over $GF(q)$. Elements of P will be called *points* and those 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 coordinates of points and lines introduced in [50]:

$$(p) = (p_1, p_{11}, p_{12}, p_{21}, p_{22}, p'_{22}, p_{23}, \ldots, p_{ii}, p'_{ii}, p_{i,i+1}, p_{i+1,i}, \ldots),$$
$$[l] = [l_1, l_{11}, l_{12}, l_{21}, l_{22}, l'_{22}, l_{23}, \ldots, l_{ii}, l'_{ii}, l_{i,i+1}, l_{i+1,i}, \ldots).$$

We now define an incidence structure (P, L, I) 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 coordinates hold:

$$\begin{aligned} l_{11} - p_{11} &= l_1 p_1 \\ l_{12} - p_{12} &= l_{11} p_1 \\ l_{21} - p_{21} &= l_1 p_{11} \\ l_{ii} - p_{ii} &= l_1 p_{i-1,i} \\ l'_{ii} - p'_{ii} &= l_{i,i-1} p_1 \\ l_{i,i+1} - p_{i,i+1} &= l_{ii} p_1 \\ l_{i+1,i} - p_{i+1,i} &= l_1 p'_{ii} \end{aligned} \qquad (2.3)$$

(The last four relations are defined for $i \geq 2$.) This incidence structure (P, L, I) we denote as $D(q)$. 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 $(p)I[l]$.

To facilitate notation in future results, it will be convenient for us to define

$$
\begin{aligned}
p_{-1,0} = l_{0,-1} = p_{1,0} = l_{0,1} &= 0,\\
p_{0,0} = l_{0,0} &= -1,\\
p'_{0,0} = l'_{0,0} &= 1,\\
p_{0,1} &= p_2,\\
l_{1,0} &= l_1,\\
l'_{1,1} &= l_{1,1},\\
p'_{1,1} &= p_{1,1},
\end{aligned}
$$

and to rewrite (2.3) in the form:

$$
\begin{aligned}
l_{ii} - p_{ii} &= l_1 p_{i-1,i}\\
l'_{ii} - p'_{ii} &= l_{i,i-1} p_1\\
l_{i,i+1} - p_{i,i+1} &= l_{ii} p_1\\
l_{i+1,i} - p_{i+1,i} &= l_1 p'_{ii}
\end{aligned}
$$

for $i = 0, 1, 2, \ldots$

Notice that for $i = 0$, the four conditions (2.3) are satisfied by every point and line, and, for $i = 1$, the first two equations coincide and give $l_{1,1} - p_{1,1} = l_1 p_1$.

For each positive integer $k \geq 2$ we obtain an incidence structure (P_k, L_k, I_k) as follows. First, P_k and L_k are obtained from P and L, respectively, by simply projecting each vector onto its k initial coordinates. The incidence I_k is then defined by imposing the first $k-1$ incidence relations and ignoring all others. For fixed q, the incidence graph corresponding to the structure (P_k, L_k, I_k) is denoted by $D(k, q)$. It is convenient to define $D(1, q)$ to be equal to $D(2, q)$. The properties of the graphs $D(k, q)$ that we are concerned with described in the following Proposition.

Proposition 2. *[57] Let q be a prime power, and $k \geq 2$. Then*

(i) $D(k, q)$ is a q–regular bipartite graph of order $2q^k$;
(ii) for odd k, $g(D(k, q)) \geq k + 5$;
(iii) for odd k and $q \equiv 1 \pmod{\frac{k+5}{2}}$, $g(D(k, q)) = k + 5$.

The following statement follows directly from Proposition 2 and the formula for a neighbour of the given vertex.

Proposition 3. *The graph $D(k, q)$, q odd, is a parallelotopic graph with the labelling π : $\pi((p)) = p_2$, $\pi([l]) = l_1$.*

$(D(k, q), [k + 3/2])$, $k = 1, 2, \ldots$ form an optimal family of linguistic coding schemes over $GF(q)$ of affine type with unbounded girth.

Scheme ($D(k,q)$, $[(k+3)/2]$) has parameters $(1,1,k,k)$, complexity coefficient 2 and quality coefficient $d_k \leq k/([k+5)/2)]) \leq 2$.

We have a natural one to one correspondence between the coordinates 2,3, ..., n, $\cdots$ of tuples (points or lines) and equations. It is convenient for us to rename by $i+2$ the coordinate which corresponds to the equation with the number i and write $[l] = [l_1, l_3, \ldots, l_n, \ldots]$ and $(p) = (p_1, p_3, \ldots, p_n, \ldots)$ (line and point in "natural coordinates").

Let η_i be the map "deleting all coordinates with numbers $> i$" from $D(q)$ to $D(i,q)$, and $\eta_{i,j}$ be map "deleting all coordinates with numbers $> i$ " from $D(j,q)$, $j > i$ into $D(i,q)$.

The following statement follows directly from the definitions:

Proposition 4. *(see, [57], [59]) $D(i,q), \eta_{i,j}$ is a free triangular linguistic folder over $GF(q)$. It has a forest $D(q)$ as a cover.*

Example 5. Let $k \geq 6$, $t = [(k+2)/4]$, and let

$$u = (u_i, u_{11}, \cdots, u_{tt}, u'_{tt}, u_{t,t+1}, u_{t+1,t}, \cdots)$$

be a vertex of $D(k,q)$. (It does not matter whether u is a point or a line). For every r, $2 \leq r \leq t$, let

$$a_r = a_r(u) = \sum_{i=0,m} (u_{ii} u'_{r-i,r-i} - u_{i,i+1} u_{r-i,r-i-1}),$$

and $a = a(u) = (a_2, a_3, \cdots, a_t)$. (Here we define $p_{0,-1} = l_{0,-1} = p_{1,0} = l_{0,1} = 0$, $p_{00} = l_{00} = -1$, $p_{0,1} = p_1$, $l_{1,0} = l_1$, $l'_{11} = l_{11}$, $p'_{1,1} = p_{1,1}$).

In [62] the following statement was proved.

Proposition 5. . *Let u and v be vertices from the same component of $D(k,q)$. Then $a(u) = a(v)$. Moreover, for any $t-1$ field elements $x_i \in GF(q)$, $2 \leq it \geq [(k+2)/4]$, there exists a vertex v of $D(k,q)$ for which*

$$a(v) = (x_2, \ldots, x_t) = (x).$$

Let us consider the following equivalence relation τ : $u\tau v$ iff $a(u) = a(v)$ on the set $P \cup L$ of vertices of $D(k,q)$ ($D(q)$). The equivalence class of τ containing the vertex v satisfying $a(v) = (x)$ can be considered as the set of vertices for the induced subgraph $EQ_{(x)}(k,q)$ ($EQ_{(x)}(q)$) of the graph $D(k,q)$ (respectively, $D(q)$). When $(x) = (0, \cdots, 0)$, we will omit the index v and write simply $EQ(k,q)$.

Let $CD(k,q)$ be the connected component of $D(k,q)$ which contains $(0,0,\ldots,0)$. Let τ' be an equivalence relation on $V(D(k,K))$ ($D(q)$) such that the equivalences classes are the totality of connected components of this graph. According to Proposition 5, $u\tau v$ implies $u\tau' v$. If char $GF(q)$ is an odd number, the converse of the last proposition is true [62].

Proposition 6. *Let q be an odd number. Vertices u and v of $D(q)$ $(D(k,q))$ belong to the same connected component iff $a(u) = a(v)$, i.e., $\tau = \tau'$ and $EQ(q) = CD(q)$ $(EQ(k,q) = CD(k,q))$.*

The condition $charGF(q) \neq 2$ in the last proposition is essential. For instance, the graph $EQ(k,4))$, $k > 3$, contains 2 isomorphic connected components. Clearly $EQ(k,2)$ is a union of cycles $CD(k,2)$. Thus neither $EQ(k,2)$ nor $CD(k,2)$ is an interesting family of graphs of high girth. But the case of graphs $EQ(k,q)$, q is a power of 2, $q > 2$ is very important for coding theory.

Corollary 4. *Let us consider a general vertex*

$$x = (x_j, x_{1,1}, x_{2,1}, x_{1,2}, \cdots, x_{i,i}, x'_{i,i}, x_{i+1,i}, x_{i,i+1}, \cdots),$$

$j = 1$ or 2, $i = 2, 3, \cdots$ of the connected component $CD(k,F)$, which contains a chosen vertex v. Then coordinates $x_{i,i}$, $x_{i,i+1}$, $x_{i+1,i}$ can be chosen independently as "free parameters" from F and $x'_{i,i}$ could be computed succesively as the unique solutions of the equations $a_i(x) = a_i(v)$, $i = 1, \ldots$.

Theorem 5.

(*i*) *$EQ(k,q), \eta_{i,j}$ is a free triangular linguistic folder with complexity coefficient 2 for the forest $EQ(k,q)$.*
(*ii*) *Let d_k be a quality coefficient for the $EQ(k,q)$. Then d_k tends to 1.5 as $k \to \infty$.*

Proof. Statement (i) follows from Theorem 7.3. The cardinality of the set of points (lines) for $EQ(k,q)$ is a linear function of the form $3/4 + c$ where c is a certain constant, and $g(EQ(k,q)) \geq k+5$. Thus $d_k = (3/4k+c)/(k+5)/2$ giving (ii). □

Remark 1. $CD(k,q)$ is the family of linguistic coding schemes of affine type with the smallest function of cycle growth among the known families of unbounded degree and increasing girth.

Remark 2. The graph $EQ(k,4)$ form the family of linguistic coding schemes with the best known quality coefficient.

The existence of folders of monotonic linguistic coding schemes of high girth over any finite field is important for the encryption of "potentially infinite texts" . For instance, if we have a text in English, we can consider an injective function f from the alphabet, which contains letters, comma, dot, empty space, to $GF(29)$. We could apply the function f to each character in our text to identify of it with an element of the finite field. After that we can use a coding by a folder of linguistic graph of increasing girth of

affine type over $GF(29)$, which will guarantee that the encoded text will be different from the initial text. Of course we can have some semantic similarity, between initial and encoded text. For instance, in the case of Galsworthy's "Forsythe Saga" the encoded text theoretically could be a translation of it into Spanish, but the probability of this happening is very small.

2.6. Parallelotopic graphs of large girth and asymmetric algorithms

2.6.1. Linguistic graphs as a public keys

Let Γ be a parallelotopic graph of girth g, i.e. the graph without cycles of length $< g$. In this case we will use a little bit different matching between arcs and strings of colors then in case of general regular graphs.

Let $w = N_a(v)$ stands now for a neighbor of vertex v in graph Γ such that the color of the neighbor is $a \in M$. As it was before

$$N(x_1, x_2, \dots, x_t) = N_{x_t}(N_{x_{t-1}}(\dots(N_{x_1}(v))\dots))$$

in variables $x_i \in M$, $v \in V(\Gamma)$. As previously we will treat an element v of $V(\Gamma)$ is a plaintext, and sequence $v = v_0, v_1, v_2, \dots, v_t$, where $v_i \Gamma v_{i+1}$ is the encryption tool. If $N_{a_{i+1}}(v_i) = v_{i+1}$ then the pass is uniquely defined by string (or word) $a_1, a_2, \dots, a_t$ over the alphabet of "colors" M and "inverse" string $a_{t-1}, a_{t-2}, \dots, a_0$, here a_0 is the color of plaintext defines "decrypting" sequence of vertices. Thus we identify walks on Γ with strings over the M.

Let us discuss the approach of section 3 in the case of *linguistic graphs* Γ of rational (polynomial) type over commutative ring K.

In case of such graphs $M = K$ and the function $N_a(v)$ of taking neighbor of vertex $v = (y_1, y_2, \dots, y_t) \in I \cup O$ is a polynomial expression from variables y_i, $i = 1, k$. A degree of polynomial linguistic graph is maximum degrees of polynomial expressions for each $N_a(v)$ in variables y_i.

In this case N^{w_i}, $i = 1, 2, \dots, s$ are polynomial expressions P_i over the commutative field K of degree d_i. For simplicity let us assume that the graph is regular, i.e. $O = I = K^n$, and polynomial expression N^w is given by the list of its coefficients. If $\deg N^w = d$ then encryption of the given vertex could be done for not more than for $O((n^d)$ elementary steps. So, if your "public key" is given as the list of coefficients of monomial expressions in N^w, then the complexity of encoding procedure for your correspondent will be proportional to a size of this list.

You can do your encryption (or decryption) fast because your know factors N_{a_i} of N^w and their inverses.

What your correspondent need for the decryption of given message $(b_1, b_2, \ldots, b_n)$? He has to solve the system of polynomial equations

$$N_w(x_1, x_2, \ldots, x_n) = (b_1, b_2, \ldots, b_n).$$

This task is a classical hard problem of algebra (see [45] and further references). The system above can be investigated for $d^{O(n^2)}$ steps, where d is the maximal degree of polynomials. We can do better (d^{Cn}) if we know that the system is consistent.

If you have a family of polynomial linguistic graph of bounded degree, you may choose the dimension n such that your correspondent could encrypt but could not decrypt and use graph encryption in "public key fashion", because we should use the gap between computations of polynomial in given point and investigation of given system of equations.

2.6.2. Graph invariants as hidden parameters, dynamical keys

Let us consider the encryption scheme $(D(k,q),t)$, $t = (k+2)/4 - 1$ for which transformation $u \to N(b_1, b_2, \ldots, b_s)(u) = c$, $u \in D(k,q)$ maps plaintext p to the ciphertext c. We can upgrade the encryption process by the following procedure.

Let us treat parameters b_i as variables and use the substitution $b_i = F_i(a_2(u), a_3(u), \ldots, a_{t+1}(u)$ where $a_2(u), \ldots a_{t+1}(u)$ are invariants of the connected component which contains u, F_i are chosen polynomial expressions from t-variables. Your correspondent knows just the transformation $u \to c$, where

$$c(u) = N(F_1(a_2(u), \ldots, a_{t+1}(u)), \ldots, F_k(a_2(u), \ldots, a_{t+1}(u))$$

is given by the list of coefficients (public key). He can encrypt but could not decrypt in case of properly chosen parameters and polynomials F_i, $i = 1, \ldots, k$. You know graph $D(k,q)$ and its invariants $a_i(k)$. If you apply to e consecutively transformations $N_{d_{s-1}}, N_{d_{s-2}}, \ldots, N_{d_1}, N_{d_0}$, where $d_i = F_i(a_2(e), \ldots, a_{t+1}(e))$, $i = 1, \ldots, k-1$ and d_0 is the colour of plaintext, you obtain the plaintext u. Here we use the fact that u and c are vertices from the same component of $D(k,q)$. In the package CRYPTIM we use this scheme in case $s = t$, $\deg F_i \leq 1$, in particular, for the problem of digital signatures.

Remark 1. The probability to have same invariants $a_2, \ldots, a_{t+1}$ for two random messages is about $1/q^t$.

Remark 2. If we want to speed up the computation of $c(u)$ we may present it to our correspondent as product of several factors. For instance, we may construct such factors as a products of several $N_{b_i}(x)$.

2.7. The jump to commutative rings, dynamical systems and fast implementations

We consider below the natural generalisation $D(n.K)$ of the family of graphs $D(n,q)$, where $n > 2$ is positive integer and K is a commutative ring. New family is obtained just by change of F_q on K. Properties of such graphs over rings were considered in [107]. Similarly to D(n, q), new graphs are linguistic graphs, the family $D(n,K)$ $n = 1, 2, \ldots$ defines the projective limit $D(K)$, connected components of $D(K)$ can be observed via quadratic invariants. The encryption algorithms as above can be formally defined for more general graphs $D(n,K)$.

Unfortunately in the case of existence of zero divisors in the ring K the infinite graph $D(K)$ is not a tree. The girth for the members of the family $D(n,K)$ $n = 1, 2, \ldots$ is bounded by certain constant. So for finite rings we obtaining a family of increasing girth if and only if $K = F_q$.

Fortunately there is an option to work with the special family of directed graphs defined in terms of simple graphs $D(n,K)$ which allow us to save a situation.

let P and L be two copies of Cartesian power K^N, where K is the commutative ring and N is the set of positive integer numbers. Elements of P will be called *points* and those 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 keep the notation for co-ordinates of points and lines of $D(q)$ for the case of general commutative ring 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 K, such that only finite number of components are different from zero.

We now define an incidence structure (P, L, I) as follows. We say the point (p) is incident with the line $[l]$, and we write $(p)I[l]$, if the relations (2. 3) between their co-ordinates hold:

For each positive integer $k \geq 2$ we obtain an incidence structure (P_k, L_k, I_k) as follows. First, P_k and L_k are obtained from P and L, respectively, by

simply projecting each vector onto its k initial coordinates with respect to the above order. The incidence I_k is then defined by imposing the first $k-1$ incidence equations and ignoring all others. The incidence graph corresponding to the structure (P_k, L_k, I_k) is denoted by $D(k, K)$.

Now we will construct aspecial directed graphs. E. Moore [75] used term *tactical configuration* of order (s, t) for biregular bipartite simple graphs with bidegrees $s + 1$ and $r + 1$. It corresponds to incidence structure with the point set P, line set L and symmetric incidence relation I. Its size can be computed as $|P|(s + 1)$ or $|L|(t + 1)$.

Let $F = \{(p, l)|p \in P, l \in L, pIl\}$ be the totality of flags for the tactical configuration with partition sets P (point set) and L (line set) and incidence relation I. We define the following irreflexive binary relation ϕ on the set F.

Let (P, L, I) be the incidence structure corresponding to regular tactical configuration of order t.

Let $F_1 = \{(l, p)|l \in L, p \in P, lIp\}$ and $F_2 = \{[l, p]|l \in L, p \in P, lIp\}$ be two copies of the totality of flags for (P, L, I). Brackets and parenthesis allow us to distinguish elements from F_1 and F_2. Let $DF(I)$ be the directed graph (double directed flag graph) on the disjoint union of F_1 with F_2 defined by the following rules

(i) $(l_1, p_1) \to [l_2, p_2]$ if and only if $p_1 = p_2$ and $l_1 \neq l_2$,
(ii) $[l_2, p_2] \to (l_1, p_1)$ if and only if $l_1 = l_2$ and $p_1 \neq p_2$.

Let $DE(n, K)$ $(DE(K))$ be the double directed graph of the bipartite graph $D(n, K)$ $(D(K)$, respectively). Remember, that we have the arc e of kind $(l^1, p^1) \to [l^2, p^2]$ if and only if $p^1 = p^2$ and $l^1 \neq l^2$. Let us assume that the colour $\rho(e)$ of arc e is $l^1_{1,0} - l^2_{1,0}$.

Recall, that we have the arc e' of kind $[l^2, p^2] \to (l^1, p^1)$ if and only if $l^1 = l^2$ and $p^1 \neq p^2$. let us assume that the colour $\rho(e')$ of arc e' is $p^1_{1,0} - p^2_{1,0}$.

If K is finite, then the cardinality of the colour set is $(|K| - 1)$.

Graph $DE(k, K)$ is the double flag graph for $D(k, K)$. We assume that $k \geq 6$and $t = [(k + 2)/4]$. Each flag f from $F_1 \cup F_2$ contains the unique point u $u = (u_{01}, u_{11}, \cdots, u_{tt}, u'_{tt}, u_{t,t+1}, u_{t+1,t}, \cdots)$ of D(k, K) . For every r, $2 \leq r \leq t$, let

$$a_r(f) = a_r(u) = \sum_{i=0,r} (u_{ii}u'_{r-i,r-i} - u_{i,i+1}u_{r-i,r-i-1}),$$

and $a = a(u) = (a_2, a_3, \cdots, a_t)$.

Let Reg$K$ be the totality of regular elements, i.e. not zero divisors. Let us delete all arrows with colour, which is a zero divisor. New graph $RDE(t, K)$ $(RD(K))$ with the induced colouring is the automaton in the alphabet Reg(K).

Let $P_t(x_{1,0}, x_{0,1}, x_{11}, \dots)$ and $R_t(x_{1,0}, x_{0,1}, x_{11}, \dots)$ are the transition function of infinite graph $RD(K)$ of taking the neighbour of vertex from the first and second copy of the flag set for $D(K)$ alongside the arrow of colour t. As it follows from the results of [118] functions P_t and R_t define the arithmetical dynamical system. It means that we can use general private and public keys algorithms which were considered in Chapter 1.

Let us consider the description of such algorithms in the particular case of the dynamical system corresponding to the family of directed graphs $RDE(n, K)$. We assume that the finite ring K contains at least 3 regular elements (non zero divisors). We start from the public key encryption.

The set of vertices of the graph $RDE(n, K)$ is a union of two copies free module K^{n+1}. Let $C(K^{n+1})$ be the Cremona group of the variety K^{n+1} containing all bijective polynomial maps for which the inverse map is also polynomial. In the simplest case of finite field F_p, where p is a prime number $C({F_p}^{n+1})$ is a symmetric group $S_{p^{n+1}}$. The Cremona group $C(K^{n+1})$ contains the group of all affine invertible transformations, i.e. transformation of kind $\mathrm{x} \to \mathrm{x}A + \mathrm{b}$, where $\mathrm{x} = (x_1, x_2, \dots, x_{n+1}) \in C(K^{n+1})$, $\mathrm{b} = (b_1, b_2, \dots, b_{n+1})$ is a chosen vector from $C(K^{n+1})$ and A is a matrix of liner invertible transformation of K^{n+1}.

Graph $RDE(n, K)$ is a bipartite directed graph. We assume that the plaintext K^{n+1} is a point flag $f = (p_1, p_2, \dots, p_{n+1})$ (a pair containing point $(p_1, p_2, \dots, p_n)$ of $D(n, K)$ and the colour p_{n+1} of neighbouring line from $D(n, K)$). Alice choses two invertible sparse affine transformations T_1 and T_2 of K^{n+1}. and the string $(\beta_1, \beta_2, \dots, \beta_l)$ of regular colours for $RDE(n, K)$, such that $\beta_i \neq -\beta_{i+1}$ for $i = 1, 2, \dots, l-1$ (irreducibility condition). This data form the key. Alice keeps chosen parameters secret in our implementation we use affine transformations which maps f to string $(p_1 + a_2p_2 + a_3p_3 + \dots + a_{n+1}p_{n+1}, p_2, p_3, \dots, p_{n+1})$. Let N_α, $\alpha \in \mathrm{Reg}(K)$ be the operator of taking the neighbour of vertex v alongside the arrow with the colour α in the directed graph $RDE(n, K)$. She computes symbolically the map $N^l = N_{\beta_1} \times N_{\beta_2} \cdots \times N_{\beta_l}$. Alice computes the public rule g as the symbolic composition of T_1, N and T_2. The case $T_2 = {T_1}^{-1}$ is a special form of general algorithm.

In the case of $RDE(k, K)$ the degree of transformation N^l is 3, independently on the choice of length l, parameter k and ring K. So the public rule is a cubical polynomial map of the free module K^{n+1} onto itself. In case of finite field the algorithm is equivalent to the public rule considered in [99]. The public user (Bob) has the cubical map g only. He sends to Alice the encrypted massage $g(f)$.

Notice that Alice can decrypt with numerically implemented

$$D = {T_2}^{-1} N_{-\beta_l} \times N_{-\beta_2} \cdots \times N_{-\beta_1} {T_1}^{-1}.$$

Functions $T_1 N^l T_2$ and E form her private key algorithm.

2.7.1. Time evaluation of the private key encryption for Alice

We have implemented computer application [54], which uses family of graphs $RDE(n, K)$ for *private key* cryptography. To achieve high speed property, commutative ring $K = Z_{2^k}$, $k \in \{8, 16, 32\}$, with operations $+, \times$ modulo 2^k. Parameter n stands for the length of plaintext (input data) and the length of ciphertext. We mark by $G1$ the algorithm with $k = 8$, by $G2$ the algorithm with $k = 16$, and by $G4$ the algorithm with $k = 32$. So $Gi, i \in 1, 2, 4$ denotes the number of bytes used in the alphabet (and the size of 1 character in the string).

The alphabet for password is the same K as for the plaintext. For encryption we use the scheme presented in the previous unit.

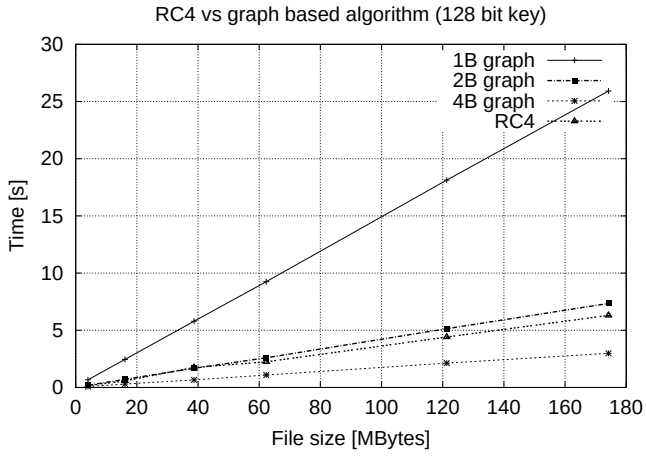
All the test were run on computer with parameters:

— AMD Athlon 1.46 GHz processor
— 1 GB RAM memory
— Windows XP operating system.

The program was written in Java language. Well known algorithms RC4 and DES which were used for comparison have been taken from Java standard library for cryptography purposes - *javax.crypto*.

2.7.2. Comparison of our symmetric algorithm with RC4

RC4 is a well known and widely used stream cipher algorithm. Protocols SSL (to protect Internet traffic) and WEP (to secure wireless networks) uses it as an option. Nowadays RC4 is not secure enough and not recommended for use in new system. Anyway we chose it for comparison, because of its popularity and high speed.



File [MB]	RC4 [s]	G1 [s]	G2 [s]	G4 [s]
4	0.15	0.67	0.19	0.08
16.1	0.58	2.45	0.71	0.30
38.7	1.75	5.79	1.68	0.66
62.3	2.24	9.25	2.60	1.09
121.3	4.41	18.13	5.14	2.13
174.2	6.30	25.92	7.35	2.98

Figure 2.1. RC4 vs high girth graph based algorithm (128 bit password)

RC4 is not dependent on password length in terms of complexity, and our algorithm is. Longer password makes us do more steps between vertices of graph. So for fair comparison we have used fixed password length equal suggested upper bound for RC4 (16 Bytes).

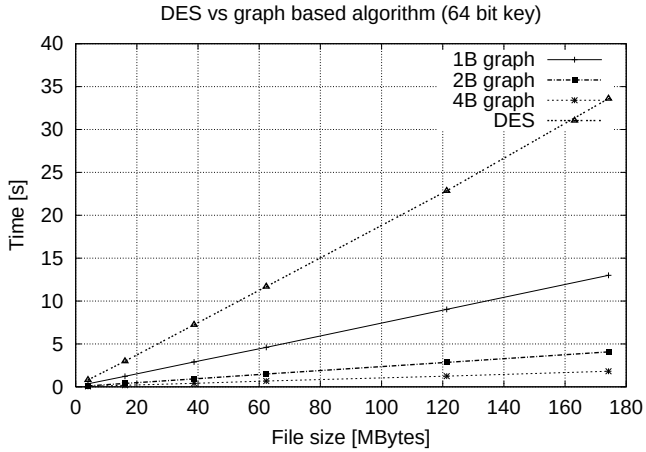
File [MB]	G1 [s]	G2 [s]	G4 [s]
4	0.04	0.02	0.01
16.1	0.12	0.10	0.08
38.7	0.32	0.24	0.20
62.3	0.50	0.40	0.30
121.3	0.96	0.76	0.60
174.2	1.39	0.96	0.74

Table 2.1. Time grow for $\mathbf{A}_n E_{\bar{a}} \mathbf{A}_n^{-1}$ for chosen operator $\mathbf{A}_n$

2.7.3. Comparison with DES

In the next test we have compared our algorithm with popular block cipher DES (Data Encryption Standard). DES is more complicated ,and have better cryptographical properties than RC4, but it is much slower.

The version of DES implemented in Java library uses 64 bit password and makes from it 56 bit key (due to documentation). In our comparison (see figure (2.2)) we used the password of the same length.

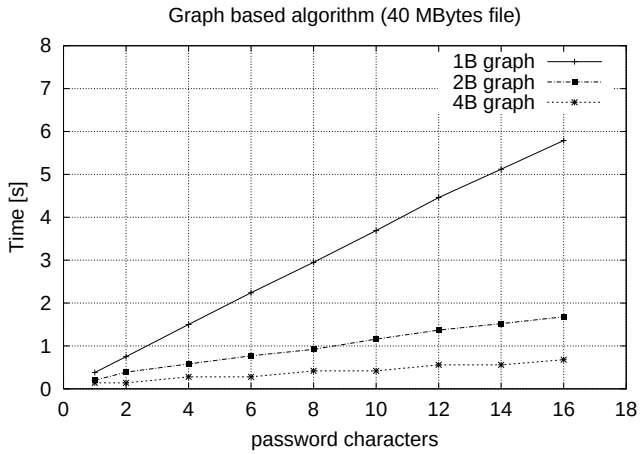


File [MB]	DES [s]	G1 [s]	G2 [s]	G4 [s]
4	0.81	0.35	0.11	0.05
16.1	2.99	1.23	0.40	0.18
38.7	7.24	2.90	0.92	0.41
62.3	11.69	4.60	1.49	0.68
121.3	22.85	9.03	2.85	1.25
174.2	33.60	13.00	4.08	1.82

Figure 2.2. DES vs high directed graph based algorithm, 64 bit password

2.7.4. Linearity from password length

It is easy to understand that with the fixed size of the plaintext, our algorithm depends linear from the password length. Each step of algorithm (taking the neighbour of the chosen colour) have fixed complexity, and the number of such steps depends on number of characters in the password.



Pass [B]	G1 [s]	G2 [s]	G4 [s]
1	0.38	0.20	0.14
2	0.75	0.39	0.14
4	1.50	0.58	0.28
6	2.24	0.77	0.28
8	2.95	0.92	0.42
10	3.69	1.16	0.42
12	4.46	1.37	0.56
14	5.12	1.52	0.56
16	5.79	1.68	0.68

Figure 2.3. Fixed file size (40 MB), comparison of our 3 algorithms.

Figure (2.3) illustrates this property, and shows the advantage of using bigger alphabet, but less number of operations. Algorithm "G4", using natural for today's computers, 32 bit arithmetics (with automatic modulo operations) behaves over 8 times faster that "G1" (8 bit arithmetics).

2.8. Statistics related to mixing properties

In our cryptographical scheme different passwords produce different ciphertexts with fixed plaintext. From the other hand when we fix the password, different plaintexts produce different cipertexts. Good cryptographical systems should ensure this difference to be big in terms of number of characters changed, looking as "randomly" as possible. These demands are known in literature as *Madryga requirements*. There are more postulates for

a good crypto system formulated by Madryga, but here we will concentrate on the mentioned two.

RC4 algorithm, as most elder stream ciphers, have the property, that with fixed password, changing one element of the plaintext leads to change one corresponding element in ciphertext. Such algorithms are not secure against the *plaintext-ciphertext* attacks.

Our basic algorithm, based on paths in graphs from the family $RDE(n, K)$, behaves similarly to RC4: changing one element of the plaintext leads to change only few elements in the ciphertext.

In order to correct this property, we can combine the algorithm with some fast, sparse matrix operations:

1. *Desynchronisation* of the graph by the automorphism.
   Let $\bar{a} = (a_1, a_2, \ldots, a_m), (a_i \in Z_{2^k})$ be the password and N_{a_i} be one step of algorithm (passing from one vertex to another using a_i element of password). We can denote our encryption algorithm as

$$E_{\bar{a}} = N_{a_1} N_{a_2} \ldots N_{a_m}.$$

   Desynchronisation can be described as:

$$\mathbf{A} N_{a_1} N_{a_2} \ldots N_{a_m} \mathbf{A}^{-1} = \mathbf{A} N_{a_1} \mathbf{A}^{-1} \mathbf{A} N_{a_2} \mathbf{A}^{-1} \ldots \mathbf{A} N_{a_m} \mathbf{A}^{-1},$$

   where $\mathbf{A}$ is some bijection. All interesting from our point properties of $E_{\bar{a}}$ are preserved.
2. *Deformation* of the graph.
   With the above notation for the deformation we use two bijections $\mathbf{A}$ and $\mathbf{B}$, changing $E_{\bar{a}}$ into $\mathbf{A} E_{\bar{a}} \mathbf{B}$. The property that different passwords lead to different ciphertexts is preserved, but there can happen the situation, that for the plaintext vector $\bar{x}$ the corresponding ciphertext, $\mathbf{A} E_{\bar{a}} \mathbf{B}(\bar{x})$ coincides with $\bar{x}$. Anyway the probability of such event is $1/|V|$, where V is the plainspace. It is very close to zero.

File [MB]	G1 [s]	G2 [s]	G4 [s]
4	0.04	0.02	0.01
16.1	0.12	0.10	0.08
38.7	0.32	0.24	0.20
62.3	0.50	0.40	0.30
121.3	0.96	0.76	0.60
174.2	1.39	0.96	0.74

Table 2.2. Time grow from mixing property $\mathbf{A}_n E_{\bar{a}} \mathbf{A}_n^{-1}$ for chosen operator $\mathbf{A}_n$

We chose the bijection A as sparse affine transformation. Its complexity is $O(n)$. Our test shows, that using for desynchronisation a properly chosen upper-triangular matrix $\mathbf{A}_n$ gives about 98.5% difference between the ciphertexts, when changing only 1 element of the plaintext (we use index n, because size of the $\mathbf{A}$ depends on the size of the plaintext). Table (2.2) shows the extra time spent by all 3 versions of our algorithm on the operation $\mathbf{A}_n$.

If instead of desynchronisation as above we apply the deformation with $\mathbf{B} = \mathbf{I}$ (identity map) and same $\mathbf{A}$, the speed of computation will be twice better and mixing properties are same.

The second Madryga requirement mentioned above (effect of the change of one character from the key) can be stated as follows: for short passwords (1B) the percentage of the change within the cipherstring is about 92%, and for longer passwords up to 96%.

2.8.1. On the time evaluation for the public rule

Recall, that we combine graph transformation N^l with two affine transformation T_1 and T_2. Alice can use $T_1 N^l T_2$ for the construction of the following public map of

$$y = (F_1(x_1, \ldots, x_n), \ldots, F_n(x_1, \ldots, x_n))$$

$F_i(x_1, \ldots, x_n)$ are polynomials of n variables written as the sums of monomials of kind $x_{i+1} \ldots x_{i_3}$, where $i_1, i_2, i_3 \in 1, 2, \ldots, n_1$ with the coefficients from $K = F_q$. As we mention before the polynomial equations $y_i = F_i(x_1, x_2, \ldots, x_n)$, which are made public, have degree 3. Hence the process of encryption and decryption can be done in polynomial time $O(n^4)$ (in one y_i, $i = 1, 2 \ldots, n$ there are $2(n^3 - 1)$ additions and multiplications). But the cryptoanalyst Cezar, having only a formula for y, has very hard task to solve the system of n equations of n variables of degree 3. It is solvable in exponential time $O(3^{n^4})$ by general algorithm based on Gröbner basis method. Anyway studies of specific features of our polynomials could lead to effective cryptoanalysis. This is an open problem for specialists.

We have written a program for generating public key and for encrypting text using generated public key [128]. The program is written in C++ and compiled with the Borland bcc 5.5.1 compiler.

We use matrix in which all diagonal elements equal 1, elements in the first row are non-zero and all other elements are zero as A, identity matrix as B and null vectors as c and d. In such a case the cost of executing affine transformations is linear.

The following table presents the time (in milliseconds) of generation of public key depending on the number of variables (n) and the password length (p).

	$p=10$	$p=20$	$p=30$	$p=40$	$p=50$	$p=60$
$n=10$	15	15	16	32	31	32
$n=20$	109	250	391	531	687	843
$n=30$	609	1484	2468	3406	4469	5610
$n=40$	2219	7391	12828	18219	24484	29625
$n=50$	5500	17874	34078	49952	66749	82328
$n=60$	12203	42625	87922	138906	192843	242734
$n=70$	22734	81453	169250	286188	405500	536641
$n=80$	46015	165875	350641	619921	911781	1202375
$n=90$	92125	332641	708859	1262938	1894657	2525360
$n=100$	159250	587282	1282610	2220610	3505532	4899657

The following table presents the time (in milliseconds) of encryption process depending on the number of bytes in plaintext (n) and the number of bytes in a character (w).

	$w=1 \quad (Z_{2^8})$	$w=2 \quad (Z_{2^{16}})$	$w=4 \quad (Z_{2^{32}})$
$n=20$	16	0	0
$n=40$	265	47	15
$n=60$	1375	188	15
$n=80$	3985	578	47
$n=100$	10078	1360	125

2.8.2. Theoretical properties, some other folders

1) The relation of the $RDE(n,K)$ - based algorithm with arithmetical dynamical system insure that in case when $l \leq n/2+2$ and fixed T_1, T_2 different passwords of length l map the plaintext to distinct ciphertext. If $T_1 = {T_2}^{-1}$ and $l \leq n+3$, then cubical encryption map has no fixed points.
2) Let us consider $DE(n,K)$-based encryption, where N_α is an operator of taking the neighbour alongside the edge of color $\alpha \neq 0$ of directed edge in $DE(n,K)$. We will prove that if $\mathrm{char}(K) \neq 2$ then the connected component containing flag $((0),[0])$ is defined by additional equations $a_1(f) = a_2(f) = \ldots a_r(f)$. All connected components are isomorphic.

Let $CDE(n,K)$ be the connected component with $|P| = |L| = |K|^{n+1-r}$ points and lines. Then $CDE(n,K)$ encryption can convert any chosen plaintext into any chosen ciphertext.

Polarities of $D(n,K)$ and related dynamical systems

Let P and L be disjoint sets, the elements of which we call *points* and *lines*, respectively. A subset I of $P \times L$ is called an *incidence relation* on

the pair (P, L). The *incidence graph* Γ of geometry (P, L, I) is defined to be the bipartite graph with vertex set $P \cup L$ and edge set $\{\{p, l\} | p \in P, l \in L, (p, l) \in L\}$.

Let $\pi : P \cup L \to P \cup L$ be a bijection for which the following hold

(i) $P^\pi = L$ and $L^\pi = P$,
(ii) for all $p \in P$, $l \in L$ $(l^\pi, p^\pi) \in I$ if and only if $(p, l) \in I$,
(iii) $\pi^2 = 1$.

We call such π a polarity of the incidence structure (P, L, I). Note that π induces an order two automorphism of the incidence graph Γ which interchanges the bipartition sets P and L. We shall use the term "polarity" and the notation "π" for the graph automorphism as well.

We now define the polarity graph Γ^π of the structure (P, L, I) with respect to polarity π. It is the graph with the vertex set $V(\Gamma^\pi) = P$ and edge set

$$E(\Gamma^\pi) = \{\{p_1, p_2\} | p_1, p_2 \in P, p_1 \neq p_2, (p_1, {p_2}^\pi) \in I\}.$$

Finally, we call point $p \in P$ an *absolute point* of the polarity π provided (p, p^π) in I.

Let N_π denote the number of absolute points of π.

Proposition 7. *(see, for instance [64])*

Let π be be a polarity of the finite incidence structure (P, L, I) and let Γ and Γ^π be the correspondent incidence and polarity graphs.

(a) $\deg_{\Gamma^\pi} = \deg_{\Gamma^\pi} - 1$ *if p is an absolute point of π, and* $\deg_{\Gamma^\pi} = \deg_\Gamma$ *otherwise.*
(b) $|V(\Gamma^\pi)| = 1/2|V(\Gamma)|$, $|E(\Gamma^\pi)| = |E(\Gamma)| - N_\pi$,
(c) *If Γ^π contains a $(2k+1)$-cycle then Γ contains a $(4k+2)$ cycle.*
(d) *If Γ^π contains a $2k$-cycle then Γ contains two vertex disjoint $2k$ cycles C and C' such that $C^\pi = C'$. Consequently, if Γ is $2k$-cycle-free then so is Γ^π.*
(e) *The girth of the two graphs are related by $g(\Gamma^\pi) \geq 1/2 g(\Gamma)$.*

It is clear that statements (c) , (d) and (e) are valid for an infinite incidence structure with polarities.

Let us consider the case of the incidence structure with paralleloopic graph (Γ, ρ) with the polarity π which is the parallelotopic morphism. We call such π a *parallelotopic polarity*. In that case we can define the *regular folding graph* $\mathrm{R}\Gamma = \mathrm{R}(\Gamma^\pi) = \{(p.p') | \rho(p) \neq \rho(p'), (p, p') \in E(\Gamma^\pi)\}$.

Let us consider the case when the set B of colours of the absolute points is a proper subset of the set of all colours C. In that case we can define an induced subgraph $\mathrm{I}\Gamma = \mathrm{I}\Gamma^\pi$ with the set of vertices $\{v \in \Gamma^\pi | \rho(v) \in$

$C - B\}$ Directly from the definitions and above proposition we are getting the following statement.

Lemma 5. *Let P, L, I be the incidence structure with the k-regular parallelotopic incidence graph Γ and parallelotopic polarity $\pi : \Gamma \to C$. Then $\mathrm{R}(\Gamma^\pi)$ is $k-1$-regular graph of girth g, where $g \geq g(\Gamma^\pi) \geq g(\Gamma)$.*

If the set B of colours for absolute points of π is different from C, then $\mathrm{I}\Gamma$ is $|C - B|$-regular graph and $g(\mathrm{I}\Gamma)) \geq g(\Gamma^\pi) \geq g(\Gamma)$.

Remark 1. Graph $\mathrm{I}\Gamma$ is a parallelotopic graph. Let S be a finite proper subset of $C - B$ of cardinality s. Then the graph $\mathrm{I}\Gamma^S$ has valency s and

$$g(\mathrm{I}\Gamma^S) \geq g(\mathrm{I}\Gamma).$$

Remark 2. Graph $R\Gamma$ is not a parallelotopic graph because of sets of colours from the neighbour hoods differs from vertex to vertex. Let S, $|S| = s$ be a subset of the colour set C of the parallelotopic graph Γ. Then parallelotopic polarity π induces a parallelotopic polarity π of $R\Gamma^S$. The graph $\mathrm{R}\Gamma^S$ shall be a graph of valency $s-1$ and $g(\mathrm{R}\Gamma^S) \geq g(\Gamma^S) \geq g(\Gamma)$.

Proposition 8. *The map π given by the close formula*

$$\mathrm{p}^\pi = [p_{10}, -p_{11}, p_{21}, p_{12}, -p'_{22}, -p_{22}, \dots, -p'_{ii}, -p_{ii}, p_{i+1,i}, p_{i,i+1}, \dots],$$
$$\mathrm{l}^\pi = (l_{01}, -l_{11}, l_{21}, l_{12}, -l'_{22}, -l_{22}, \dots, -l'_{ii}, -l_{ii}, l_{i+1,i}, l_{i,i+1}, \cdots)$$

is a parallelotopic polarity of $D(n, K)$. It preserves blocks of the equivalence relation τ. It is restriction on $V(\mathrm{CD}(n, K))$ is a parallelotopic polarity of $\mathrm{CD}(n, K)$.

Let $L(n, K)$ be regular folding graph corresponding to the parallelotopic polarity π induced on the vertices of the graph $C(n, K)$. In case of $\mathrm{char}K = 2$ the colours of absolute points of the polarity graph of $C(n, K)$ corresponding to the polarity π form the set $B = \{x | x^2 = 0\}$. Thus colours of the vertices of $B(n, K)$ are elements of $K - B$.

Directly from the fact $g(D(n, F_q)) \geq 2[(n+5)/2]$, proposition 7 and lemma 5 we are getting

Proposition 9.

(i) The girth of the graph $L(n, F_q) = L(n, q)$ and $B(n, F_q) = B(n, q)$, q is even is, at least $2[(n+5)/2]$. They are regular graphs of degrees $q-1$ and q^t with q^t and $(q-1)q^{t-1}$ vertices, respectively.

(ii) For each q they form a families of graphs of large girth with the $\gamma = 2/3\log_{q-1}(q)$.

(iii) Let S be a subset of nonzero elements of F_q, $|S| = s$ then $L(n, F_q)^S$ and $B(n, F_q)^S$ (q is even) are graphs of the order sq^{t-1}, girth $\geq 2[(n + 5)/2]$ and degrees $s - 1$ and s, respectively.

Arithmetical dynamical systems defined by projective limits of $L(n, K)$ and $B(n, K)$ were considered in [118]. The implementation of related stream ciphers the reader can find in [54].

Chapter 3

Groups and geometries as source of graphs with special walks

3.1. Incidence systems and groups

Let us recall some standard definitions.

An *incidence system* over a type set Δ is a triple (Γ, I, t), where Γ is a set (whose elements are called *objects*), I is a symmetric and reflexive binary relation on Γ (called the *incidence relation*) and t is a map from Γ into Δ (called the *type function*). The *rank* of the incidence system is defined to be $|\Delta|$. It is convenient to write Γ in place of (Γ, I, t) when doing so does not lead to confusion. Let Γ and Γ' be incidence systems defined over the same type set Δ. A *morphism* of Γ into Γ' is a map $\phi : \Gamma \to \Gamma'$ which preserves incidence. We say ϕ is *type–preserving* if, in addition, $t(A) = t(A^\phi)$ for all $A \in \Gamma$.

An important example of the above is the so–called *group incidence system* $\Gamma(G, G_s)_{s\in S}$. Here G is an abstract group and $\{G_s\}_{s\in S}$ is a family of distinct subgroups of G. The objects of $\Gamma(G, G_s)_{s\in S}$ are the cosets of G_s in G for all possible $s \in S$. Cosets α and β are incident precisely when $\alpha \cap \beta \neq \emptyset$. The type function is defined by $t(\alpha) = s$ where $\alpha = xG_s$ for some $x \in G$.

The rank of the incidence system (Γ, I, t) is the cardinality of Δ.

An *incidence structure* (P, L, I) is an incidence system of rank 2. In this case the set Γ is $P \cup L$, where P and L are two disjoint sets (the set of points and the set of lines, respectively). As usual, we impose the following restrictions on the incidence relation I: two points (lines) are incident if and only if they coincide.

The graph $B((P, L, I))$ of the symmetrical incidence relation I referred to as the incidence graph for (P, L, I). We will not make much distinction between the incidence structure and the corresponding *incidence graph*. We will say that an incidence structure is r_1, r_2-*biregular*, if every point is incident to r_1 lines and every line is incident to r_2 points.

Let (P, L, I) be an incidence structure, P' and L' subsets of P and L respectively and I' the restriction of the relation I to the set $P \cup L$. We shall refer to the incidence structure (P', L', I') as a *substructure* of (P, L, I). Obviously, $B((P', L', I'))$ is the induced subgraph in (P, L, I).

Let G be a group with proper subgroups G_1 and G_2 such that $< G_1, G_2 >= G$. Let us consider the incidence structure $\Gamma(G) = \Gamma(G)_{G_1,G_2}$ with a set of points $P = (G : G_1)$ and set of lines $L = (G : G_2)$.

The group G is a subgroup of the automorphism group of $\Gamma(G)$, the action of G on the set of edges being equivalent to its action on $(G : G_1 \cap G_2)$ by right shifting.

The following elementary statement is well known.

Lemma 6. *The incidence graph for* $\Gamma = \Gamma(G)_{G_1,G_2}$ *is connected if and only if the subgroups* G_1 *and* G_2 *generate* G*. Every connected component of* Γ *is isomorphic to* $\Gamma(G')_{G_1,G_2}$ *where* $G' =< G_1, G_2 >$.

We will say that a bipartite parallelotopic graph is a *group parallelotopic graph* if it is isomorphic to the incidence graph of an incidence structure $\Gamma(G)_{G_1,G_2}$ where G is a group with subgroups G_1 and G_2.

We will consider further the case of a *unipotent-like factorization*, i.e. a factorization of a group U into 3 subgroups U_1, U_2 and U_3 such that $U_1 \cap U_2 = 1$, $U_1 \cap U_3 = 1$, $U_2 \cap U_3 = 1$, and U_3 contains $[U_1, U_2]$. Thus, there are unique decompositions $u \in U$ of the kinds $u = u_1u_2u_3$ and $u = u_2u_1u'_3$ where $u_1 \in U_1$, $u_2 \in U_2$, $u_3, u'_3 \in U_3$.

Let us consider the incidence structure $\Gamma = \Gamma(U)_{U_1,U_2}$. Directly from definitions we obtain:

(1) For every coset U_1u there is a canonical representative u_2u_3, $u_2 \in U_2$, $u_3 \in U_3$. Let us call $u_2 = \pi(U_1u)$ the colour of the coset U_1u.
(2) For every coset U_2u' there is a canonical representative $u_1u'_3$, $u_1 \in U_1$, $u'_3 \in U_3$. Let us call $u_1 = \pi(U_2u')$ the colour of the coset U_2u'.

Lemma 7. *Let* $U = U_1U_2U_3$ *be a unipotent-like factorization of the group* U*. Then the incidence graph of* $\Gamma(U) = \Gamma(U)_{U_1,U_2}$ *is a group parallelotopic graph with color set* $U_1 \cup U_2$ *and with parallelotopic colouring* π*.*

Proof. Without loss of generality we can consider only the case of neighbor U_2u' of coset U_1u. Let u_1 be the color of U_2u' and let u_2 be the color of U_1u. Let g be common element of cosets U_1u and U_2u'. Then $g = b_2u_1u_3 = b_1u_2u'_3$ for some $b_1 \in U_1$ and $b_2 \in U_2$ and $u_3, u'_3 \in U$. We can rewrite this equation in the following form $u_1b_2wu_3 = b_1u_2u'_3$, where $w = [b_2, u_1] \in U$. From the uniqueness of the decomposition of g into product of elements from U_1, U_2 and U_3 we obtain $b_i = u_i$, $i = 1, 2$ and $wu_3 = u'_3$.

Thus, $u_3 = [u_1, u_2]u'_3$ and neighbor has been determined uniquely.

□

Remark. Graph $\Gamma(U)_{U_1,U_2}$ is a biregular with bidegrees $a = |U_1|$ and $b = |U_2|$ and $r = |U_3|$.

The following statement is useful for making parallelotopic quotients of group parallelotopic graphs.

Lemma 8. *Let* $U = U_1U_2U_3$ *be a unipotent like factorization of group* U *and let* F *be a normal subgroup of* U*, which is a proper subgroup in* U_3*. Let* ϕ *be the canonical homomorphism of* U *onto* U/F*. Then*

$$\phi(U) = \phi(U_1)\phi(U_2)\phi(U_3)$$

is a unipotent like factorization of $\phi(U)$.

Let Γ' be a parallelotopic quotient of a group parallelotopic graph Γ. If Γ' is also a group parallelotopic graph we call it a group parallelotopic quotient of Γ.

Reference for rest of section is [68].

Let us recall the definition of free product and some of its basic properties.

Let $A =< a_1, \cdots, a_n | R_1, \ldots, R_d >$ and $B =< b_1, \cdots, b_m | S_1, \ldots S_t >$ be subgroups with generators a_i and b_k and generic relations R_j and S_l, respectively. Free product $F = A * B$ of A and B is the subgroup $< a_1, \cdots, a_n, b_1, \cdots b_m | R_1, \cdots, R_d, S_1, \cdots, S_t >$. We will treat A and B as a subgroups of F in a usual way.

Call the nonidentity elements of A and B syllables. A syllabic word is simply a product of the form $a_1 b_1 a_2 b_2 \cdots$ or $b_1 a_1 b_2 a_2 \cdots$ with $a_i \in A$, $b_j \in B$ syllables. The length of a syllabic word w is merely its number of syllables, and the syllabic length of w is the smallest length of any syllabic word equivalent to w. We define $\lambda(g)$ to be the syllabic length of g.

Proposition 10. *Let G be a free product of finite nontrivial groups G_1 and G_2. Let G_3 be the group $[G_1, G_2]$. Then $G = G_1 G_2 G_3$ is a unipotent-like factorization of G.*

Proof. Consider the case $g = g_1 h_1 g_2 h_2$. Then

$$g_1 h_1 g_2 h_2 = g_1 h_1 h_2 g_2 [g_2, h_2] = g_1 h g_2 [g_2, h_2],$$

where $h \in G_2$. This equals $g_1 g_2 h [h, g_2][g_2, h_2] = g h [h, g_2][g_2, h_2]$ where $g \in G_1$.

Finally this equals $hg[g, h][h, g_2][g_2, h_2]$ which is a word in $G_1 G_2 G_3$. For longer words this trick can be repeated inductively.

Let us prove that the decomposition of g above is unique. Suppose that $g = v_1 v_2 v_3$ and $g = u_1 u_2 u_3$ are two decompositions of g. Then $v_1 v_2 v_3 = u_1 u_2 u_3$ and $(u_1^{-1} v_1) v_2 (v_3 u_3)^{-1} = u_2$. It is clear that $\lambda(u_2) \leq 1$. If $u_3 \neq v_3$ the syllabic length of the left-hand-side of the last relation is ≥ 2 because the nontrivial elements of G_3 have syllabic length at least 4 (since elements $[a, b] = aba^{-1}b^{-1}$ are free generators of $[A, B]$). Thus $v_3 = u_3$ and we immediately have $u_2 = v_2$ and $v_1 = u_1$.

□

A *filtration* of group G is an infinite decreasing sequence F_n, $n = 1, \ldots$ of distinct normal subgroups of G such that

(*i*) $F_1 = G$,
(*ii*) the commutator $[F_n, F_m] < F_{n+m}$

An important fact for us is that the free product $G = A * B$, with A and B finite groups, has infinitely many filtrations F_n such that $|G : F_n| \leq \infty$ for all n.

Lemma 6 provides a tool for making group parallelotopic quotients $\Gamma(U/F)_{U_1,U_2}$. Namely, we have

Theorem 6. *Let G_i, $i = 1, 2$ be a finite group, $G = G_1 * G_2$ and F_i a filtration with $|G : F_i| \leq \infty$ for all i. Then the graphs $\Gamma_i = \Gamma(G/F_i)_{G_1,G_2}$ form an infinite sequence of parallelotopic graphs of increasing order and unbounded girth. Moreover each mapping $\Gamma_i \to \Gamma_{i+1}$ is a group parallelotopic morphism.*

Proof. The group G is the projective limit of factor groups G/F_i (see [87]) and the graph $\Gamma(G)_{G_1,G_2}$ is the projective limit of $\Gamma(G/F_i)_{G_1,G_2}$. The graph $\Gamma(G)_{G_1,G_2}$ is a tree because G is a free product of G_1 and G_2 and its girth is infinite. Thus the girth of $\Gamma(G/F_i)$ is unbounded.

□

Proposition 11. *Let G be a simple group of Lie type of rank 2 over the finite field of characteristic p , U a Sylow p-subgroup of G , and U_1 and U_2 root subgroups corresponding to positive simple roots.*

Then there exist elementary abelian subgroups U_1, U_2 of U such that $\Gamma(U)_{U_1,U_2}$ is a group parallelotopic graph. Moreover, the girth of $\Gamma(U)_{U_1,U_2}$ is at least $2m$, where $m = 3, 4, 6$, or 8 depending on the Weyl group D_m of G, $G \in \{A_2(q), B_2(q), G_2(q),^2 A_3(q),^2 A_4(q),^3 D_4(q),^2 F_4(q)\}$.

Proof. Let U_1 and U_2 be root subgroups of U (see [22]) corresponding to fundamental roots. Then $U =< U_1, U_2 >$ and U_1, U_2 are elementary abelian subgroups with $U_1 \cap U_2 = e$. Let us consider the free product $H = U_1 * U_2$ and the uniquely defined surjective homomorhism map $\phi : H \to U$ such that $\phi(v) = v$ for $v \in U_i$, $i = 1, 2$. It is clear, that $\mathrm{Ker}\phi \cap U_i = e$. Thus $H/Ker\phi = U$ and we have unipotent like factorizations $U_1U_2[U_1, U_2]$ for both U and H. (Note that $[U_1, U_2]$ has different meanings inside U and H. If we set $U_3 = [U_1, U_2]$ (a subgroup of U) and $U_3' = [U_1, U_2]$ (a subgroup of H), then $U_3 = U_3'/F$, where $F = \mathrm{Ker}\phi$. Thus $\Gamma(U)_{U_1,U_2}$ is a parallelotopic graph.

Let r_1 and r_2 be fundamental roots corresponding to U_1 and U_2, respectively. Then the root subgroups U_{-r_1}, U_{-r_2} generate $U^- =< U_{-r_1}U_{-r_2} >$ and U_{-r_1}, U_{-r_2} and U^- are isomorphic to U_1 , U_2 and U, respectively. Let P_1 and P_2 are maximal parabolic subgroups containing U^-, i.e, maximal proper subgroups of G containing the normalizer of U^- in G. Then we have $U \cap G = U$ and $U \cap P_i = U_i$, $i = 1, 2$. Thus $\Gamma(U)_{U_1,U_2}$ is an induced subgraph of $\Gamma(G)_{P_1,P_2}$ which is the incidence graph of a generalized m-gon corresponding to G (see [20]). The girth of the generalized m- gon is $2m$.

The girth of the induced subgraph is at least the girth of the initial graph. Thus the girth of $\Gamma(U)_{U_1,U_2}$ is at least $2m$.

□

Remark. The graph $\Gamma(U)_{U_1,U_2}$ is called an affine m-gon. Generalized m-gons and their use in Cryptography will be considered in the next unit independently.

Let us consider the characterizations of $\Gamma(U)_{U_1,U_2}$ in terms of linear algebra.

Let G be a finite simple group of Lie type of rank 2 defined over a field of characteristic p, and U a Sylow p-subgroup of G. Then the normalizer B of U can be written as a semidirect product $U\lambda T$, where T is the maximal torus. There are exactly two maximal parabolic subgroups P_1 and P_2 of G, i.e. maximal proper subgroups containing the group B. The geometry $\Gamma(G)$ of G is the tactical configuration with the set of points $P = (G : P_1)$ and the set of lines $L = (G : P_2)$, a point p and line l are incident pIl if the set-theoretical intersection of the cosets a and b is non empty.

The Weyl group W of G is the dihedral group D_m, i.e. Coxeter group with the generators r_1 and r_2 and generic relations $(r_1)^2 = e$, $(r_2)^2 = e$ and $(r_1r_2)^m = e$, where m can be 3, 4, 6, or 8 depending on W. Thus we have factorizations $G = BWB$, $P = B < s_i > B$, $i = 1, 2$. The geometry $\Gamma(W)$ of W is the tactical configuration with the set of points $(W :< r_1 >)$ and $(W :< r_2 >)$, the point p is incident to the line l if and only if the intersection of cosets p and l is not empty. It is clear that $\Gamma(W)$ can be identified with the set of vertices (points) and edges (lines) of ordinary m-gon on the plane with natural incidences (as drawn).

The set of invariant points of the transformation group $(T, \Gamma(G))$ with the restriction of the incidence relation to it is the Weyl geometry $\Gamma(W)$. Thus we have the standard embedding of $\Gamma(W)$ into $\Gamma(G)$.

Let us consider orbits of B acting on $\Gamma(G)$ (so called Schubert cells). Every Schubert cell S contains a unique element α of $\Gamma(W)$, so we can use α as natural index of $S = S_\alpha$. We write $Cox(a) = \alpha$ if the element a of $\Gamma(G)$ is a representative of the Schubert cell S_α and say that $\mathrm{Cox}(a)$ is a *Coxeter trace* of a. Actually, the map $a \to \mathrm{Cox}(a)$ is a homomorphism of tactical configurations $\Gamma(G)$ onto $\Gamma(W)$, so called retraction map.

For each α in $\Gamma(W)$ there is a unique subgroup $U(\alpha)$ of U which acts regularly on S_α. Thus we have a natural matching between elements of $U(\alpha)$ and S_α. Thus any element a of $\Gamma(G)$ can be identified with the pair (α, x), where $\alpha = \mathrm{Cox}(a)$ and x is the element of $U(\alpha)$ corresponding to $a \in S_\alpha$. We will say that (α, x) are the *group coordinates* of a and write simply $a = (\alpha, x)$.

We follow the interpretation of $\Gamma(G)$ described in [105]. It is shown there that there exists a skew-symmetric linear algebra $(L, *)$ over the finite field K and a bijective map $\log : U \to L$ such that $\log U(\alpha)$ is a subspace $L(\alpha)$ of the vector space L and such that elements $a = (\alpha, x)$ and $b = (\beta, y)$ are incident in $\Gamma(G)$ if and only if α is incident with β as elements of $\Gamma(W)$ and $d_\alpha(\log x) - d_\beta(\log y) = \log x * \log y$, where d_α and d_β are certain linear operators on L, depending on α and β only. Neighbours of an element (α, x_0) of the kind (β, y) can be defined by system of linear equations. There is a convenient *canonical* basis in which all $L(\alpha)$ are spanned by basis elements and neighbours of an element (α, x_0) of the kind (β, y) are given by a system of linear equations with its matrix in row-echelon form.

If G is defined over $GF(q)$ of sufficiently large characteristic, then $(L, *)$ can be identified with the upper root space L^+ of the Borel subalgebra $H + L^+$ in the Lie algebra corresponding to G where H is the Cartan subalgebra. Multiplication in this Lie algebra is called the Lie product and denoted by [,].

Elements d_α can be treated as vectors in H acting on L^+. Thus an element $a = (\alpha, x)$ can be identified with the vector $y = d_\alpha + \log x$ of the Borel subalgebra. We will refer to y as the algebraic coordinates of a. Such an embedding of a Lie geometry into the Borel subalgebra is considered in [108] (see also [105] in case of a Lie group normal type, and in [104] for twisted groups). The relation $\exp(\log x) = x$ explains our choice of notation for the map from U onto L, the word logarithm is a reasonable term for this map.

Now, let us consider the structure of the subgroups U and $U(\alpha)$ and operators d_α in more details.

We say that $G = G_1 G_2 \cdots G_t$, $t \geq 2$ is a standard factorization of group G if each element $g \in G$ can be presented uniquely as $g = g_1 g_2 \cdots g_t$, where $g_i \in G_i$, $i = 1, \ldots, t$.

Let us consider $R = R_1 \cup R_2$, $R_i = \{{r_i}^w | w \in W\}$, $i = 1, 2$, i.e. R is the set of elements of W which are conjugate to one of the two Coxeter generators r_1, r_2. If $|W| \neq 8$ there is a natural one to one correspondence between R and the set $Root^+$ of all positive roots in R^2 corresponding to W, because we may identify each element of R with the reflection map relative to the line orthogonal to a certain root in $Root^+$. Hence we identify R with $Root^+$. In what follows, for each element $r \in R$ we may consider a root subgroup U_r, which is isomorphic to the additive subgroup of some finite vector space V. Thus U_r is an elementary abelian group. We have $U_r = \{x_r(t) | t \in V\}$, where $x_r(t)$ is the so called root element, see [22].

Note that if G is a Lie group of normal type defined over the field F_q, then V is a one-dimensional vector space over F_q. In the case of a twisted

group the dimension of V can be larger; in fact one may have that U_r and $U_{r'}$ have different order for $r \in R_1$ and $r' \in R_2$.

There are standard factorizations $U = U_{r_1}U_{r_2}\dots U_{r_m}$ and $U = U_{r_2}U_{r_1}\dots U_{r_m}$, where $r_3, \dots, r_m$ is an ordering of the positive roots other than r_1, r_2.

For a group of normal type each element u of U can be written uniquely as $u = x_{r_1}(t_1)x_{r_2}(t_2)\dots x_{r_m}(t_m)$ and $\log(u) = (t_1, t_2, \dots, t_m)$ in the basis $e_{r_1}, \dots, e_{r_m}$ of L^+ which is the restriction of the Chevalley basis of the Lie algebra. This is the canonical basis mentioned earlier. Set $L_{r_i} = \log(U_{r_i})$. Then one has $L^+ = L_{r_1} + L_{r_2} + \dots + L_{r_m}$. This is called the *root decomposition* of L^+.

There are standard factorizations of group $U(\alpha)$ into subgroups U_{r_i}, $r_i \in R_\alpha$, where R_α is a certain subset of R. More precisely, we may identify the coset $\alpha =< r_i > w$, $w \in W$, $i = 1, 2$ with the unique word $a_1a_2\dots a_t$ in α of shortest length in the alphabet $\{r_1, r_2\}$ with the initial letter $a_1 \neq r_i$. Then $R_\alpha = \{a_t, (a_t)^{a_{t-1}}, \dots, (a_t)^{a_{t-1}a_{t-2}\dots a_1}\}$. Obviously L_α is the direct sum of L_r, $r \in R_\alpha$.

Finally, we may consider the following interpretation of $\Gamma(W)$ in the case of sufficiently large characteristic. Let h_1, h_2 be the basis of the Cartan subalgebra H which is the restriction of the Chevalley basis on it. L^+ is an invariant subspace for H under the Lie product and we may identify the linear operator $d_\alpha(x)$, $\alpha =< r_i > w$ with $[{h_i}^w, x]$ and the element α with $h_\alpha = {h_i}^w$. Thus ${h_i}^w + \log(u)$ is an element of the Borel subalgebra corresponding to the object from $\Gamma(G)$ with group coordinates (α, u). Elements $a = (\alpha, u)$ and $b = (\beta, v)$ are incident if $\alpha I \beta$ in the Weyl geometry and $[h_\alpha + \log u, h_\beta + \log v] = 0$.

Let us consider the action of U on $\Gamma(G) = P \cup L$. There are m orbits (Schubert cells) of the transformation group (U, P) ((U, L), respectively). Let us consider the largest Schubert cell on each $P = (G : P_1)$ and $L = (G : P_2)$. It contains the unique coset $\alpha(i)$ from $\Gamma(W)$ containing an element D_m of maximal length. For such an $\alpha(i)$ we have $h_{\alpha(i)} = -h_i$ and $R_{\alpha(i)} = R - \{r_i\}$, $i = 1, 2$. Let us consider the induced subgraph $A(\Gamma(G))$ of $\Gamma(G)$ on the Schubert cells containing $\alpha(1)$ and $\alpha(2)$. We may identify elements with their algebraic coordinates, which are actually vectors of the Borel subalgebra, and get the following description of $A(\Gamma(G))$:

$$P = (-h_1)^* + x, \quad x \in \sum_{r \in R - \{r_1\}} L_r,$$

$$L = (-h_2)^* + y, \quad y \in \sum_{r \in R - \{r_2\}} L_r$$

$$-{h_1}^* + xI - {h_2}^* + y \iff [-{h_1}^* + x, -{h_2}^* + y] = 0.$$

For a group of normal type we may write the incidence condition defined above by linear equations over the ground field K. To define a group G of twisted type we need to take a certain field automorphism v. We may still use the equations above to define the incidence of $\Gamma(G)$ (see [104]). They are not linear over K but they are linear over a certain subfield $K' \subset K$. In fact K' could be the invariant subfield of the field automorphism v if $m \neq 8$, and $K' = GF(2)$ for $m = 8$. Thus we proved the following statement.

Proposition 12. *Let G be a simple group of Lie type of rank 2, defined over a field K, U a unipotent subgroup of G, and U_1 and U_2 root subgroups corresponding to positive simple roots. Then the incidence graph $\Gamma = \Gamma(U)_{U_1,U_2}$ is a linguistic graph of affine type over some subfield K' of K.*

Remark. Examples 1-5 of optimal linguistic graphs of affine type are graphs $\Gamma(U)_{U_1,U_2}$. They are affine m-gons corresponding to groups the A_2, B_2, G_2 and ${}^2A_3$, respectively. Similarly we can write the linear equations, which give us ${}^3D_4$, ${}^2A_4$ and ${}^2F_4$. The case of F_4 is more sophisticated because the number of equations depends on the size of the ground field $GF(2^k)$, where k is odd.

Lemma 6 is a tool for making group parallelotopic quotients $\Gamma(G/F)_{U_1,U_2}$. It can be generalized as follows.

Let G be a group with subgroups, H, G_1, G_2, such that $H \cap G_1 =< e >$ and $H \cap G_2 =< e >$. Let $\Gamma(G \swarrow H)_{G_1,G_2}$ be the incidence graph of the incidence structure, whose points are double cosets G_1gH and lines are double cosets G_2gH and where a point and a line are incident iff their intersection, as subsets of G, is not empty. Let $\mu_H : \Gamma(G)_{G_1,G_2} \to \Gamma(G \swarrow H)$ be the map such that $\mu_H(G_ig) = G_igH$.

We say that $G = G_1G_2 \dots G_t$ is a *strong factorization* of the group G if the decomposition $g = g_1g_2 \dots g_n$, $g_i \in G_i$ is uniquely defined for any $g \in G$.

Lemma 9 (107)**.** *. Let $G = U_1U_2U_3$ be a unipotent-like factorization of the group G and H a proper subgroup of U_3. Then*

(i) $\Gamma(G \swarrow H)_{U_1,U_2}$ is a group parallelotopic structure and μ_H is a parallelotopic homomorphism of $\Gamma(G)_{U_1,U_2}$ onto $\Gamma(G \swarrow H)_{U_1,U_2}$.

(ii) If $U_3 = KH$ is a strong factorization of U_3, then μ_H is a triangular homomorphism.

Remark. The graphs $CD(k, q)$ are graphs of the triangular folder $\Gamma(G \swarrow H_i)_{G_1,G_2}$, $G = G_1 * G_2$, where G_1 and G_2 are two copies of the elementary abelian group $GF(q)^+$, $q = p^n$ and $|H_i : H_{i+1}| = q$ [107]. In particular, $CD(2k, q)$ are isomorphic to $\Gamma(G/H)_{G_1,G_2}$, where H_i is a normal subgroup of G.

We will consider also the more general case of a *parabolic-like factorization*, i.e. a factorization of group G into 4 subgroups U_1, U_2, U_3 and U_4 such that

a) $U_i \cap U_j = 1$ for $i \neq j$
b) U_2 and U_3 are finite
c) U_4 contains $[U_1, U_2]$ and
d) there are unique decompositions $g \in G$ of the kind

where $u_i \in U_i$, $i = 1, 2, 3, 4$.

Let us consider the incidence structure $\Gamma = \Gamma(G)_{U_{1,3},U_{2,3}}$, where

$$U_{1,3} =< U_1, U_3 >, U_{2,3} =< U_2, U_3 >$$

Directly from the definitions we find that

(1) For every point $(p) = gU_{1,3}$ there is a canonical representative $g' = u_2 u$, $u_2 \in U_2$, $u \in U_4$ such that $U_{1,3}g = U_{1,3}g'$. Let us call $u_2 = \pi((p))$ the "labelling of the point".
(2) For every line $[l] = U_2 g$ there is a canonical representative $g' = u_1 u$, $u_1 \in U_1$, $u \in U_4$ such that $U_{2,3}g = U_{2,3}g'$. Let us call $u_1 = \pi([l])$ the "labelling of the line".

Lemma 10. *Let $G = U_3U_1U_2U_4$ be the parabolic-like factorization of group G. Then the incidence structure $\Gamma(G) = \Gamma(G)_{U_{1,3},U_{2,3}}$ is a group parallelotopic structure over (U_1, U_2) with the labelling π.*

Proof. Without loos of generality we can consider only the case of a neighbour $[l]$ of the point (p). Let u_1 be the first coordinate of $[l] = U_{2,3}u_1x$ and $(p) = U_{1,3}u_2u_4$, $u_2 \in U_2$, $u \in U_4$. Let g be a common element of the cosets $[l]$ and (p). Then $g = b_2u_1x = b_1u_2u$ for some $b_1 \in U_1$ and $b_2 \in U_2$. We can rewrite this equation in the following form: $u_1b_2(wx) = b_1u_2u$, where $w = [b_2, u_1] \in U_4$. From the uniqueness of the decomposition of g into a product of elements from U_1, U_2 and U_4 we obtain

$$b_i = u_i,\ i = 1, 2 \text{ and } wx = u.$$

Thus, $x = [u_1, u_2]u$ and the neighbour has been determined uniquely. □

The proof of the following lemma is straightforward.

Lemma 11. *Let $G = U_3U_1U_2U_4$ be a parabolic-like factorization of group G and H be a proper subgroup of U_4. Then*

(i) $\Gamma(G \swarrow H)_{U_{1,3}U_{2,3}}$ is a group parallelotopic structure and μ_H is a parallelotopic homomorphism of $\Gamma(G)_{U_{1,3},U_{2,3}}$ onto $\Gamma(G \swarrow H)_{U_{1,3},U_{2,3}}$.

(ii) If $U_4 = KH$ is a factorization such that the decomposition $g = kh$, $k \in K$, $h \in H$ is uniquely defined, then μ_H is a triangular homomorphism.

(*iii*) *If H is normal subgroup of G and ϕ is a canonical homomorphism of G onto G/H, then μ_H is induced by ϕ and $\phi(U_3)\phi(U_1)\phi(U_2)\phi(U_4)$ is a parabolic-like factorization of G/H.*

Let U be a unipotent subgroup of a simple group G of Lie type. Consider two Schubert cells R_α and R_β, such that $\alpha I \beta$ in the Weyl geometry. Let U_α and U_β be the stabilizers of elements α and β in U, and $U_3 = U_\alpha \cap U_\beta$. Then there are a standard factorizations $U_\alpha = U_3U_1$, $U_\beta = U_3U_2$ and a parabolic-like factorization $U = U_3U_1U_2U_4$. It follows from interpreting of the geometry $\gamma(G)$ as a blow up of the Weyl geometry that the parallelotopic graph $\Gamma(U)_{U_\alpha,U_\beta}$ is a linguistic graph of affine type over some field K'. In particular, we have the following generalization of Proposition 9.

Proposition 13. *Let G be a simple group of Lie type, defined over field K, U a unipotent subgroup of G, and $U_1, \dots, U_n$ root subgroups corresponding to simple roots. Let J and I be subsets of $N = \{1, 2, \dots, n\}$ with $I \cup J = N$. Then*

(*i*) *There exist standard factorizations*

$$U_I = U(I, J)U(I),$$
$$U_J = U(I, J)U(J),$$

where $U(I, J) = U_I \cap U_J$, and the factorization

$$U = U(I, J)U_IU_JU'$$

is a parabolic-like factorization and the

(*ii*) *the incidence graph $\Gamma = \Gamma(U)_{U_I,U_J}$ is a linguistic graph of affine type over some subfield K' of K.*

Remark 1. Let U'' be a subgroup of U' in the last statement.

Then $\Gamma' = \Gamma(U//U'')_{U_I,U_J}$ is a parallelotopic quotient of Γ.

If $U' = U''U^1$ is a strong factorization for U'. Then Γ' is a triangular quotient of Γ.

Remark 2. Let L be the Lie algebra over the K corresponding to G, L' a subalgebra of L over the subfield K' and $U'' = \{exp(adx)|x \in K'\}$ a subgroup of U'. Then $\Gamma(U//U")_{U_I,U_J}$ is a linguistic graph of affine type over K'.

We will consider below examples of folders of graphs of unbounded girth connected with amalgamated products of finite groups. Let us recall the definition of this operation and some of its basic properties.

Definition [68] Let

$$G = < a_1, \ldots, a_n, b_1, \ldots, b_m; R(a_\nu), \ldots, R(b_\mu), \ldots, S(b_\nu), \ldots,$$
$$U_1(a_\nu) = V_1(b_\mu), \ldots, U_l(a_\nu) = V_l(b_\mu) > .$$

Let A' be the subgroup of G generated by $a_1, \ldots, a_n$,
B' the subgroup of G generated by $b_1, \ldots, b_m$,
H' the subgroup of A' generated by elements $U_1(a_\nu, \ldots, U_l(a_\mu)$, and
K' the subgroup of B' generated by $V_1(b_\mu), \ldots, V_l(b\mu)$
We may identify the isomorphic subgroups H' and K'. We then say that G is the free product of A' and B' with the amalgamated subgroup H'.

Remark. In the definitions above upper case letters with primes and G denote groups, other capital letters denote words or relations, and lower case denote generators.

Finally we formulate some straightforward generalizations of proposition 9 and 10.

Proposition 14. *Let G be a free product of finite groups A' and B' with amalgamation subgroup H' such that $A' = H'A"$ and $B' = H'B"$ are standard group factorizations. Then for some subgroup K the decomposition $G = H'A"B"K'$ is a parabolic-like factorization of G.*

Theorem 7. *Under the assumptions of the proposition above let us assume that D_i, $D_i > D_{i+1}$ is a family of subgroups of K' of the finite index containing almost all members of some filtration of G. Then the graphs $\Gamma_i = \Gamma(G//D_i)_{A',B'}$ of different order form an infinite folder of parallelotopic graphs of unbounded girth.*

If $D_i = D_{i+1}D_i'$ is a standard factorization of D_i, then Γ_i is a triangular folder. $\Gamma(G)_{G_1,G_2}$ is the projective limit of Γ_i defined by the canonical parallelotopic morphisms $\Gamma_i \to \Gamma_{i+1}$.

3.2. On graph theoretical absolutely secure encryption

Let us refer to a pair $(\Gamma, \{n_i | i \in J\})$, where J is the set of possible length for an encryption arc in the graph Γ, as a J-graphic scheme. Let us assume that the plaintext of bipartite graph will be one of partition sets. If graph is not a bipartite we will Is there a graph of girth g such that some J-graphic scheme defines the absolutely secure algorithm?

Examples of such graphs can give us idea, what objects are good tools for the encryption in practical situation, when size of the key is essentially

smaller than size of the plaintext. Besides absolutely secure algorithms can be used as blocks in real life encryption algorithms,

Let us consider a J-graphic scheme of the graph Γ og girth g where $\max\{j \in J\} \leq [g/2]$.

If Γ be an *absolutely secure scheme* then the ratio $p_{\text{key}}(i)/p_{\text{mes}}(i)$ of probabilities $p_{\text{key}}(i)$ and $p_{\text{mes}}(i)$ to guess the encoding sequence and to guess the message (plaintext) in the scheme (Γ, t), respectively, equals to 1, because we use a finite probabilistic space.

We will introduce below absolute secure algorithms with a certain resistance to an attack of type (ii), used walks on bipartite graphs which do not satisfy to the parallelotopic property. To mark the walks on these graphs we shall use the coloring of vertices in similar way with the parallelotopic case together with the special use of symbol ∞.

Lemma 12. *If a biregular incidence structure with bidegrees $r+1$ and $s+1$ and parameters p, l has girth $g \geq 2k+2$, then the following inequalities hold*

(1) *If $k = 2t+1$ then*

$$1 + r + rs + r^2s + r^2s^2 + \cdots + r^{t+1}s^t \leq p \tag{3.1}$$

$$1 + s + sr + s^2r + s^2r^2 + \cdots + s^{t+1}r^t \leq l \tag{3.2}$$

(2) *If $k = 2t$ then*

$$1 + r + rs + r^2s + r^2s^2 + \cdots + r^ts^t \leq p \tag{3.3}$$

$$1 + s + sr + s^2r + r^2s^2 + \cdots + s^tr^t \leq l \tag{3.4}$$

Proof. Let us consider chosen point P. The pass of length $h \leq k$ between two chosen vertices is unique. Thus counting of vertices at distance h can be done by branching process.

So we have $l_1 = r + 1$ lines at distance 1 from P, $p_1 = (r + 1)s$ is the number of points at distance 2 from $P \ldots$, $l_3 = (r + 1)rs$ is the number of points at distance 3 from P. Let $k = 2t + 1$. Then

$l_{2h+1} = (r + 1)r^hs^h$, $p_{2h+2} = (r + 1)r^hs^{h+1}$, where $h = 0, 1, \ldots, t$.

Obviously $l_1 + l_2 + \cdots + l_{2t+1} \leq l$ and this inequality is equivalent to (3.1).

If we change the points and lines in the computation above, we will get (3.2) by branching process starting from chosen line L.

In case of $k = 2t$ inequalities

$$p_0 + p_2 + \cdots + p_{2t} \leq p$$

$$l_0 + l_2 + \cdots + l_{2t} \leq l$$

are equivalent to (3.3) and (3.4).

□

If $t+1 = s+1 = k$ then the order of the graph is $v = 2p = 2l$ inequalities as above are equivalent to well known the Tutte's inequality for regular graphs.

$$v \geq 2(1 + (k-1) + \ldots (k-1)^{(g-2)/2} \tag{3.5}$$

Let us refer as *extaspecial* to incidence structures for which both inequalties of the lemma above turn out to be equalities. We will use term *extraspecial* for graphs of extraspecial structures and regular graphs (not necessarily bipartite) of order on Tutte's bound.

Proposition 15. *Let Γ be an extraspecial structure of girth $g = 2k$. Then graph encryption scheme (Γ, J), $J = \{0, \ldots, k-1\}$, defines absolutely secure algorithm.*

Proof. The plainspace here is P. The cipherspace is P or L depending on the value $k$mod2 the type of element (point or line). Thus for the decryption adversary has to try all passes of the length from $\{0, 1, \ldots, k-1\}$. The number of such passes is same with $V(\Gamma)$. More than that application of different passes produce different ciphertexts and all possible ciphertexts can be obtained by application all passes to the certain plaintext. Thus $p_{mes} = P_{key} = 1/|P|$.

□

Remark. If Γ is an extraspecial structure with hidden type of elements (point or line), then we may consider set $V(\Gamma)$ as the plainspace and take $S = \{0, 1, \ldots, k-1\}$. It gives us an example of absolutely secure algorithm of encryption. Below we consider some representations of extraspecial configurations with hidden type function.

It is important that the totality of extraspecial graps is nonempty. Generalized m-gons defined by J. Tits in 1959 (see [96], [97]) as a tactical configurations of bidegrees $s+1$ and $t+1$ of girth $2m$ and diameter m. The pair (s, t) is known as order of generalized m-gon. Extraspecial graphs of odd girth are known as Moore graphs [17].

From the definition one can deduce (see, for instance, [17]) that in case of generalized m-gon parameters p and l equal to lefthandsides of inequalities (3.1)-(3.4). Thus, generalized m-gons are extraspecial tactical configurations.

The following statement is well known (see [17])

Theorem 8. *A finite generalized n-gon of order (s, t) has $n \in \{3, 4, 6, 8, 12\}$ unless $s = t = 1$. If $s > 1$ and $t > 1$, then*

(1) $n \neq 12$;

(2) *if* $n = 4$, *then* $s \leq t^2$, $t \leq s^2$;
(3) *if* $n = 6$, *then* st *is a square and* $s \leq t^3$, $t \leq s^3$;
(4) *if* $n = 8$, *then* $2st$ *is a square and* $s \leq t^2$, $t \leq s^2$;

Apart of the inequalities, this is the original Feit-Higman theorem.

The known examples of generalized n-gons of of bidegrees ≥ 3 up to parameters are rank 2 incidence graphs of geometries of finite simple groups of Lie type. The regular incidence graphs are m=3 (group $A_2(q)$), $m = 4$ (group $B_2(q)$ or $C_2(q)$), $m = 6$ (group $G_2(q)$), in all cases $s = t = q$, where q is prime power.

The biregular but not regular generalized n-gons have parameters $s = q^\alpha$ and $t = q^\beta$, where q is some prime power. The list is below.
$n = 4$: $s = q, t = q^2$ and q is arbitrary prime power or $s = q^2, t = q^3$ and q is arbitrary prime power;
$n = 6$: $s = q^2$, $t = q^3$ and $q = 3^{2k+1}$, $k > 1$;
$n = 8$: $s = q$, $t = q^2$ and $q = 2^{2k+1}$.

It means that generalized $m-$gons related to simple Lie groups $G(GF(q))$ with chosen Dynkin diagramm over the finite field $GF(q)$, $q = p^n$, $n \geq 1$, p is prime, produce an infinite family of one-time pads. They have a certain resistance to attacks of type (ii). The best resistance given by the constant of complexity would be in the case of ${}^2F_4(q)$, $q = 2^{2k+1}$ - problem of finding the pass between 2 vertices of general position in generalized octagon is known hard problem of Algebraic Combinatorics.

The set of points (lines, respectively) of generalized m-gone can be considered as a disjoint union of vector spaces over the $GF(q)$, . It is convenient to treat elements of $GF(q)$ as tuples over the fixed alphabet $GF(p)$, so we may encrypt of "potentially infinite" text over $GF(p)$. We may consider say a real-life encryption schemes with flexible keys if we restrict our passes to the set of passes for the m-gone related to $G(GF(t))$ where $p \leq t \leq q$ is chosen power of p. We may vary the resistance $f(n)$ of such a scheme to attacks of type (i) (known ciphertext), it can be as close to one-time pad as we want, we may chose increasing $f(n)$, but the resistance to attack of type (ii) is bounded by some constant.

Thus we need families of increasing girth to construct graph theoretical schemes of encryption for the case of increasing resistance to attacks of type (ii).

Let us consider in details an example of the implementation of the encryption process for the generalized m-gons related to a simple group $G = G(q)$ defined over the finite field $GF(q)$ of char$GF(q) = p$.

Let U be the Sylow subgroup of G, i.e. so called unipotent subgroup of the standard Borel subgroup for G. This group can be factorized as $U = U_1U_2U' = U_2U_1U'$, where U_1, U_2 are so called root subgroups of U

of order $s = q$ and $t = q^h$, respectively, $s + 1$ and $p + 1$ are bidegrees of our bipartite graph. The orbits of (U, P), (U, L) are in one-to one corespondents with monomial terms in the expansion (1) for $|P|$, such that size of the orbit coincides with the numerical value of related monomial term $s^k t^j$. The action on each orbit of (U, P) can be identified with the regular representation of subgroup U_1U'', $U'' < U'$ of group U. Similarly each orbit of (U, L) can be identified with the subgroup U_2U'', where U'' depends on the chosen orbit. Thus each vertex of generalized m-gon can be identified with element gg'', $g \in U_i$, $i = 1, 2$, $g'' \in U''$. Let us consider the labbelling $c(gg') = g$, which is the "color" of the element, our spectra of colors is the disjoint union $U_1 \cup U_2$.

Let v be a point from the orbit for the term $s^j t^i$. It has exactly t-neighbors from the orbit (U, L) which is related to term $s^{j-1}t^{j+1}$. There is exactly one neighbor of chosen color there.

Remark. Besides generalized polygons related to simple groups of Lie type, which we consider above, there are important "nonclassical examples" of projective planes, non classical generalized quadrangles and hexagons ([17], [20]) and further references). In case of known examples , we can consider a similar interpretation of vertices and arcs: choose the edge $e = \{p, l\}$ of the graph and change the orbit of U on totality of points (lines) at given distance from e.

3.3. Correlation with expansion properties

Our applications of the Graph theory to Cryptography based on the use of graphs of high girth. Other cryptographic application use expansion properties of graphs, which is also important for parallel computations and some other area of Computer Science (see [66] and further references). One of the application expanding graphs to Image Processind the reader can find in [93].

In fact, there is an interesting correlation between these two properties:

(1) *All infinite families of k-regular graphs of given degree t, which have been considered above, are infinite families of expanders with the second largest eigenvalue bounded by constant* $2\sqrt{t}$.
(2) *All infinite families of graphs of unbounded degree k and bounded girth, which have been considered above are Ramanujan graphs, i.e. graphs with the second largest eigenvalue bounded by function* $2\sqrt{t-1}$.

Let us consider these facts in more details.

Let A be a set of vertices of a graph X. We define ∂A to be the set of all elements $b \in X - A$ such that b is adjacent to some $a \in A$.

We say that t-regular graph with n vertices has an expansion constant c if, for each set $A \subset X$ with $|A| \leq n/2$, $|\partial A| \geq c|A|$.

One says that the infinite family of graph X_i is a family of expanders constant c, if there exists a constant c such that every X_i has the expansion constant c.

An explicit construction of infinite families of t-regular expanders (t-fixed) turns out to be difficult.

Gregory Margulis ([69], [70], [71]) constructed the first family of expanders. He use representation theory of semisimple groups.

It can be shown that if $\lambda_1(X)$ is the second largest eigenvalue of the adjacency matrix of the graph X, then $c \geq (t - \lambda_1)/2t$. Thus, if λ_1 is small, the expansion constant is large. A well-known result of Alon and Bopanna says that, if X_n is an infinite family of t-regular graphs (t fixed), then $\lim \lambda_1(X_n) \geq 2\sqrt{t-1}$.

This statement was the motivation of Ramanujan graphs as special objects among t-regular graphs. A finite t-regular graph Y is called Ramanujan, if for every eigenvalue λ of Y, either $|\lambda| = t$ or $|\lambda| \leq 2\sqrt{t-1}$. So, Ramanujan graphs are, in some sense, best expanders. There is an interest to families of Ramanujan graph of unbounded degree too.

Lubotzky, Phillips and Sarnak ([66]) proved that graphs defined by Margulis are Ramanujan graphs of degree $p+1$ for all primes p. M. Morgenstern proved that, for each prime degree q, there exists a family of Ramanujan graphs of the degree $q - 1$.

Unipotent-like factorizations give us other source of infinite families of expanders. The following 2 statements have been formulated in [110].

Theorem 9. *Let G be a finite group, let G_1 and G_2 be isomorphic subgroups of G such that $G =< G_1, G_2 >$, $G = G_1G_2G'$ be a unipotent-like factorization, and set $T = |G_1|$. Let $\Gamma = \Gamma(G)_{G_1,G_2}$ has no cycles of length 4. Then the second largest eigenvalue of Γ is bounded by $2\sqrt{t}$.*

Theorem 10. *Let G_1, G_2 are two copies of a finite group G of order $|t|$. Then the free product $F = G_1 * G_2$ contains infinitely many normal subgroups H of finite index, such that graphs $\Gamma(F/H)_{G_1,G_2}$ form an infinite family of expanders with embedded spectra for which the second largest eigenvalue is bounded by $2\sqrt{t}$.*

As it has been mentioned before in the remark after Lemma 8 graphs $CD(2k, q)$ coincide with $\Gamma(G/H)_{G_1,G_2}$, where $G_1 * G_2$ for $G_1 = G_2 = {F_q}^+$. So we may apply the theorem 11.2. Spectra of $CD(l, k)$ is a subset of

spectra $CD(l+1,q)$, because they are consecutive members of a folder and the following statement is true.

Theorem 11. *The second largest eigenvalue of $CD(k,q)$ is bounded by $2\sqrt{q}$.*

Theorem 12. *Let G be a Chevalley group of rank 2 over the finite field $GF(q)$, $q>2$ of characteristic p, U a Sylow p-subgroup of G and U_1 and U_2 root subgroups corresponding to simple roots. Then affine generalized polygon $\Gamma=\Gamma(U)_{U_1,U_2}$ is a q-regular Ramanujan graph.*

Proof. The geometry $\Gamma(G)$ is a reular generalized m-gon ($m\in\{3,4,6\}$) which is a Ramanujan graph. The second largest eigenvalue of it is $a=2\sqrt{q}\cos(2\pi/m)$ (see [17]). Affine generalized polygon is a q-regular induced subgraph of $\Gamma(G)$. Thus its eigenvalues are interlacing with eigenvalues of $\Gamma(G)$, so the second largest eigenvalue of Γ is bounded by a. Finally, if $q>2$, then $a\le 2\sqrt{q-1}$.

□

Lemma 13. *Let Γ be a k-regular parallelotopic graph with the second largest eigenvalue λ and ${}^{\mathrm{B}}\Gamma$ be an induced subgraph of Γ which contains all vertices which colors belong to a subset B, $t=|B|>\lambda$ in the set of colors C. Then ${}^{\mathrm{B}}\Gamma$ is a t-regular graph with the second largest eigenvalue bounded by λ.*

Proof. We just mention that ${}^{\mathrm{B}}\Gamma$ is a regular induced subgraph of Γ.

□

The lemma above can be useful for the construction of graphs of given degree satisfying additional restrictions on the expansion property and the girth. Let us consider the following example.

Theorem 13. *For each positive integer $n>3$ there is n-regular Ramanujan graph of girth >8.*

Proof. Let us consider the sequence 2^i+1, $i>1$. For given number n we can find an integer i such that $2^i+1<n\le 2^{i+1}+1=q$. If $n=q$ then generalized quadrangle $GQ(q)$ over $GF(q)$ satisfies to requested properties. If not, we can consider ${}^{\mathrm{B}}\Gamma$ (or its connected component), where Γ is affine generalized quadrangle and $|B|=n$. The second largest eigenvalue of ${}^{\mathrm{B}}\Gamma$ is bounded by $2\sqrt{2^{i+1}}\times(\sqrt{2}/2)$, which is $\le 2\sqrt{n-1}$, because of $n>2^i+1$.

□

Remark. If q is "sufficiently large" then affine generalized quadrangle has a diameter 5 ([17]). Thus for such q we can chose graph ${}^{\mathrm{B}}\Gamma$ of diameter 5 in the construction above.

We used the expansion properties of graphs as above to organize encryption process based on methods of parallel computations.

3.4. On small world semiplanes with generalised Schubert cells

3.4.1. On the small world graphs without small cycles

The well known "real -life examples" of small world graphs, including the graph of binary relation: "two persons on the earth know each other" contains cliques, so they have cycles of order 3 and 4. Some problems of Computer Science require explicit construction of regular algebraic graphs with small diameter but without small cycles. The well known examples here are generalised polygons, which are small world algebraic graphs i.e. graphs with the diameter $d \leq c \log_{k-1}(v)$, where v is order, k is the degree and c is the independent constant, semiplanes (regular bipartite graphs without cycles of order 4); graphs that can be homomorphically mapped onto the ordinary polygons. The problem of the existence of regular graphs satisfying these conditions with the degree $\geq k$ and the diameter $\geq d$ for each pair $k \geq 3$ and $d \geq 3$ is addressed in the paper. This problem is positively solved via the explicit construction. Generalised Schubert cells are defined in the spirit of Gelfand-Macpherson theorem for the Grassmanian. Constructed graph, induced on the generalised largest Schubert cells, is isomorphic to the well-known Wenger's graph. We prove that the family of edge-transitive q-regular Wenger graphs of order $2q^n$, where integer $n \geq 2$ and q is prime power, $q \geq n$, $q > 2$ is a family of small world semiplanes. We observe the applications of some classes of small world graphs without small cycles to Cryptography and Coding Theory.

It is well known that the diameter of a k-regular graph (or graph with the average degree k) of order v is at least $\log_{k-1}(v)$ and that the random k-regular graph has diameter close to this lower bound (see [12], c.X, [4], [20],[14],[15]). Only several explicit constructions of families of k-regular graphs with diameter close to $\log_{k-1}(v)$ are known [12], c.X, sec.1, [24]. Most of them have cycles C_3 or C_4.

The problem of constructing infinite families of given degree with small diameter (i.e. with diameter at most $c \log_{k-1}(v)$, $c \geq 1$ is a constant) with certain additional properties is far from trivial. This problem has many remarkable applications in economics, natural sciences, computer sciences and even in sociology. For instance, the "small world graph" of binary relation "two person know each other" on the set of people in the world, has small diameter.

The restriction of this problem on the class of bipartite graphs has additional motivations because such problem for random graphs has been studied by Klee, Larman and Wright, Harary and Robinson, Bollobś and others (see the survey in [12], c. X, sec.5).

Recall that graph Γ be the algebraic graph over K if the set of vertices $V(\Gamma)$ and the neighbourhood of each vertex u are algebraic quasiprojective varieties over the ring K (see [7]).

One of the most important classes of algebraic small world bipartite graphs with additional geometric properties is a class of regular generalised m-gon, i.e. regular tactical configurations of diameter m and girth $2m$. For each parameter m, a regular generalised m-gon has degree $q+1$ and order $2(1+q+\cdots+q^m)$. Up to parameters as above all known examples of regular generalised m-gons are geometries of finite Shevalley group $A_2(q)$, $B_2(q)$ and $G_2(q)$ for $m=3$, 4 and 6, respectively (see, previous section). According to the famous Feit-Higman theorem regular thick (i.e. degree ≥ 3) generalised m-gons exist for $m=3,4$ and 6 only. Thus Generalised Pentagon does not exist, in particular. Each generalised m-gon has a homomorphism (retraction) onto the geometry of diheadral group D_m, which is ordinary m-gon.

We underline the following natural generalizations of regular generalised polygons.

(i) The class of graphs with logarithmic diameter $d \leq c_1 \log_{k-1}(v)$ and logarithmic girth $g \geq c_2 \log_{k-1}(v)$, where c_1, c_2 are some constants. Such graphs are important for communication networks (see [6]). The problem of existence of an infinite family of such graphs with constant degree k has been solved explicitly by Margulis ([69], [70], [71]) and Lubotzky, Phillips and Sarnak [66]. These graphs are not bipartite, they are $q+1$-regular Cayley graphs of $PSL_2(p)$ (p and q are special distinct primes) introduced in [69] and investigated in [66]. In this construction the diameter is bounded by $2\log_{k-1}(v)+2$ and the girth $g \approx \frac{4}{3}\log_{k-1}(v)$. This construction supports the existence of graphs with unbounded logarithmic diameter and logarithmic girth $\geq g$ of degree $\geq k$ for each pair (k,g). Notice that these graphs are not algebraic, because the neighbourhood of each vertex is not a quasiprojective manifold over the field F_p of dimension > 1.

(ii) Other generalisation of generalised m-gon is a flag system with the Coxeter metric of dihedral group D_m (for the definition, see [17], [20]). This class of combinatorial objects is very close to the generalised m-gons. The examples of such systems different from generalised m-gons are unknown.

(iii) Let us consider the class of regular semiplanes, which are bipartite small world graphs and can be epimorphically mapped onto the ordinary polygons. These two conditions are not so restrictive as existence of flag systems with Coxeter metrics. The existence of a homomorphism onto

the ordinary polygon allows to define naturally so called Schubert cells and small Schubert cells on the vertex-set of the graph.

The purpose of this paper is to prove the existence of graphs from this class with the diameter $\geq d$ and degree $\geq k$ for each pair (d, k) via explicit constructions. Our main result is the following statement.

Theorem 14. *For each integer m, $m \geq 2$, and any prime power q, there exists an algebraic over F_q semiplane $SP_m(q)$ of diametrer d, $m \leq d \leq 2m-1$, of order $2(1+q+\dots q^{m-1})$ and degree $q+1$, which can be homomorphically mapped onto the geometry of the dihedral group D_m.*

Note that $SP_3(q)$ and $SP_4(q)$ are isomorphic to geometries of groups $A_2(q)$ and $B_2(q)$, respectively. Semiplane property insures that the girth of the graphs $SP_m(q)$ is ≥ 6. The Schubert geometry of $SP_m(q)$, i.e the totality of all points and lines at maximal distance from standard flag, turns out to be Wenger graph $W_n(q)$, $q > 2$ [127], which is useful for applications in Computer Science. Some of such applications of graphs with small diameter and without cycles C_3, C_4 (Cryptography, Coding theory) we will observe in the last section of the paper.

The problem to evaluate the diameter of Wenger graph was open since graphs had been defined by equations over F_q. In section 5 we prove that graph $W_n(q)$, $q \geq n$ form a family of small world graphs of unbounded degree and diameter.

Notice that the vertex set of $SP_n(q)$ graphs and neighbourhood of each vertex are projective varieties over F_q. In case of Wenger graph these sets are affine varieties over F_q.

Graphs $SP_m(q)$ are defined via equations over F_q written in terms of field addition and multiplication. If we change F_q onto general commutative field K we will get graphs $SP_m(K)$. If K is infinite then $SP_m(K)$ are infinite graphs of diameter $\geq m$ such that we can find a pass of length t, $t \leq 2m+1$ fast, i.e. with $O(m^2)$ arithmetic operations.

An incidence structure is a semiplane if two distinct lines are intersecting not more than in one point and two distinct points are incident not more than in one line. As it follows from the definition, graphs of the semiplane have no cycles C_3 and C_4.

The graph is k-regular if each of its vertex has degree k, where k is a constant.

Let us consider an incidence structure with point set P and line set L, which are two copies of n-dimensional vector space over F_q. It will be convenient for us to denote vectors from P as

$$\mathrm{x} = (x) = (x_{1,0}, x_{1,1}, x_{2,1}, x_{3,1}, \dots x_{i,1}, \dots)$$

and vectors from L as

$$\mathrm{y} = [y] = [y_{0,1}, y_{1,1}, y_{2,1}, y_{3,1}, \ldots, y_{i,1}, \ldots].$$

We say that point (x) is incident with the line $[y]$ and we write it xIy or $(x)\mathrm{I}[y]$ if and only if the following conditions are satisfied:

$$y_{i,1} - x_{i,1} = y_{i-1,1}x_{1,0}$$

where $i = 1, 2, \ldots$.

Let $\mathrm{W}(q)$ be the incidence graph of the structure $\Gamma(F_q) = (P, L, I)$. For each integer $k \geq 2$ let $\Gamma(l, F_q) = (P(k), L(k), I(k))$ be the incidence system, where $P(k)$ and $L(k)$ are the images of P and L under the projection of these spaces on the first k -coordinates and binary relation $I(k)$ is defined by the first k equations. Finally, let $\mathrm{W_k(q)}$ be the incidence graph for $\Gamma(k, F_q)$. This is exactly the graph which has been defined by Wenger [127]. Graph $\mathrm{W}(q)$ is a projective limit of $\mathrm{W}_k(q)$ when k goes to infinity.

Let P_m be the incidence graph of the incidence structure of points (vertices) and lines (edges) of the ordinary m-gon.

For the prime powers k and integers $m \geq 3$ define a family $F(k, m)$ of incidence structures satisfying the axioms (A1)-(A5) below.

(A1) $F(k, m)$ is a family of small world graphs;
(A2) Each $\gamma \in F(k, m)$ is a $k + 1$-regular tactical configuration;
(A3) $\gamma \in F(k, m)$ is a semiplane;
(A4) For each $\gamma \in F(k, m)$ there is a homomorphism $\phi : \gamma \to P_m$ and monomorphism $\eta : P_m \to \gamma$ such that $\phi \circ \eta$ is the identity map and $\eta(P_m)$ is the set of fixed points of $\eta \circ \phi$;
(A5) there is a flag $\{p, l\} \in P_m$ such that $dist(u, \eta(p)) = dist(u, \eta(p))$ and $dist(u, \eta(l)) = dist(u, \eta(l))$ if and only if $\phi(u) = \phi(v)$;

Remark 1. Members of the family as above depend on two parameters k and m. The axiom (A1) is the property of the whole family. Axioms (A2), (A3), (A4) are requirements for each member of the family, so in fact we have infinitely many axioms $Ai(k, m)$, $i = 2, 3, 4, 5$, $k = p^n > 2$, p is a prime number, $n \geq 1$, $m \geq 3$.

Remark 2. If we add the axioms $(A6)(k, m)$: $\mathrm{diam}F(k, m) = \mathrm{diam}P_m$, $\mathrm{girth}(F(k, m)) = \mathrm{girth}(P_m) = 2m$ then the list of axioms will be contradictory, because of Feit-Higman theorem (regular generalised m-gons of degree $r > 2$ exist only for $m \leq 12$).

In the next section we construct explicitly a family of graphs satisfying the axioms $A(1) - A(5)$, so the list of axioms is not contradictory.

The axioms (A4) and (A5) alow us to define the *generalised Schubert cells* in the following way: vertices u and v are in the same cell if and only if $\phi(u) = \phi(v)$ (or distances from u and v to the elements of standard flag $\{p, l\}$ are the same). We can also consider *generalised small Schubert cells*: u and v are in the same cell if $dist(u, x) = dist(v, x)$ for each $x \in \eta(P_m)$. Last equivalence relation is defined in the spirit of Gelfand-MacPherson theorem for the Grassmanian [32], [33]. Note, that axioms (A4) and (A5) guarantee that the graph is connected: there is a walk from each vertex v to the vertex of standard flag.

Remark 3. For each prime power $k \geq 2$ infinite family of graphs $F(k, m)$, $m = 3, 4, \ldots$ is the family of small world graphs of bounded degree.

Remark 4. We can fixe the parameter m in our requirements (A1) - (A6) (formally we will get the list L_m of axioms $A1_m$, $A2_m$, $A3_m$, $A4_m$, $A5_m$, $A6_m$ and define the family $F(k, m)$, where k runs throw all prime powers but m is chosen. Note, that the whole properties are corollaries from $A2_m$ and $A6_m$, which are axioms of regular generalised m-gon which order is a prime power. So this list is contradictory if and only if m is not 3, 4 or 6. The geometries of simple groups $A_2(k)$, $B_2(k)$ and $G_2(k)$ are models for the lists of axioms L_3, L_4 and L_6, respectively.

3.4.2. Main construction

Let us consider the dihedral group D_m and its geometry. The Coxeter group D_m is defined as a group with generators a and b and generic relations $(ab)^m = e$, $a^2 = e$ and $b^2 = e$. The order of D_m is $2m$. The point set and the line set for the geometry D_m is the totality of cosets $D_m : (a)$ and $D_m : (b)$, respectively. Two classes α and β are incident $\alpha I \beta$ if and only if $|\alpha \cap \beta| = 0$ It is easy to see that the geometry is just the incidence structure P_m of vertices (points) and edges (lines) of the ordinary m-gon.

The totality of mirror symmetries (reflections) of ordinary m-gon is the set of the elements with odd length with respect to thee irreducible decomposition into letters of the alphabet $\{a, b\}$. It contains the words $a, b, aba, bab, \ldots$, and the longest element is $(ab)^r a = (ba)^r b$, $2r+1 = 2[m/2]$.

Let $l(\alpha)$, $\alpha \in (D_m : (a)) \cup (D_m : (b))$ be the length of the coset α, i. e. the minimal length of the irreducible representation for representatives of α. Let Δ be the totality of all reflections of the Coxeter group D_m. To each element $\alpha \in \Gamma(D_m)$ we construct the set $\Delta(\alpha) = \{w \in \Delta | l(w\alpha) \leq l(\alpha)\}$. and the vector space $V(\alpha) = (F_q)^{\Delta(\alpha)} = \{f : \Delta(\alpha) \to F_q\}$. We can consider such a vector space as a subspace of ${F_q}^{\Delta}$ consisting of elements satisfying

condition $f(x) = 0$ for $x \in \Delta - \Delta(\alpha)$. The natural basis of ${F_q}^{\Delta}$ is the totality of e_r, where $e_r(r) = 1$ and $e_r(r') = 0$, $r \neq r'$. Let us use "double index notation" for the basis elements: $e_a = e_{1,0}$, $e_b = e_{0,1}$, $e_{aba} = e_{2,1}$, $e_{bab} = e_{3,1}$, $\dots$, $e_{ab^{[m/2]}a} = e_{m-2,1}$.

We can turn ${F_q}^{\Delta}$ into an alternating linear algebra with the multiplication $*$, such that $e_{1,0} * e_{0,1} = e_{1,1}$, $e_{1,0} * e_{i,1} = e_{i+1,1}$, $i = 1, \dots, m-3$ and product of other basis elements is zero. Note that this operation is not associative. In fact it is a Lie bracket (see the last section of the paper).

Let us consider now the following new incidence structure on the set $\tilde{\Gamma}(D_n)$ of elements (α, x), $\alpha \in \Gamma(D_n)$ (element of ordinary n-gon), $\mathrm{x} \in F_q$, We shall assume that (α, x) is a point if and only if α is a point of ordinary n-gon. Two pairs (α, x) and (β, y) are incident (relation I') if and only if the following two conditions hold

(i) $\alpha I \beta$ within geometry of ordinary m-gon
(ii) $\mathrm{x} - \mathrm{y}|_{\Delta(\alpha) \cap \Delta(\beta)} = \mathrm{x} * \mathrm{y}$.

The graph of the incidence relation I' will be denoted as $SP_m(q)$.

We can identify elements of kind $(\alpha, 0)$, where $0(x) = 0$ for each $x \in \Delta$ with the elements of $\tilde{\Gamma}$. Thus we have a natural embedding η of Γ into $\tilde{\Gamma}$. Let us use the term *standard flag* for $(a), (b)$.

Proposition 16. *The degree of each element of $\tilde{\Gamma}$ is $q+1$. The diameter of $\tilde{\Gamma}$ is bounded by $2m-1$. The map $\phi : \tilde{\Gamma} \to \Gamma$, $\phi(\alpha, \bar{x}) = \alpha$, is the homomorphism onto the geometry of ordinary m-gon, the map $\eta : \Gamma \to \tilde{\Gamma}$ is monomorphism, $\phi \circ \eta$ is an identity map and $\eta(\Gamma)$ is the set of fixed elements of $\eta \circ \phi$.*

Proof. The definition of the incidence relation for $\tilde{\Gamma}$ implies that ϕ is an epimorphism. Let (α, f) be the vertex of $\tilde{\Gamma}$. The element α has two neighbors α_1 and α_2 in the polygon. Without loss of generality we may assume that $l(\alpha_1) < l(\alpha_2)$. It is clear that $\Delta(\alpha_1) \subset \Delta(\alpha_2)$ and, as it follows from the definition of the incidence, there is a unique neighbour u of (α, f) such that $\phi(u) = \alpha$. In fact, it is $(\alpha_1, f|_{\Delta(\alpha_1)})$. For the neighbour of (α, f) of kind α_2 we have two different cases: if $l(\alpha_2) > l(\alpha_1)$, then $\Delta(\alpha_2)$ includes $\Delta(\alpha_1)$ and $|\Delta(\alpha_2) \setminus \Delta(\alpha_1)| = 1$ and we have q-neighbours of kind (α, g), such that $g|_{\alpha_1} = f$. Let $l(\alpha_2) = l(\alpha)$, i.e. the cosets α_2 and α have maximal length. Then $|\Delta(\alpha_2)| = |\Delta(\alpha)| = m-1$ and $|\Delta(\alpha) \cap \Delta(\alpha)| = m-2$. As it follows from the definition of the incidence relation, the neighbour of kind (α_2, g) is uniquely determined by $g(w)$, where $\{w\} = \Delta(\alpha_2) \setminus (\Delta(\alpha_2) \cap \Delta(\alpha))$. Thus we have exactly q options there. It means that the degree of each vertex of $\tilde{\Gamma}$ has degree $q+1$.

Let v and u be the vertices of $\tilde{\Gamma}$, their minimal distance to some element of the standard flag is restricted by $m-1$. If v and u are elements of the

same type then the shortest walks from them to elements of the standard flag have the same last element. Thus $\mathrm{dist}(u,v) = 2m-2$. If these elements are of different type then we can combine the shortest walk from the first element, edge of the standard flag and reverse for the shortest walk from the second element to the standard flag. It means that the $\mathrm{dist}(u,v) \leq 2m-1$. □

Proposition 17. *The graph $SP_m(q)$ is a semiplane.*

Proof. We have to prove that the common neighbourhood for two distinct vertices u and v of the same type (both points or both lines) contains at the most one element. Let us consider the case $\phi(u) \neq \phi(v)$. Without loss of generality we may assume that $\Delta(\phi(u))$ contains $\Delta(\phi(v))$ and write u as (α, f). There is a unique common neighbour β of $\phi(u)$ and $\phi(v)$ and $\Delta(\beta)$ is a subset of $\Delta(\alpha)$. It means that the only possible option for the common neighbour is $(\beta, f|_{\Delta(\beta)})$. In fact, the condition of the existence of the unique common neighbour is $v = (\phi(v), f_{\Delta(\phi(v))})$.

Let β be one of the two common neighbours for α in the pentagon. We can write u and v as (α, f_1) and (α, f_2), respectively. Then a possible common neighbour of u and v can be written as (β, g). Consider the following cases:

(*i*) If $l(\beta) > l(\alpha)$ then $\Delta(\beta)$ contains $\Delta(\alpha)$ and $f_1 = f_2 = g|_{\Delta(\alpha)}$. Thus $u = v$ and we get a contradiction in this case.

(*ii*)) Let $l(\beta) < l\alpha$, then possible neighbours have form $(\beta, f_1|_{\Delta(\alpha)})$. The condition of the existence of common neighbour for u and v is $f_1(x) = f_2(x)$ for $x \in \Delta$. Then the unique neighbour of u and v exist in the case $f_1(x) = f_2(x)$ for $x \in \Delta(\beta)$. Note that $f_1(r) \neq f_2(r)$ for the single root r in $\Delta(\alpha) \setminus \Delta(\beta)$.

(*iii*) Let $l(\beta) = l(\alpha)$ and $g(r') = x$ for $r' \neq \Delta(\alpha)$. The values f_1 and f_2 are the following tuples $(a_r, a_{1,1}, \ldots a_{m-2,1})$ and $(b_r, b_{11}, \ldots, b_{m-2,1})$, where r is a simple root different from r'. Let $e(r) = 1$ for $r = (1,0)$ and $e(r) = 0$ for $r = (0,1)$. If $a_r \neq b_r$ then possible x is uniquely defined from the system of two equations

$$a_{1,1} - x_{1,1} = e(r)a_r x, \;\; b_{1,1} - x_{1,1} = e(r)b_r x.$$

Note that in this case $a_{1,1} \neq b_{1,1}$ and there is no neighbour w with $l(\phi(w)) = m-1$. Let $a_r = b_r$ then from the incidence equations we are getting $f_1 = f_2$ which contradicts to $u \neq v$.

Thus u and v have at most one common neighbour.

□

Proposition 18. *The Schubert substructure of $\tilde{\Gamma} = SP_m(q)$ is well defined. It is isomorphic to the Wenger graph $\mathrm{W}_{m-1}(q)$.*

Proof. Let us consider point p and lines l of $\tilde{\Gamma}$ with the property $l(\phi(p)) = m-1$, $l(\phi(l)) = m-1$ and pIl. Then distances from p and l to the nearest vertex from the standard flag equal $m-1$. Thus the generalised largest Schubert cells are well-defined. Let $p = (\alpha, f)$, $l = (\beta, g)$, f and g are defined by tuples $(a_{1,0}, a_{1,1}, \dots, a_{m-2,1})$ and $(b_{0,1}, b_{1,1}, \dots, b_{m-2,1})$. Then incidence condition of $\tilde{\Gamma}$ implies

$$a_{i+1,1} - b_{i+1,1} = a_{1,0}b_{i,1}, i = 0, 1, \dots, m-3.$$

These are the equations that define the Wenger graph.

□

Propositions 3.1-3.3 imply immediately Theorem 1.1 and show that the family of graphs $SP_m(q)$ satisfies to the axioms $A(1) - A(5)$.

Remark. Ordinary m-gon is the incidence geometry for diheadral group D_m of order $2m$. The generalisation of our construction for the case of any Coxeter group the reader can find in [102] r [112]. It is more general, than the blow-up operation in [104] and [105], so in [102] the reader can find wide family of graphs which contaiins incidence geometries of simple groups of Lie type.

3.4.3. Schubert transitivity

Let us consider the affine Kac-Moody Lie algebra $L = \tilde{A}_1$ over the field K defined via 2×2 symmetric extended Cartan matrix (a_{ij}) with $a_{11} = a_{22} = 2$ and $a_{12} = -2$ see [35]. It has a Cartan decomposition $L^- \oplus H \oplus L^+$, where H and $H \oplus L^+$ are the Cartan and the Borel algebras respectively. The algebra L^+ is a direct sum of one dimensional root subalgebras corresponding to positive roots. The set of positive roots in the standard basis of simple roots α_1 and α_2 can be written as tuples $(i+1, i)$, (i, i), $(i, i+1)$, $i = 0, 1, \dots$. Let $<$ be the lexicographical order on the set of positive roots. Let e_α be the basic element from the root subalgebra L_α. We choose a basis of L such that $[e_\alpha, e_\beta] = e_{\alpha+\beta}$ if $\alpha < \beta$ and $\alpha + \beta$ is a root, and identify the elements of L with the tuples in this basis.

For each positive root α and $l \in K$ we consider the automorphism $t_\alpha(l) = \exp(\mathrm{ad}(le_\alpha))$ of the infinite dimensional Lie algebra L^+. This automorphism can change infinitely many components of the vector from L^+, but the close formulae for the i-th component of $t_\alpha(l)(x)$, $x \in L^+$, is the polynomial expression in variables $x_1, \dots x_i$.

Let us consider the direct sum $L(\alpha)$ of L_β such that $\beta \leq \alpha$. Then $t_r(l)$ acts naturally on $L(\alpha)$. Let U and $U(\alpha)$ be the groups generated by $t_r(l)$ where $e_r \in L^+$ and $e_r \in L(\alpha)$, respectively. Then U an $U(\alpha)$ act regularly,

i.e. transitively with a trivial point stabilizer, on the vector spaces L and $L(\alpha)$, respectively.

Consider the subalgebra P of L generated by elements e_{α_1} and e_β, where $\beta = \alpha_1 + \alpha_2$. Then P is a direct sum of L_r, where $r = (i+1, i)$ and (i, i). Let $P(\alpha) = P \cap L(\alpha)$, where $e_\alpha \in P$. Groups $U(P) =< t_r(l) | e_r \in P >$ and $UP(\alpha) = U(P) \cap U(\alpha)$ act regularly on P and P_α, respectively. We will denote any root $\alpha = l\beta + \alpha_1$ corresponding to a root subspace from P as $(l, 1)$. We will also restrict the order $<$ on this set of roots: $(l, 1) < (l', 1)$ if and only if $l < l'$.

The following statement is immediate corollary from the definitions.

Proposition 19. *The Lie algebra $(F_q^\Delta, *)$, which defines the graphs $SP_m(q)$, is isomorphic to $L(\alpha)$ for $\alpha = (m-2, 1)$, considered as a Lie algebra over the ground field F_q.*

Next statement is equivalent to the flag transitivity of the Schubert substructure (*Schubert transitivity*) for the semiplane $SP_m(q)$.

Theorem 15. *Wenger graph $W_m(q)$ is an edge transitive.*

Proof. Consider first the case of $\mathrm{char} F_q \geq m$. Let α^* be the dual root for $\alpha = (1, 0)$. Then α^* is a basis element of the Cartan subalgebra H. The multiplication rule in $H \oplus L^+$ for α is $[\alpha^*, e_r] = 2e_r$, where $r \neq (0, 1)$ and $[\alpha^*, e_{0,1}] = 0$.

Let us consider the external derivation β^* which is "dual" to $\beta = (0, 1)$: $[\beta^*, e_r] = \beta^*(r)e_r$, where $\beta^*(i, 1) = i$ and consider the the subalgebra $\tilde{L} =< \alpha^*, \beta^*, L^+ >$. We shall identify points $(x_{1,0}, x_{1,1}, \ldots, x_{m-1,1})$ and lines $[y_{0,1}, y_{1,1}, \ldots, y_{m-1,1}]$ with the elements

$$\tilde{x} = \alpha^* + \sum_{i=1}^{m-1} \frac{1}{i} x_{i,1} e_{i,1} + x_{1,0} e_{1,0}$$

and

$$\tilde{y} = \beta^* + \frac{1}{2} \sum_{i=0}^{m-1} y_{i,1} e_{i,1},$$

respectively.

We can rewrite the incidence condition of Wenger graph in the form $[\tilde{x}, \tilde{y}] = 0$. Elements $u = t_r(l)$ preserve the Lie bracket and the group $UP(\alpha)$, $\alpha = (m-1, 1)$ acts regularly on the set of pairs $(\tilde{x}, \tilde{y})$ such that $[\tilde{x}, \tilde{y}] = 0$ according to the rule: $\tilde{x} \to \tilde{x}^u|_{(L^+ - L_{0,1})}$, $\tilde{y} \to \tilde{y}^u|_{(L^+ - L_{1,0})}$. Thus Wenger graph is an edge transitive for $p = \mathrm{char}(F_q) \geq m$.

We can write close formula for each transformation $t_\alpha(l)$ acting on $P \cup L$ in the form $x_r \to x_r + f_r(x_{1,0}, \ldots, x_{r'})$, $y_r \to y_r = g_r(y_{0,1}, \ldots, y_{r'})$, $r' < r$,

which preserve the incidence relation for the case of small characteristic as well.

These transformation generate the group which acts regularly on the vertices of Wenger graph.

□

The *spanning tree* of the graph G with the vertex set $V(G)$ and edge st E is the tree T with vertex set V and the edge set which is subset of E.

Let us remove all edges between elements from the "largest Schubert cells" i.e. elements of kind (α, x), where $l(\alpha) = m-1$. After the completion of this operation we shall get the graph $ST_m(q)$.

Lemma 14. *Graph $ST_m(q)$ is a spanning tree for the graph $SP_m(q)$.*

Proof. Let us consider the process of walking from one of the vertices $(< a >, 0)$ or $(< b >, 0)$ which does not contain the edge between these two vertices. This branching processes produce rooting trees $T_{<a>}$ and $T_{<b>}$. They do not contain common vertices. So adding extra edge between $(< a >, 0)$ and $(< b >, 0)$ leads to the tree $ST_m(q)$, which contains all vertices of $PC_m(q)$.

□

Let us simplify the notations for basic elements of $L_n(q)$.

It can be considered as Lie algebra with the basis $e_1, e_2, \ldots, e_n$ such that $e_i \times e_1 = e_{i+1}$, $e_1 \times e_i = -e_{i+1}$, $i \geq 2$, $e_i \times e_j = 0$ if 1 is not an element of $\{i, j\}$. The Wenger graph $W_{n-1}(q)$ corresponds to the following incidence relation I on $L \cup P$ where pIl holds for $l = [x_2, x_3, \ldots, x_n]$ and $p = (y_1, y_3, \ldots, y_n)$ if and only if $y_m - x_m = y_1 x_{m-1}$ for each $m = 3, \ldots, n$ (see [36]).

Let us consider the polynomial bijections $t_i(x)$, $x \in F_q$, $i = 1, \ldots, n$ defined on $P \cup L$ by the following formulae:

$$
\begin{aligned}
(a_1, a_3, \ldots, a_n)^{t_2(x)} &= (a_1, a_3 + a_1 x, a_4, \ldots, a_n),\\
[b_2, b_3, \ldots, b_n]^{t_2(x)} &= [b_2 + x, b_3, \ldots, b_n];\\
(a_1, \ldots, a_k, \ldots a_n)^{t_k(x)} &= (a_1, \ldots, a_{k-1}, a_k + x, a_{k+1} + a_1 x, \ldots, a_n),\\
[b_2, \ldots, b_k, \ldots, b_n]^{t_k(x)} &= [b_2, \ldots, b_{k-1}, b_k + x, b_{k+1}, \ldots, b_n],\ k \geq 3;\\
(a_1, \ldots a_n)^{t_1(x)} &= (a_1 + x, a_3, \ldots, a_k - {C^1}_{k-3} a_{k-1} x + \ldots\\
&\quad + {C^i}_{k-3} x^i (-1)^i + \ldots),\\
[b_2, b_3, \ldots, b_n]^{t_1(x)} &= [b_2, b_3 - b_2 x, \ldots, b_k - {C^1}_{k-2} b_{k-1} x + \ldots\\
&\quad + {C^i}_{k-2} b_{k-i} - x^i \ldots],
\end{aligned}
$$

where $C^i{}_k$ is a binomial coefficient.

The following statement follows immediately from the above formula.

Proposition 20. *For every $x \in K$ the transformations $t_i(x)$, $i = 1, \ldots, n$ are automorphisms of $W_{n-1}(q)$.*

In fact the group $U =< t_i(x) | i = 1, \ldots, n >$ coincides with the edge transitive group used in the proof of Theorem 4.2.

3.5. On the diameter of Wenger graph

Theorem 16. *If $q \geq n$, then the diameter of the Wenger graph $W_n(q)$ is bounded by $2n + 2$.*

Proof. We shall work with the incidence structure (P, L, I), where

$$P = \{(x) = (x_1, x_2, \ldots, x_n) | x_i \in F_q, i = 1, \ldots, n\},$$
$$L = \{[y] = [y_1, \ldots, y_n] | y_i \in F_q, i = 1, \ldots, n\}$$

are poin-set and line-set, $(x)I[y]$ if and only if

$$y_i - x_i = y_1 x_{i-1}, \quad i = 2, \ldots, n.$$

The map

$$(a_1, a_3, \ldots, a_n) \to (-a_1, -a_3, \ldots, -a_{n+1}),$$
$$[b_2, b_3, \ldots, b_{n+1}] \to [-b_2, -b_3, \cdots - b_{n+1}]$$

establish the isomorphism between Wenger graph $W_n(q)$ and incidence graph of binary relation I. Let us investigate the walks

$$[y^0]I(x^0)I[y^1]I(x^1)I \ldots I[y^{n-1}]I(x^{n-1})I[y^n]I[y^i]I(x^i)I[y^{i+1}]$$

for $i = 0, \ldots, n-1$ of length $2n$ between two lines of the incidence structure. We can assume that the first line $[y_0]$ is zero line $[0]$ because we proved the edge transitivity of the graph. Let z_1, $i = 1, \ldots, n$ be the first component of tuples $[y_1]$. We write the first components of $[y^i]$ in the form $y_1^{i-1} + z_i$. Then the first component of $[y^n]$ will be $z_1 + z_2 + \ldots z_n$. Let us assume that the first component of the point (x^i), $i = 0, \ldots, n-1$ is a_i. Note that the starting point and first components of points and lines define our walk. We can use the incidence equations and write the components $[b_1, b_2, \ldots b_n] = [y^n]$ of the last line explicitly:

$$\begin{aligned} z_1 + z_2 + \dots z_n &= b_1 \\ a_0 z_1 + a_1 z_2 + \dots a_{n-1} z_n &= b_2 \\ {a_0}^2 z_1 + {a_1}^2 z_2 + \dots + {a_{n-1}}^2 z_n &= b_3 \\ &\dots \\ {a_0}^{n-1} z_1 + (a_1)^{n-1} z_2 + \dots {a_{n-1}}^{n-1} z_n &= b_n \end{aligned}$$

So we get the system of linear equations in variables $z_1, \dots z_n$ with the Vandermonde determinant D, which is the product of $(n-1)n/2$ differences $(a_i - a_j)$, where i, j are distinct elements of $\{0, 1, \dots, n-1\}$. It means that if $q \geq n$ we can take arbitrary line $[b_1, b_2, \dots, b_n]$ chose distinct elements $a_0, a_1, \dots, a_n$, get the solution of the above system of linear equations $z_1 = c_1, c_2, \dots, c_n$. So we have a walk from $[0]$ to $[b_1, \dots, b_n]$ defined by the sequence of first components $0, a_0, c_1, a_1, c_1 + c_2, a_2, \dots c_1 + c_2 \dots, +c_{n-1}, a_{n-1}, c_1 + c_2 + \dots + c_n$. So we constructed the walk between zero line and $[b_1, dots, b_n]$ and proved that the maximal length between two lines is bounded by $2n$. Note that we can add zero point (0) to our walk between $[0]$ and $[b_1, \dots, b_n]$ and edge transitivity of the graph allows us to bound the maximal distance between point and line by $2n + 1$. Finally, we can take arbitrary point (p) and its neighbour $[n]$ and get the pass and construct the walk from (0) to (p) of length $\leq 2n + 2$.

□

Corollary 5. *The family of Wenger graphs $W_n(q)$, $q \geq n \geq 2$, $q > 2$ is family of small world graphs.*

Remark. It can be shown that for each q the family of graphs $W_n(q)$, $n = 2, 3, \dots$ is not a family of small world graphs, but the "enveloping graphs" $SP_n(q)$ form the family of small world graphs without cycles C_4 for each q. So such a family do not satisfy to the axiom $A4$ and $A5$.

3.6. Automorphisms and connected components of $D(n, K)$ in case of general commutative ring K

We need the following well known results on groups acting on graphs.

Let G be a group with proper distinct subgroups G_1 and G_2. Let us consider the incidence structure with the point set $P = (G : G_1)$ and the line set $(G : G_2)$ and incidence relation $I : \alpha I \beta$ if and only if the set theoretical intersection of cosets α and β is nonempty set. We shall not distinguish the incidence relation and corresponding graph $\Gamma(G)_{G_1, G_2}$. Let $l(g)$ be the minimal length of representation of g in the form of products

of elements from G_1 and G_2 The following statement had been formulated first by G. Glauberman.

Lemma 15. *Graph I is connected if and only if $< G_1, G_2 >= G$. The diameter of I is* $\max l(g)$, $g \in G$.

Let

$$A =< a_1, \dots, a_n | R_1(a_1, \dots, a_n), \dots, R_d(a_1, \dots, a_n) >,$$
$$B =< b_1, \dots b_m | S_1(b_1, \dots b_m), \dots, S_t(b_1, \dots, b_m) >$$

are subgroups with generators a_i, $i = 1, \dots, n$ and b_j, $j = 1, \dots, m$ and generic relations R_i, $i = 1, \dots, d$ and S_j, $j = 1, \dots, t$, respectively. Free product $F = A * B$ of A and B be the subgroup

$$< a_1, \dots, a_n, b_1, \dots, b_m | R_1, \dots R_d, S_1, \dots, S_t > \quad \text{(see [68])}.$$

The definition of an operation of free product F_H of groups A and B amalgamated at common subgroup H can be found in [68]. If $H =< e >$, then $F_H = A * B$.

Theorem 17. (see, for instance [68]) *Let G acts edge transitively but not vertex transitively on a tree T. Then G is the free product of the stabilizers G_a and G_b of adjacent vertices a and b amalgamated at their intersection.*

Corollary 6. *Let G acts edge regularly on the tree T, i. e. $|G_a \cap G_b| = 1$. Then G is the free product $G_a * G_b$ of groups G_a and G_b.*

In section 2. 5 we define the family of graphs $D(k, K)$, where $k > 2$ is positive integer and K is a commutative ring, such graphs have been considered in [59] for the case $K = F_q$.

The incidence relation motivated by the linear interpretation of Lie geometries in terms their Lie algebras [105] (see [104]). Let us define the "root subgroups" U_α, where the "root" α belongs to the root system

$$\text{Root} =\{(1,0), (0,1), (1,1), (1,2), (2,1), (2,2), (2,2)' \dots,$$
$$(i,i), (i,i)', (i,i+1), (i+1,i) \dots\}.$$

The "root system above" contains all real and imaginary roots of the Kac-Moody Lie Algebra $\tilde{A}_1$ with the symmetric Cartan matrix. We just doubling imaginary roots (i,i) by introducing $(i,i)'$.

Group U_α generated by the following "root transformations" $t_\alpha(x)$, $x \in K$ of the $P \cup L$ given by rules $p_\beta = p_\beta + r_\beta(x)$, $l_\beta = l_\beta + s_\beta(x)$, where $\beta \in$ Root and the functions $r_\beta(x)$, $s_\beta(x)$ are consist of summands defined by the following tables ($i \geq 0$, $m \geq 1$).

	$s_{0,1}(x)$	$s_{1,0}(x)$	$s_{m,m+1}(x)$	$s_{m+1,m}(x)$	$s_{m,m}(x)$	$s'_{m,m}(x)$
$l_{i,i}$		$-l_{i,i-1}x$	$+l_{r,r-1}x$, $r-m>1$		$-l_{r,r}x$, $r-m>0$	
$l_{i,i+1}$		$(l_{i,i}+l'_{i,i})x$ $+l_{i,i-1}x^2$	$+l'_{r,r}x$, $r=i-m\geq 0$		$-l_{r,r+1}x$, $r=i-m\geq 0$	
$l_{i+1,i}$	$+l_{i,i}x$			$-l_{r,r}x$, $r=i-m>0$		$+l_{r+1,r}x$, $r=i-m>0$
$l'_{i,i}$	$l_{i-1,i}x$	$l_{i,i-1}x$		$-l_{r-1,r-1}x$, $r=i-m>1$		$+l'_{r,r}$, $r=i-m>0$

TABLE 1

	$r_{0,1}(x)$	$r_{1,0}(x)$	$r_{m,m+1}(x)$	$r_{m+1,m}(x)$	$r_{m,m}(x)$	$r'_{m,m}(x)$
$p_{i,i}$	$+p_{i-1,i}x$	$p_{i,i-1}x$	$+p_{r,r-1}x$ $r = i - m > 1$		$-p_{r,r}x$ $r = i - m < 0$	
$p_{i,i+1}$		$+p'_{i,i}x$	$+p'_{r,r}x$ $r = i - m > 0$		$-p_{r,r+1}x$ $r = i - m > 0$	$+p_{r+1,r}x,$ $r = i - m \geq 0$
$p_{i+1,i}$	$(p_{i,i} + p'_{i,i})x$ $+p_{i-1,i}x^2$			$-p_{r,r}x,$ $r = i - m \geqslant 0$		
$p'_{i,i}$	$p_{i-1,i}x$			$-p_{r-1,r}x,$ $r = i - m > 1$		$+p'_{r,r},$ $r = i - m < 0$

TABLE 2

Proposition 21. (*i*) *For each pair* (α, x), $\alpha \in \mathrm{Root}$, $x \in K$ *the transformation* $t_\alpha(x)$ *are automorphisms of* $D(K)$. *The projections of these maps onto the graph* $D(n, K)$, $n \geq 2$ *are elements of* $\mathrm{Aut}(D(n, K))$.

(*ii*) *Group* $\mathrm{U}(K)$ *acts edge regularly on the vertices of* $\mathrm{D}(K)$.

(*iii*) *Group* $\mathrm{U}(n, K)$ *generated by projections of* $t_\alpha(x)$ *onto the set of vertices* V *of* $\mathrm{D}(n, K)$ *acts edge regularly on* V.

Proof. Statement (i) follows directly from the definitions of incidence and closed formulas of root transformations $t_\alpha(x)$. Let $<$ be the natural lexicographical linear order on roots of kind (i,j), where $|i-j| \le 1$. Let us assume additionally that $(i,i) < (i,i)' < (i,i+1)$. Then by application of transformations $t_\alpha(x_\alpha)$, $\alpha \ne (0,1)$ to a point (p) consecutively with respect to the above order, where parameter x_α is chosen to make α component of the image equals zero, we are moving point (p) to zero point (0). A neighbour $[a,0,\dots,0]$ of the zero point can be shifted to the line $[0]$ by the transformation $t_{(1,0)}(-a)$. Thus each pair of incident elements can be shifted to $((0),[0])$ and group U acts edge regularly on vertices of $D(K)$. This action is regular $((ii))$ because the stabilizer of the edge $(0),[0]$ is trivial. Same arguments about the action of $U(n,K)$ justify (iii).

□

Remark. For $K = F_q$ this statement had been formulated in [59].

Let $k \ge 6$, $t = [(k+2)/4]$, and let

$$u = (u_\alpha, u_{11}, \dots, u_{tt}, u'_{tt}, u_{t,t+1}, u_{t+1,t}, \dots)$$

be a vertex of $\mathrm{D}(k,K)$ ($\alpha \in \{(1,0),(0,1)\}$, it does not matter whether u is a point or a line). For every r, $2 \le r \le t$, 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 $a = a(u) = (a_2, a_3, \cdots, a_t)$.

Proposition 22. *(i) The classes of equivalence relation*

$$\tau = \{(u,v) | a(u) = a(v)\}$$

form the imprimitivity system of permutation groups $\mathrm{U}(K)$ *and* $\mathrm{U}(n,K)$

(ii) For any $t-1$ *ring elements* $x_i \in K$*),* $2 \le t \ge [(k+2)/4]$*, there exists a vertex* v *of* $\mathrm{D}(k,K)$ *for which*

$$a(v) = (x_2, \dots, x_t) = (x).$$

(iii) The equivalence class C *for the equivalence relation* τ *on the set* $K^n \cup K^n$ *is isomorphic to the affine variety* $K^t \cup K^t$ *,* $t = [4/3n] + 1$ *for* $n = 0,2,3 \bmod 4$*,* $t = [4/3n] + 2$ *for* $n = 1 \bmod 4$*.*

Proof. Let C be the equivalence class on τ on the vertex set $\mathrm{D}(K)$ ($\mathrm{D}(n,K)$ then the induced subgraph, with the vertex set C is the union of several connected components of $\mathrm{D}(K)$ ($\mathrm{D}(n,K)$).

Without loss of generality we may assume that for the vertex v of $C(n, K)$ satisfying $a_2(v) = 0, \dots a_t(v) = 0$. We can find the values of components $v'_{i,i)}$ from this system of equations and eliminate them. Thus we can identify P and L with elements of K^t, where $t = [3/4n] + 1$ for $n = 0, 2, 3 \mod 4$, and $t = [3/4n] + 2$ for $n = 1 \mod 4$.

□

We shall use notation $C(t, K)$ $(C(K))$ for the induced subgraph of $D(n, K)$ with the vertex set C.

Remark. If $K = F_q$, q is odd, then the graph $C(t, k)$ coincides with the connected component $CD(n, q)$ of the graph $D(n, q)$ (see [62]), graph $C(F_q)$ is a q-regular tree. In other cases the question on the connectedness of $C(t, K)$ is open. It is clear that $g(C(t, F_q))$ is $\geq 2[2t/3] + 4$.

Let $U_\alpha =< t_\alpha(x) | x \in K >$ be a subgroup of $U(K)$. It is isomorphic to the additive group K^+ of the ring K. Let U^C be subgroup generated by $t_\alpha(x)$, $x \in K$, $\alpha \in \{(0, 1, (1, 0), \dots, (i, i), (i, i + 1), \dots\}$. Let ${U_n}^C$ be the subgroup generated by transformations $t_\alpha(x)$ from U^C onto the graph $D(n, K)$ (or $C(n, K)$).

Proposition 23. *(i) The connected component $CD(n, K)$ of the graph $D(n, K)$ (or its induced subgraph $C(t, K)$) is isomorphic to $\Gamma({U_n}^C)_{U_{(0,1)}, U_{(1,0)}}$.*

(i) Projective limit of graphs $D(n, K)$ (graphs $C(t, K)$, $CD(n, K)$) with respect to standard morphisms of $D(n + 1, K)$ onto $D(n, K)$ (their restrictions on induced subgraphs) equals to $D(K)$ $(C(K)$, $CD(K) = {U^C}_{U_{(0,1)}, U_{(1,0)}}$, respectively).

If K is an integrity domain, then $D(K)$ and $CD(K)$ are forests. Let C be the connected component, i.e tree.

Group U^C acts regularly on $CD(K)$. So we can apply theorem on group acting regular on the tree and get the following statement.

Proposition 24. *If K is integrity domain then group $U^C(K)$ is isomorphic to the free product of two copies of K^+.*

Theorem 18. *The diameter of the graph $C_m(K)$, $m \geq 2$, where K is a commutative ring with unity of odd characteristic, is bounded by function $f(m)$, defined by the following equations:*

$$f(m) = \begin{cases} (32/3)(4^{(m+1)/3} - 1) - m + 7, & \text{for } m = 2 \pmod 3 \\ (32/3)(4^{(m-1)/3} - 1) + 4^{(m+5)/3} - m + 7 & \text{for } m = 1 \pmod 3 \\ (32/3)(4^{m/3} - 1) + 32 \times 4^{(m-3)/3} - m + 7, & \text{for } m = 0 \pmod 3 \end{cases}$$

Proof. Let $C = C_t(K)$ be the block of equivalence relation τ, containing zero point and zero line. Let us consider the stabiliser of this block. It is clear that group G generated by elements $t_{i,i+1}(x)$, $t_{i+1,i}(x)$, $i \geq 0$, $t_{1,1}(x)$ and $t_i(x) = t_{i,i}(x)t'_{i,i}(x)$, $i \geq 2$, $x \in K$ stabilises C and acts regularly on this set.

Let $l(g)$ be the minimal length of irreducible representation of $g \in G$ in the form

$$T_1(x_1)T_2(x_2)\dots T_d(x_d), x_i \in K, \tag{3.1}$$

where consecutive elements $T_i(x_i)$ and $T_{i+1}(x_{i+1})$ belong to different subgroups U_1 and U_2.

As it follows from the group theoretical interpretation of lemma 3 the diameter of group G is equal to the maximal length $l(g)$.

Let $G_{1,1}$ be the totality of all commutator elements $[t_{0,1}(x), t_{1,0}(y)] = t(x, y)$. Then applications of $T_{1,1}(y) = t(1, y)$ to zero point (0) (or line) do not change its first component. For the second component $u_{1,1}$ of (u) $= (0)^{T_{1,1}(y)}$ we have $u_{1,1} = y$. In fact, $(O)^{T_{1,1}(y)} = (O)^{t_{1,1}}(y)$ and $l(\mathrm{u}) \leq 4$.

Let us consider the totality $G_{1,2}$ of the commutators

$$t(x, y) = [t_{0,1}(x), T_{1,1}(y)].$$

Then its action of on zero line (point) does not change its first, second components. The third component will be $2xy$. Let us consider $T_{1,2}(y) = t(x/2, y)$. Let $u = [O]^{T_{1,2}(y)}$, then $u_{1,2} = y$. Similarly, we construct the totality $G_{2,1}$ of commutators $t(x, y)[t_{1,0}(x)T_{1,1}(y)]$ containing element $T = T_{2,1}(y)$, such that $O^T = O^{T_{2,1}(y)} = [0, 0, 0, y, \dots]$. We can write the irreducible presentation of $g \in G$ in the form (3.1) starting either with element from U_1 or U_2. It means that $l(g) \leq 8$ for $g \in G_{1,2} \cup G_{2,1}$

Let us define $G_{2,2}$ as totality of commutators $[t_{1,0}(x), T_{1,2}(y)]$ (or equivalently as set of elements of kind $[t_{0,1}(x), T_{2,1}(y)]$. Then for element $t \in G_{2,2}$ we have $O^t = O^{t_2(xy)} = (0, 0, 0, 0, xy, xy, \dots)$. We have $l(g) \leq 16$ for $g \in G_{2,2}$.

We can define recurrently $G_{i,i+1}$, $G_{i+1,i}$ and $G_{i+1,i+1}$, $i \geq 2$ as totalities of elements of kind $[t_{0,1}(x), T_{i,i}(y)]$, $[t_{1,0}(x), T_{i,i}(y)]$ and
$[t_{0,1}(x), T_{i,i+1}(y)]$, respectively. The length of elements from $G_{i,i+1}$ and $G_{i+1,i}$ are bounded by 2^{2i+1} and $l(g) \leq 2^{2i+2}$ for $g \in G_{i+1,i+1)}$. Notice, that the element $g \in G_\alpha$ acting on element v (point or line) changing only components v_β , $\beta > \alpha$. We can find an element $g \in G_\alpha$, such that for $u = v^g$ the component u_α equals zero.

Let $u \in G$ be element such that $O^u = v$. Then by consecutive applications of appropriate transformations $g \in G_\alpha$ with respect to natural order on roots we can move v to O It means that each element $g \in G$ can be

presented as product $g_{0,1}g_{1,0}g_{1,1}\dots g_\alpha\dots$, where $g_\alpha \in G_\alpha$. Let $d(\alpha)$ be the length of g_α. We can bound the length of g by the sum S of d_α. In case when α is not simple root we have a choice to write irreducible representation of g_α, is with the first character from U_1 or the one from U_2. It allows slightly improve the bound foe the diameter - get $S-m+1$ instead of S.

Let us count S for the case $m = 2\mathrm{mod}\ 3$. If $m = 2$ then $S = 6$. In case of $m \geq 5$ each triple of roots $(i, i+1)$, $(i+1, i)$, $(i+1, i+1)$, $i \geq 1$ contributes summands 2^{2i+1}, 2^{2i+1} and 2^{2i+2}. So we can count S via the sum of the geometrical progression.

Let $m = 2\mathrm{mod}\ 3$ then each triple as above contribute summand 2^{2i+3}. So we have the geometrical progression 2^{2i+3}, $i = 1, \dots (m-2)/3$. The roots $(0,1)$, $(1,0)$ and $(1,1)$ contribute 6.

In case $m = 0\mathrm{mod}\ 3$ we have a geometrical progression 2^{2i+3}, $i = 1, \dots, m/3-1$ and last root contributes $32 \times 4^{m/3-1}$.

In case $m = 1\mathrm{mod}3)$ we have a geometrical progression 2^{2i+3}, $i = 1, \dots, (m-4)/3$ and two last roots contributes $64 \times 4^{(m-4)/3}$

This way we are getting the formulae for the bound.

□

Remark. Theorem 1 follows directly from theorem 12 and Proposition 3.

3.7. On some applications

The idea to use simple graphs of large girth in Cryptography had been explored in [107], [108], [109], [110], [111], [113], [114], [116], [117], [118], [119], in particular.

The definitions of family of graphs of large girth, small world graphs for the class of irreflexive binary relation graphs will be f0rmulated in chapter 4, where more general encryption scheme for the "potentially infinite" text based on the graphs of binary relations with special "rainbow-like" coloring of arrows will be proposed.

For this purpose we identify the vertex of the graph with the plaintext, encryption procedure corresponds to the chain of adjacent vertices starting from the plaintext, the information on such chain is given by the sequence of colours (passwords). We assume that the end of the chain is the ciphertext.

The important feature of such encryption is the resistance to attacks, when adversary intercepts the pair plaintext - ciphertext. It is true because the best algorithm of finding the pass between given vertices (by Dijkstra , see [26] and latest modifications) has complexity $n\ln n$ where n is the order of the graph, i.e. size of the plainspace. The situation is similar to the checking of the primality of Fermat's numbers $2^{2^m}+1$: if the input given

by the string of binary digits, then the problem is polynomial, but if the input is given by just a parameter m, then the task is NP-complete.

We have an encryption scheme with the flexible length of the password (length of the chain). If graphs are connected then we can convert each potentially infinite plaintext into the chosen string "as fast as it is possible".

Finally, in the case of "algebraic graphs" (see [7]) with the special "rainbow-like" coloring (symbolic rainbow-like graphs of section 3) there is an option to use symbolic computations in the implementation of graph based algorithm. We can create public rules symbolically and use the above algorithm as public key tool (for the example of implementation look at [40]). Notice, that for the use of graphs in public key mode the girth property is not so important. The Wegner graphs can be used effectively for such purpose (see [99]).

As we mentioned before, the first explicit examples of families with large girth with arbitrary large valency were given by Margulis. The constructions were Cayley graphs $X^{p,q}$ of group $SL_2(Z_q)$ with respect to special sets of $q+1$ generators, p and q are primes congruent to 1 mod4. The family of $X^{p,q}$ is not a family of algebraic graphs because the neighbourhood of each vertex is not an algebraic variety over F_q. For each p, graphs $X^{p,q}$, where q is running via appropriate primes, form a family of small world graph of unbounded diameter.

The fist family of connected algebraic graphs over F_q of large girth and arbitrarily large degree had been constructed in [17]. These graphs $CD(k,q)$, k is an integer ≥ 2 and q is odd prime power had been constructed as connected component of graphs $D(k,q)$ defined earlier. For each q graphs $CD(k,q)$, $k \geq 2$ form a family of large girth with $\gamma = 4/3\log_{q-1}q$.

Some new examples of simple algebraic graphs of large girth and arbitrary large degree the reader can find in [118].

3.8. On Lie geometries their flag systems and applications in Coding Theory and Cryptography

We propose some cryptographical algorithms based on finite BN-pair G defined over the fields F_q. We convert the adjacency graph for maximal flags of the geometry of group G into a finite Tits automaton by special colouring of arrows and treat the largest Schubert cell Sch $= {F_q}^N$ on this variety as a totality of possible initial states and a totality of accepting states at a time. The computation (encryption map) corresponds to some walk in the graph with the starting and ending points in Sch. To make algorithms fast we will use the embedding of geometry for G into Borel subalgebra of corresponding Lie algebra. We consider the induced subgraph of adjacency graph obtained

by deleting all vertices outside of largest Schubert cell and corresponding automaton (Schubert automaton). We consider the following symbolic implementation of Tits and Schubert automata. The symbolic initial state is a string of variables x_α, where roots α are listed according Bruhat order, choice of label will be governed by linear expression in variables x_α, where α is a simple root.

Conjugations of such nonlinear map with element of affine group acting on ${\mathrm{F_q}}^N$ can be used in Diffie-Hellman key exchange algorithm based on the complexity of group theoretical discrete logarithm problem in case of Cremona group of this variety. We evaluate the degree of these polynomial maps from above and the maximal order of this transformation from below. For simplicity we assume that G is a simple Lie group of normal type but the algorithm can be easily generalised on wide classes of Tits geometries. In a spirit of algebraic geometry we generalise slightly the algorithm by change of linear governing functions for rational linear maps.

According to Hilbert's approach to Geometry it is a special incidence system (or multipartite graph). Felix Klein thought that the Geometry was a group and proposed his famous Erlangen program. J. Tits combined those two ideas for the development of concept of a BN-pair, its geometry and flag system [96], [97]. He created an axiomatic closure for such objects based on the definition of building [98].

Finite geometries $\Gamma(G(q))$ of BN-pair $G(q)$ with Weyl group W defined over finite field F_q, $q \to \infty$ form a family of small world graphs. Really, the diameters of the incidence graphs for $\Gamma(G(q))$ coincide with the diameter of Weyl geometry $\Gamma(W)$, but average degree is growing with the growth of parameter q. The problem of constructing infinite families of small world graphs has many remarkable applications in economics, natural sciences, computer sciences and even in sociology. For instance, the "small world graph" of binary relation "two person shake hands" on the set of people in the world has small diameter.

The algorithm of finding the shortest pass between two arbitrarily chosen vertexes of $\Gamma(G(q))$ is much faster than the action of general Dijkstra algorithm. One can find the pass in $\Gamma(G(q))$ for the time c, where c is a constant independent on q. Regular graphs of simple groups of Lie type of normal type of rank 2 (generalised m-gons for $m \in \{3, 4, 6\}$ support the sharpness of Erdös' bound from Even Circuit Theorem in cases of cycles of length $4, 6$ and 10 (see [11]).

One of the constructions which provide for each $k_0 \geq 2$ the infinite family of regular graphs of degree $k, k \geq k_0$ of large girth (length of minimal cycle) is based on the properties of the geometry of Kac-Moody BN-pair $G(q)$ with diagram $\tilde{A}_1$ (see [57], [58], [59])

The geometries of finite BN-pairs are traditionally used in classical

Coding Theory. Foundations of this theory are based on the concept of finite distance-transitive or distance-regular metrics (distance regular and distance transitive graphs in other terminology [17]). Large number of known families of distance transitive graphs are constructed in terms of the incidence geometry of BN-pair or geometry of its Weyl group. Known constructions of families of distance - regular but not distance transitive graphs are also based on the properties of BN-pair geometries (see [17], [20]). Linear codes are just elements of projective geometry and all applications of Incidence Geometries to Coding Theory are hard to observe (see [82], [34], [78] and further references). Notice that some nonclassical areas like LDPS codes and turbocodes use objects constructed via BN-pair geometries: for the first constructions of LDPS codes Tanner [94] used finite generalised m-gons, the infinite family of graphs of large girth defined in [59] have been applied to constructions of the LDPS codes ([82], [34], [45], [46], [39] and further references)

Quite recent development gives an application of linear codes and their lattices to cryptography. Incidence geometries were used in [110] and [36] for the development of cryptographical algorithms.

In next units we generalise some 0f these encryption algorithms of and consider the key exchange protocols based on geometries of BN-pairs.

3.8.1. On Coxeter systems and BN-pairs

An important example of the incidence system as above is the so-called *group incidence system* $\Gamma(G, G_s)_{s\in S}$. Here G is the abstract group and $G_{ss\in S}$ is the family of distinct subgroups of G. The objects of $\Gamma(G, G_s)_{s\in S}$ are the left cosets of G_s in G for all possible $s \in S$ Cosets α and β are incident precisely when $\alpha \cap \beta \neq \emptyset$. The type function is defined by $t(\alpha) = s$ where $\alpha = gG_s$ for some $s \in S$.

Let (W, S) be a Coxeter system, i.e. W is a group with set of distinguished generators given by $S = \{s_1, s_2, \ldots, s_l\}$ and generic relation $(s_i \times s_j)^{m_{i,j}} = e$. Here $M = (m_{i,j})$ is a symmetrical $l \times l$ matrix with $m_{i,i} = 1$ and off-diagonal entries satisfying $m_{i,j} \geq 2$ (allowing $m_{i,j} = \infty$ as a possibility, in which case the relation $(s_i \times s_j)^{m_{i,j}} = e$ is omitted). Letting $W_i =< S - \{s_i\} >$, $1 \leq i \leq l$ we obtain a group incidence system $\Gamma_W = \Gamma(W, W_i)_{1\leq i\leq l}$ called the Coxeter geometry of W. The W_i are referred to as the *maximal standard subgroups* of W (see [16]).

Let G be a group, B and N subgroups of G, and S a collection of cosets of $B \cap N$ in N. We call (G, B, N, S) a *Tits system* (or we say that G *has a* BN*-pair*) if

(*i*) $G =< B, N >$ and $B \cap N$ is normal in N,
(*ii*) S is a set of involutions which generate $W = N/(B \cap N)$,

(*iii*) sBw is a subset in $BuB \cup BswB$ for any $s \in S$ and $w \in W$,
(*iv*) $sBs \neq B$ for all $s \in S$.

Properties (1)-(iv) imply that (W, S) is a Coxeter system (see [16], [17]). Whenever (G, B, N, S) is a Tits system, we call the group W the Weyl group of the system, or more usually the Weyl group of G. The subgroups P_i of G defined by BW_iB are called the *standard maximal parabolic subgroups* of G. The group incidence system $\Gamma_G = \Gamma(G, P_i)_{1 \leq ilel}$ is commonly referred to as the *Lie geometry* of G (see [17]). Note that the Lie geometry of G and the Coxeter geometry of the corresponding Weyl group have the same rank. In fact there is a type preserving morphism from Γ_G onto Γ_W given by $gP_i \to wW_i$, where w is determined from the equality $BgP_i = BwP_i$. This morphism is called a *retraction* (see [98]).

3.8.2. Tits and Schubert automata for symbolic computations

The geometry $\Gamma(G)$ of BN-pair G is the set of all left cosets by the standard maximal subgroups i.e. maximal subgroups P_i, $i = 1, 2, \ldots, n$i of G containing standard Borel subgroup B. Two cosets $C_1 = gP_i$ and $C_2 = hP_j$ are incident C_1IC_2 if and only if their intersection is not empty. It is clear, that $gP_i \cap hP_j \neq 0$ implies $i \neq j$. The maximal flag of the geometry is a subset $F = \{C_1, C_2, \ldots, C_n\}$ such that C_iIC_j for each pair (i, j), $i \neq j$. Maximal flags form the set $\mathrm{F}\Gamma(G)$, they are in one to one correspondence with the left cosets by standard Borel subgrop. The largest Schubert cell Sch is the orbit of B acting on $\mathrm{F}\Gamma(G)$ containing largest number of elements. In case of group of normal type variety $\mathrm{Sch} = \mathrm{Sch}(G)$ is isomorphic to vector space ${F_q}^N$, where N is the number of positive roots.

We assume that two maximal flags F_1 and F_2 are adjacent if their intersection contains $n - 1$ elements of geometry. Let $AF(G)$ be the simple graph of symmetric adjacency relation (flag graph for $\Gamma(G)$. The order of this simple regular graph is $|(G : B)|$, the degree is nq and diameter is n. Let us restrict the adjacency relation as above on the largest Schubert cell $\mathrm{Sch}(G)$. We obtain new graph $AS(G)$ which is a regular induced subgraph of $AF(G)$ of order q^N and degree $q - 1$. We refer to $AS(G)$ as Schubert subgraph of the flag graph.

We convert the directed graph of adjacency relation of flags into the following automaton.

Let (F_1, F_2) be the ordered pair of adjacency flags such that $t(F_1 \cap F_2) = \{1, 2, \ldots, n\} - \{s\}$. So flags differs by geometry elements $C_1 = {C_s}^1$ and $C_2 = {C_s}^2$ of type s from (F_1, F_2), respectively. The following situations are possible.

(*i*) Element C_1 and C_2 are from the same Schubert cell. In that case there unique a transformation $u = x_\alpha(t)$, $t \neq 0$, shifting C_1 to C_2. Root α depends on $\mathrm{Retr}(F_1)$ only.

(*ii*) Elements C_1 and C_2 are from different Schubert cells and there is a group U_α such that $(F_1 \cap F_2) \cup \{u(C_2)\}$ is an adjacent flag to F_1 for each $u = x_\alpha(t)$. Notice, that case $t = 0$ is a possibility here. Root α depends on $\mathrm{Retr}(F_1)$ again.

(*iii*) Elements C_1 and C_2 are from different Schubert cells and Schubert cell contains C_2 as unique representative C such that flag $(F_1 \cap F_2) \cup \{C\}$ is adjacent to F_1.

Let us consider the following labelling of $F_1 \to F_2$ for cases of (i), (ii) and (iii) separately:

(*i*) put the label (s, t). where $t \neq 0$.

(*ii*) the label is (s, t), where $t \in F_q$ is defined by condition

$$x_\alpha(t)\mathrm{Retr}(C_2) = C_2$$

(*iii*) put the label ∞.

So for fixed F_1 and fixed type s the label (s, t) in direction to s-adjacency flag is defined by parameter t taken from the "acceptable" set $\mathrm{Ac}(F_1) = F_q \cup \{\gamma\}$ where γ is one of the symbols 0 and ∞. We add the formal loop on state F_1 labelled by the unique symbol from $\{0, \infty\} - \{\gamma\}$.

So the transition function $T_{s,t}$ of taking the s-adjacent element of colour (s, t) for general flag is defined for each $t \in F_q \cup \{\infty\}$ We assume that the initial state can be any flag from the largest Schubert cell Sch and this cell is the totality of all accepting states.

So algorithm can be given by the string of labels (s_1, t_1), (s_2, t_2), ..., (s_d, t_d) such that the composition $T = T(s_1, t_1)T(s_2, t_2)T(s_d, t_d)$ maps Sch into itself. We are interested only in irreducible computations for which $s_i \neq s_{i+1}$ for $i = 1.2, \ldots d - 1$

In case of group of normal type the alphabet contains exactly $n(q + 1)$ symbols. The computation corresponds to special walks in the graph $AF(G)$ with the starting and ending point in $\mathrm{Sch}(G)$. Notice that C may be not a bijection. For instance $T(s, O)$, which image for Sch lays outside of the largest large Schubert cell, is not invertible.

We refer to such automaton as *Tits automaton* for group G. We would like to use it as tool for symbolic computations.

The unipotent group U acts regularly on Sch. So we cam identify $v \in \mathrm{Sch}$ with certain product of $X_\alpha(t_\alpha)$, and positive roots $\alpha \in \mathrm{Root}$ are taken in Bruhat order. In fact, we identify the string $\mathrm{v} = t_\alpha \in F_q$, $\alpha \in \mathrm{Root}^+$ with the accepting state v.

We refer to the list $(t_{\alpha_1}, t_{\alpha_2}, \ldots, t_{\alpha_n})$, where $\alpha_1, \alpha_2, \ldots, \alpha_n$ is the set of all simple roots, as the color of v from plainspace. So we are colouring accepting states now but not arrows.

Let us consider irreducible computation within Tits automaton of kind $v \to v_s$, $v_1 = T(i_1, a_1)(v)$, $v_2 = T(i_2, a_2)(v_1), \ldots, v_s = T(i_s, a_s)(v_{s-1})$, where $i_k \neq i_{k+1}$, $k = 1, \ldots, s-1$, $a_k \in F_q \cup \infty$, element $\mathrm{Retr}(v) = \mathrm{Retr}(v_s)$ equals to the element $w \in W$ of maximal length. Notice, that in the sequence $\mathrm{Retr}(v_1), \mathrm{Retr}(v_2), \ldots, \mathrm{Retr}(v_k)$ consecutive elements are adjacent in $\mathrm{F\Gamma}(W)$ or equal.

The computation is conducted into several steps. Each time we have one of the situations i, (ii) or (iii). In cases of kind (i) and (ii) when the corresponding root α is simple parameters a_j will be chosen as linear functions of kind $l(t_{\alpha_1}, t_{\alpha_2}, \ldots t_{\alpha_n}) = c_1 t_{\alpha_1} + c_2 t\alpha_2 \ldots, c_n t_{\alpha_n} + b$, where $c_1, c_2, \ldots, c_n$ and b are elements of F_q and $(t_{\alpha_1}, t_{\alpha_2}, \ldots, t_{\alpha_n})$ is a colour of our initial state. If α is not a simple root, we choose a_j as $c_j t_{\beta_j} + f_j((t_{\alpha_1}, t_{\alpha_2}, \ldots t_{\alpha_n})$, where $c_j \neq 0$.

After the completion of our computation we get the accepting state $u = v_s$. It has a colour $(d_{\alpha_1}, d_{\alpha_2}, \ldots, d_{\alpha_n}) = (t_{\alpha_1}, t_{\alpha_2}, \ldots t_{\alpha_n})A + (b_1, b_2, \ldots, b_n)$, where the matrix A is defined by some linear expressions of kind

$a_i = l_i(l(t_{\alpha_1}, t_{\alpha_2}, \ldots t_{\alpha_n})$, which we used during the computation. We will require that the matrix A is invertible. Notice that we may use symbol ∞, where the design of algorithm allows such option.

After the completion of algorithm we obtain accepting state of colour $(d_{\alpha_1}, d_{\alpha_2}, \ldots, d_{\alpha_n})$. The invertibility of A allows us to compute

$(t_{\alpha_1}, t_{\alpha_2}, \ldots t_{\alpha_n})$ as $((d_{\alpha_1}, d_{\alpha_2}, \ldots d_{\alpha_n}) - (b_1, b_2, \ldots, b_n))A^{-1}$. So we can compute all parameters a_i and create the reverse walk in the graph and compute the inverse map T^{-1} which sends the final accepting state to initial state.

Let us restrict Tits automaton on the largest Schubert cell, i. e delete all states outside $\mathrm{Sch}(G)$ together with corresponding output arrows. We obtain Schubert automaton over the alphabet (i, a), where $a \in F_q$, $1 \leq i \leq n$. Notice, that $a = 0$ corresponds to taking the loop.

3.8.3. Tits and Schubert automata and related symmetric encryption

Correspondents Alicia and Bob may use the following symmetric encryption based on the Tits automaton. The plainspace is a vector space $\mathrm{Sch} = {F_q}^N$. The plaintext p we identify with the string v= $t_\alpha \in F_q$, $\alpha \in \mathrm{Root}^+$. We may think that this is a function p : $Root^+ \to F_q$. Alicia has to compute the restriction of this function onto subsets of all simple roots and get the colour $(t_{\alpha_1}, t_{\alpha_2}, \ldots, t_{\alpha_d})$ of the plainspace.

Correspondents share symbolic string of labels $(s_1, l_1), (s_2, l_2), \ldots, (s_d, l_d)$, where $l_i.i = 1, 2, \ldots, d$ is a linear expression of formal variables z_α, for each simple root α or ∞ and two affine invertible transformations τ_1 and τ_2. The vector space of all maps from the totality of simple roots to F_q has to be not invariant subspace for τ_i, $i = 1, 2$. Alice executing the specialization $z_\alpha = p_\alpha$ computing Corresponding numerical string t $= (t_1, t_2, \ldots, t_d)$. She has to hide that string by applications of affine maps τ_i. So she is adding to symbolic key two invertible Linear transformations τ_1 and τ_2 of the plainspace ${F_q}^N$ and compose τ_1, the automaton map corresponding to t and τ_2.

She sends to Bob the ciphertext

$$\mathrm{c} = \tau_1(T(s_1, t_1)T(s_2, t_2)\ldots T(s_d, t_d)(\tau_2(\mathrm{p}))$$

Bob decrypt applying to c consequently ${\tau_2}^{-1}$, T^{-1}, where

$$T = T(s_1, t_1)T(s_2, t_2)\ldots T(s_d, t_d)$$

and ${\tau_1}^{-1}$.

Remark 1. If correspondents do not use ∞ in the shared symbolic key then T is the computation in Schubert automaton. Bob can simply compute T^{-1} as $T(s_d, -t_d)T(s_{d-1}, -t_{d-1})\ldots T(s_1, -t_1)$.

Remark 2. We may generalise the above algorithms by changing affine maps τ_1, τ_2 and $(t_1, t_2, \ldots, t_n) \to (t_1, t_2, \ldots, t_d)A + (b_1, b_2, \ldots, b_n)$ for general invertible polynomial maps.

3.8.4. Key exchange protocols based on incidence geometries

The automata as above can be considered over the general ground field F We can see that the computations in both automata do not use division. What is going on during the computations on a symbolic level. Let us assume now that the initial state is a formal string of variables x_α, where α is running through the list of all positive roots. It is convenient for us to expand the ground field F_q to the field R of rational functions $r(x_1, x_2, \ldots, x_N) = f(x_1, x_2, \ldots, x_N)/g(x_1, x_2, \ldots, x_N)$, where f and g are elements $F_q[x_1, x_2, \ldots, x_N]$. Formal variables x_α and governing linear expressions $l(x_{\alpha_1}, x_{\alpha_2}, \ldots, x_{\alpha_n}, x_\alpha)$, where α is not a simple root are elements of subring $F_q[x_1, x_2, \ldots, x_N]$ in R. During its work Tits automaton newer use division. So after getting accepting state over R we got the vector of dimension N with polynomial components f_α. So the numerical encryption map is regular automorphism of ${F_q}^N$ (element of Cremona group for ${F_q}^N$) of kind.

$$x_i \to f_i(x_1, x_2, \ldots, x_N), i = 1, 2, \ldots, N$$

Special choice of symbolic key guarantee that the above transformation is bijective. Symbol ∞ play just formal role. Linearity of governing functions leads to rather small degree of the nonlinear map.

Such a walk produces a bijective transformation T of variety $\mathrm{Sch}(G)$ which is its regular automorphism (polynomial map of the variety into itself such that its inverse is also polynomial). We will conjugate T by invertible affine transformation $\tau \in AGL_N(F_q)$ and use $Y = \tau^{-1}T\tau$ as the instrument for the key exchange based in modified Diffie - Hellman method. So the Alice is computing a standard form for Y

$$t_1 = f_1(t_1, t_2, \ldots, t_N), t_2 = f_2(t_1, t_2, \ldots, t_N), \ldots, t_N = f_N(t_1, t_2, \ldots, t_N),$$

where $f_i \in F_q[t_1, t_2, \ldots, t_N]$, $i = 1, 2, \ldots, N$, and sending the map to Bob via open communication channel. Correspondents Alice and Bob (as usually) are choosing their keys k_A and k_B, respectively. They are executing computations $D_A = Y^{k_A}$ and $D_B = Y^{k_B}$. They exchange the outputs via the open channel.

Finally Alice and Bob are computing collision maps ${D_B}^{k_A}$) and ${D_A}^{k_B}$. So correspondents are getting common element.

We can modify the above scheme:

Alice chooses the maximal flag F from the largest large Schubert cell $\mathrm{Sch}(G)$ and sends it to Bob via open channel. Correspondence may use common flag ${D_A}^{k_B}(F) = {D_B}^{k_A}(F)$ as the key for their private key algorithm.

The security of the above key exchange algorithm based on the complexity of discrete logarithm problem for the Cremona group of variety $\mathrm{Sch}(G)$. In case of finite field F_q this group coincides with the symmetric group S_{q^N}. it is important that we use description of permutations in terms of polynomial algebra. So related discrete logarithm problem is formulated in terms of algebraic geometry.

Method allows various modification: we can use nonlinear invertible maps instead of affine transformation τ, the base of discrete logarithm can be non invertible polynomial map and etc. An interesting modifications can be obtained if we will allow not bijective transformations of the variety. For instance we may consider fractional linear governing function l_i for the step i looks like $(a_1x_{\alpha_1} + a_2x_{\alpha_2} + \ldots a_{\alpha_n}x_{\alpha_n})/(b_1x_{\alpha_1} + b_2x_{\alpha_2} + \ldots b_{\alpha_n}x_{\alpha_n})$ if the root α on step i is simple, and l_i is a fraction of two linear combinations of x_{α_i}, if α is not a simple root. In case of such governing functions we refer to corresponding automata as birational Tits and Schubert automaton, respectively.

3.8.5. Embedding of the flag variety into the Lie Algebra and some complexity estimates

Throughout this section (G, B, N, S) is a Tits system which arises in connection with Chevalley group G, although we point that the results of this section remain valid in a far more general setting (see [35],[20], [108]). We write $G = X_l(K)$ to signify that G is the Chevalley group over the field K, with associated Dynkin diagram X_l. We are most interested in the case when K is finite, and we shall write $X_l(q)$ instead of $X_l(F_q)$ in that case.

So, fix Chevalley group $G = X_l(K)$ with corresponding Weyl group W. As in the previous section Γ_W and Γ_G their associated Coxeter and Lie geometries. Let $L = H + L^+ + L^-$ be the Lie algebra corresponding to G.

Following convention, we refer to H, L^+, L^- and $H+L^+$ as, respectively, the *Cartan subalgebras, positive root space, negative root space* and *Borel subalgebra* with respect to the given decomposition of L. We also use the familiar bracket notation $[,]$ to indicate Lie product [16], [87],

Below we turn out our attention to a method of embedding Γ_W and Γ_G in L. As the reader shall see, this method actually embeds Γ_W in the Cartan subalgebra H of L. Let us consider the embedding more precisely.

Let $A = (a_{i,j})$ be the Cartan matrix corresponding to the root system Ω of W. We consider the lattice R which is generated by simple roots $\alpha_1, \alpha_2, \ldots, \alpha_l$ and the reflection $r_1, r_2, \ldots r_l$ of R defined by the equality $(\alpha_i)^{r_j} = \alpha_i - a_{i,j}\alpha_j$.

Let $S = \{r_1, r_2, \ldots, r_l\}$ is the set of Coxeter generators of Weyl group W. Let ${\alpha_1}^*, {\alpha_2}^*, \ldots, {\alpha_l}^*$ be a dual basis of $\alpha_1, \alpha_2, \ldots, \alpha_l$, i.e. ${\alpha_i}^*$ is the linear functional on R which satisfies ${\alpha_i}^*(\alpha_j) = \delta_{i,j}$. We define the action of W on the dual lattice R^* by $l(x)^s = l(x^s)$, where $l(x) \in \mathrm{R}^*$ and $s \in S$.

Consider the orbit $H_i = \{w({\alpha_i}^*) | w \in W\}$ of permutation group (W, R^*), which contains ${\alpha_i}^*$. Let H be the disjoint union of H_i. We give the set H the structure of an incidence system as follows. Linear functionals $l_1(x)$ and $l_2(x)$ are incident if and only if products $l_1(\alpha)l_2(\alpha) \geq 0$ for all $\alpha \in \Omega$. The type function t is defined by $t(l(x)) = i$ where $l(x) \in H_i$. It can be shown that (H, I, t) is isomorphic to Coxeter geometry Γ_W. (In fact there is a unique isomorphism of Γ_W with (H, I, t) which sends W_i to α_i, $1 \leq i \leq l$.) This gives the desired embedding since H is a subset in R^* and $\mathrm{R}^* \subset L_0$. Moreover this embedding still valid for a field K of sufficiently large characteristic, since, in that case H is a subset of $\mathrm{R} \times K = L_0$.

We now consider an analogous embedding of the Lie geometry Γ_G into the Borel subalgebra $U = L_0 + L^+$ of L. Let $d = {\alpha_1}^* + {\alpha_2}^* + \ldots {\alpha_l}^*$. Than we can take $\Omega^+ = \alpha \in \Omega | d(\alpha) \geq 0$ to be our set of positive roots in Ω. For any $l(x) \in \mathrm{R}^*$ define $\eta^-(L) = \alpha \in \Omega^+ | l(\alpha) < 0$.

Let L_α be the root space corresponding to positive root α. For each

$h \in H$ we define the subalgebra L_h as the sum of L_α, $\alpha \in \eta^-(h)$. Let $U_i = \{h + v | h \in H_i, v \in L_h\}$ and U is a disjoint union of U_i. We give U the structure of an incident system as follows. Elements $h_1 + v_1$ and $h_2 + v_2$ are incident if and only if each of the following hold:

(i) $h_1(\alpha)h_2(\alpha) \geq 0$ for all $\alpha \in \Omega$, i.e. h_1 and h_2 are incident in (H, I, t).
(iii) $[h_1 + v_1, h_2 + v_2] = 0$

Element $h + v$ has type i if $h + v \in U_i$.

In [105] it is shown that this newly defined incident system is isomorphic to the Lie geometry Γ_G, provided that the characteristic of K is zero or sufficiently large to ensure the isomorphism at the level of the subgeometries (H, I, t) and Γ_W. Then analogous to the Weyl case, there exists a unique isomorphism Retr of $\Gamma(G)$ into (U, I, t) which sends P_i to α_i, $1 \leq i \leq l$.

Proposition 25. *Let $\Gamma = \Gamma(G)$ be the geometry of group $G = X_n(q)$. The above interpretation of $\Gamma(G)$ allows*

- (i) *generate Γ in $O(|\Gamma|)$ elementary steps and check whether or not two elements of Γ are incident for time $O(N^2)$, where N is the number of positive roots.*
- (ii) *complete the computation in Tits and Schubert automaton consisting of k elementary steps for time $O(kN)$*

Graphs of degree q and $SF(X_n(q)$, $q \geq 4$ of degree $q - 1$ have orders $|X_n(q)|/|B|$ and q^N, respectively. They form families of small world graphs depending on two parameters n and q.

3.8.6. On the discrete logarithm problem with polynomial or birational base

Let F_p, where p is prime. be a finite field. Affine transformations x $\rightarrow$ Ax + b, where A is invertible matrix and $b \in (F_p)^n$, form an affine group $AGL_n(F_p)$ acting on ${F_p}^n$. It is known that polynomial transformation of kind $x_1 \rightarrow g_1(x_1, x_2, \ldots, x_n), x_2 \rightarrow g_2(x_1, x_2, \ldots, x_n), \ldots, x_n \rightarrow g_n(x_1, x_2, \ldots, x_n)$ form a symmetric group S_{p^n}.

In the simplest case F_p, affine transformations form an affine group $AGL_n(F_p)$ of order $(p^n - 1)(p^n - p) \ldots (p^n - p^{n-1})$ in the symmetric group S_{p^n} of order $(p^n)!$. In [76] the maximality of $AGL_n(F_p)$ in S_{p^n} was proven. So we can present each permutation π as a composition of several "seed" maps of kind $\tau_1 g \tau_2$, where $\tau_1, \tau_2 \in AGL_n(F_p)$ and g is a fixed map of degree ≥ 2. One may choose quadratic map of Imai - Matsumoto algorithm in case $p = 2$ (see [53], [80] for its description and cryptanalysis by J. Patarin) or graph based cubical maps [128] for general p.

We can choose the base of ${F_p}^n$ and write each permutation $g \in S_{p^n}$ as a "public rule":

$$
\begin{aligned}
&x_1 \to g_1(x_1, x_2, \ldots, x_n),\\
&x_2 \to g_2(x_1, x_2, \ldots, x_n),\\
&\ldots,\\
&x_n \to g_n(x_1, x_2, \ldots, x_n).
\end{aligned}
$$

Let $g^k \in S_{p^n}$ be the new public rule obtained via iteration of g. Discrete logarithm problem of finding solution for k for $g^k = b$ can be difficult if the order of g is "sufficiently large". We have to avoid the linear growth of the degree g^k, when k is growing. Obvious bad example is the following: g sends x_i into ${x_i}^t$ for each i. In this case the solution is just a ratio of deg$b$ and deg$g$.

Let us consider the Cremona group $C(n, q)$ of all invertible polynomial automorphisms of the vector space ${F_q}^n$, where $q = p^m$, the semigroups $PC(n, q)$ and $BC(n, q)$ of polynomial and birational maps of ${F_q}^n$ into itself, respectively.

To avoid such trouble one can look at families of subgroups of increasing orger G_n, $n \to \infty$ of S_{p^n} such that maximal degree of its element equals c, where c is independent constant (groups of degree c or groups of stable degree). We refer to an element g such that all its nonidentical powers are of degree c as element of stable degree.

It is clear that the family of affine subgroup $AGL_n(p)$ is a subgroup of stable degree for $c = 1$ and all nonidentical affine transformations are of stable degree. Notice that if g is a linear diagonalisable element of $AGL_n(p)$, then discrete logarithm problem for base g is equivalent to the classical number theoretical problem.

One can take a subgroup H of $AGL_n(p)$ and consider its conjugation with nonlinear bijective polynomial map f. Of course the group $H' = f^{-1}Hf$ will be also a stable group, but for most pairs f and H group H' will be of degree degf $\times$ deg$f^{-1} \geq 4$ because of nonlinearity f and f^{-1}. So the problem of construction an infinite families of subgroups G_n in S_p^n of degree 2 and 3 may attract some attention.

The following questions are important because of Diffie Hellman type protocols (see [25]).

$Q1$; How to construct stable subgroups C of small degree c ($c = 2$ and $c = 3$ especially) of increasing order in $C(n, q)$?

We say refer to a semigroup Se generated by single elements as monogenetic semigroup of order |Se|.

$Q2$; How to construct stable monogenetical subsemigroups in $PC(n,q)$ and $BC(n,q)$ of small degree c ($c=2$ and $c=3$ especially) of increasing order in $C(n,q)$ of large order?

Finally, we announce the following statement

Theorem 19. *Let $X_n(F)$, $n \geq 2$ be a simple group of Lie type over the field F. Let $L(X_n(q)$ be a group of all invertible computations in Schubert automaton.*

In case of classical groups (diagrams A_n, B_n, C_n and D_n) groups $L(X_n(F))$, $n \to \infty$ form families of stable degree.

Remark. Groups $L(X_n(F))$ are of degree 3 in case of diagram B_n, C_n and D_n, and $L(A_n(F))$ are groups of degree 2.

We can demonstrate the existence of elements in $L(X_n(q))$ of rather large order. Really, take a permutation $i_1, i_2, \ldots, i_n$ on the nodes of Dynkin diagram and compute a composition g of generators $Z^{i_1}(l_1(x)), Z^{i_2}(l_2(x)), \ldots$

$Z^{i_n}(l_n(x))$, where $l_i(x)$ are linear forms corresponding to the rows of Singer cycle matrix of order $q^n - 1$ (see, for instance, [36]). As it follows from the description of algorithm the order of g will be at least $q^n - 1$.

Similarly we can use Singer cycle to generate by Tits automata a stable monogenetic subgroup in $PC(n,q)$ and $BC(n,q)$.

Chapter 4

On the directed graphs without commutative diagrams, related encryption automata and optimistion problems

4.1. Directed graphs and related automata

The missing theoretical definitions on directed graphs the reader can find on [114]. Let Φ be an irreflexive binary relation over the set V, i.e. $\Phi \in V \times V$ and for each v pair (v, v) is not the element of Φ.

We say that u is the neighbour of v and write $v \to u$ if $(v, u) \in \Phi$. We use term *balanced binary relation graph* for the graph Γ of irreflexive binary relation ϕ over finite set V such that for each $v \in V$ sets $\{x|(x, v) \in \phi\}$ and $\{x|(v, x) \in \phi\}$ have same cardinality. It is a directed graph without loops and multiple edges. We say that graph Γ is k-regular if for each vertex $v \in \Gamma$ the cardinality of $\{x|(v, x) \in \phi\}$ is k.

Let Γ be the graph of binary relation. The *pass* between vertices a and b is the sequence $a = x_0 \to x_1 \to \dots x_s = b$ of length s, where x_i, $i = 0, 1, \dots s$ are distinct vertices.

We say that the pair of passes $a = x_0 \to x_1 \to \dots \to x_s = b$, $s \geq 1$ and $a = y_0 \to y_1 \to \dots \to y_t = b$, $t \geq 1$ form an (s, t)- commutative diagram $O_{s,t}$ if $x_i \neq y_j$ for $0 < i < s$, $0 < j < t$. Without loss of generality we assume that $s \geq t$.

We refer to the number $\max(s, t)$ as the rank of $O_{s,t}$. It is ≥ 2, because the graph does not contain multiple edges.

Notice, that the graph of antireflexive binary relation may have a directed cycle $O_s = O_{s,0}$: $v_0 \to v_1 \to \dots v_{s-1} \to v_0$, where v_i, $i = 0, 1, \dots, s - 1$, $s \geq 2$ are distinct vertices.

We will count directed cycles as commutative diagrams.

For the investigation of commutative diagrams we introduce *girth indicator* gi, which is the minimal value for $\max(s, t)$ for parameters s, t of commutative diagram $O_{s,t}$, $s + t \geq 3$. Notice, that two vertices v and u at distance $<$ gi are connected by unique pass from u to v of length $<$ gi.

In case of symmetric binary relation gi $= d$ implies that the girth of the graph is $2d$ or $2d - 1$. it does not contain even cycle $2d - 2$. In general case gi $= d$ implies that $g \geq d + 1$. So if the case of family of graphs with unbounded girth indicator, the girth is also is unbounded. We have also gi $\geq g/2$.

We assume that the *girth* $g(\Gamma)$ of directed graph Γ with the girth indicator $d+1$ is $2d+1$ if it contains commutative diagram $O_{d+1,d}$. If there are no such diagrams we assume that $g(\Gamma)$ is $2d + 2$.

In the case of symmetric irreflexive relations the above general definition of the girth agrees with the standard definition of the girth of simple graph i.e the length of its minimal cycle.

We will use term *the family of graphs of large girth* for the family of regular graphs Γ_i of degree k_i and order v_i such that gi(Γ_i) is $\geq c \log_{k_i} v_i$, where c' is the independent of i constant.

As it follows from the definition $g(\Gamma_i) \geq c' \log_{k_i}(v_i)$ for appropriate constant c'. So, it agrees with the well known definition for simple graphs.

4.1.1. Algebraic Graphs with special colouring of edges, general algorithm

We shall use term *the family of algebraic graphs* for the family of graphs $\Gamma(K)$, K belongs to some infinite class F of commutative rings, such that the neighbourhood of each vertex of $\Gamma(K)$ and the vertex set itself are quasiprojective varieties over K of dimension ≥ 1 (see [7] for the case of simple graphs).

Such a family can be treated as special Turing machine with the internal and external alphabet K.

We say that the graph Γ of binary relation Φ has a rainbow-like colouring ρ of edges over the set of colours C if for each arrow (u, v) of the graph $((u, v) \in \Phi)$ the colour $\rho(u, v) \in C$ is defined and the following properties hold:

(i) For each pair (u, c) such that u is a vertex and $c \in C$ there is a unique vertex $v = N_c(u)$ satisfying condition $\rho(u, v) = c$.

(ii) For each pair (v, c) such that v is a vertex and $c \in C$ there is a unique vertex $u' = N_c(v)$ satisfying condition $\rho(u, v) = c$.

We have $N'_c(N_c(u)) = u$. So operators N and N' are bijections on the set of vertices of Γ.

Let us consider the encryption algorithm corresponding to the graph Γ with the chosen invertible rainbow like colouring of edges.

Let $\rho(u, v)$ be the colour of arrow $u \to v$. The set C is the totality of colours and $N_c(u)$ is the operator of taking the neighbour of u with the colour c.

The password is the string of colours key $= (c_1, c_2, \ldots, c_s)$ and the encryption procedure is the composition $N_{c_1} \times N_{c_2} \ldots N_{c_s}$ of bijective maps $N_{c_i} : V(\Gamma) \to V(\Gamma)$. So if the plaintext $v \in V(\Gamma)$ is given, then the encryption procedure corresponds to the following chain in the graph: $x_0 = v \to x_1 = N_{c_1}(x_0) \to x_2 = N_{c_2}(x_1) \to \cdots \to x_s = N_{c_s}(x_{s-1}) = u$. The vertex u is the ciphertext.

The decryption procedure corresponds to the composition of maps $N'_{c'_s}$, $N'_{c'_{s-1}}, \ldots, N'_{c'_1}$. The above scheme gives a symmetric encryption algorithm with flexible length of the password (key). Let $A(\Gamma, \rho)$ be the above encryption scheme. The following statement is immediate corollary from definitions.

Lemma 16. *Let Γ be the invertible rainbow-like graph of girth g and $A = A(\Gamma, \rho)$ be the above encryption scheme with the length of password s,*

1, $s <$ (gi). *Then different passwords produce distinct ciphertexts, plaintext and corresponding ciphertext are different.*

Let $|C| \geq 3$ is finite. For the encryption algorithm A depending on the key key from the keyspace we consider the constants $dk = dk(A)$ which is the minimal number of different ciphertexts obtained from the same plaintext by application of all passwords from the keyspace (coefficient of direct key impact).

Let $\mathrm{mk}(A)$ be the cardinality of the minimal set containing all vertices which one can obtained from some vertex v by the compositions of the encryption transformations in $A(\Gamma, \rho)$. If the graph Γ is strongly connected then $\mathrm{mk}(A)$ is the minimal size of the connected component in Γ.

For the Turing machine T of block-cipher working with potentially infinite text both coefficients $\mathrm{dk}(T)$ and $\mathrm{mk}(T)$ are bounded by the size of block.

We say that the graph of binary relation is k-regular if each vertex has exactly k inputs and k outputs

Let Γ_n, $n = 1, 2, \ldots$ be an infinite family of rainbow-like k_n-regular graphs ($k_n \geq 3$) with the increasing girth indicator gi_n. Let us consider the Turing machine A corresponding to the sequence of $A_n = A(\Gamma_n, \rho_n)$, $n = 1, 2, \ldots$, It is clear that $\mathrm{dk}(A_n)$ is the sum of $k(k-1)^{s_i}$, $s_i \leq \mathrm{gi}_n$. So this is unbounded function and the Turing machine is not a block cipher.

Notice that for simple k-regular graphs Γ_n of the girth g_n and order v_n, $k \geq 3$ and $n = 1, 2, \ldots$ the parameter $\mathrm{dk}(n) = \mathrm{dk}(A_n)$ is the minimal number of vertices at a distance d, $1 < d \leq g_n$ from the chosen vertex. The fastest possible grow of $\mathrm{dk}(n)$ will be in case of the family of *graphs of large girth* when $g_i \geq c\log(v_i)$ for the constant $0 < c < 2$ (see [8]). Parameter $\mathrm{mk}(n) = \mathrm{mk}(A_n)$ is the minimal order of the connected component of Γ_n. If each Γ_n is connected then $\mathrm{mk}(n) = v_n$.

The fastest possible grow of $\mathrm{mk}(n)$ will be in the case of *small world graphs* i.e graphs with the diameter $O(log_{k-1}(v_n))$ (see [12]).

We will use term the folder of graphs for the family Γ_n of k-regular rainbow-like graphs for which there is a colour preserving graph homomorphism from Γ_n onto Γ_{n-1}. For such a family there is well define projective limit which is an infinite k-regular graph. It corresponds to graph based Turing machine. In case of the folder of connected k-regular graphs of increasing girth the projective limit is k-regular tree. The example of a folder has been considered in the section 5.

GENERAL ALGORITHM.

Let us assume that all members of the above family Γ_n of raibow like graphs over the set of colours C_n are strongly connected algebraic graphs over the commutative ring K (finite or infinite) and the vertex set $V_n(K)$ is an open variety in Zarissky topology. Let us chose to biregular automorphisms τ_1 and τ_2 i.e. polynomial bijections on $V_n(K)$ such that their inverses are polynomial maps also. Let f_i, $1 \leq i \leq m(n)$ be the sequence of functions which are invariant on connected components of Γ_n: for v and u from the same connected component and each i we have $f_i(u) = f_i(v)$. Let v be the vertex of Γ_n (the plaintext. We can take the *symbolic key* $(F_1, F_2, \ldots, F_l(n))$, $l(n) \leq \text{gi}_n$ formed by elements of $K[y_1, \ldots, y_{m(n)}]$, compute the *numerical key* $\text{k} = (k_1, k_2, \ldots, k_{l(n)})$, where $k_i = F_i(f_1(v), f_2(v), \ldots, f_{m(n)}(v))$, $i = 1, 2, \ldots, l(n)$, create the transformation $T_\text{k} = N_{k_1} \times N_{k_2} \times \ldots N_{l(n)}$ and take the composition $E(\tau_1, \tau_2, F_1, F_2, \ldots F_{l(n)})$ of τ_1, T_k and τ_2 as encryption map on $V_n(K)$.

4.1.2. Some examples

Example 1: Cayley graphs

Let G be the group and S be subset of distinct generators, then the binary relation $\phi = \{(g_1, g_2) | g_i \in G, i = 1, 2, g_1 {g_2}^{-1} \in S$: admit the rainbow like colouring $\rho(g_1, g_2) = g_1 {g_2}^{-1}$

This rainbow like colouring is invertible because the inverse graph $\phi^{-1} = \{(g_2, g_1) | g_1 {g_2}^{-1} \in S\}$ admit the rainbow-like colouring $\rho^{'}(g_2, g_1) = g_2 {g_1}^{-1} \in S^{-1}$.

The first explicit examples of families with large girth were given by Margulis [69], with for some infinite families with arbitrary large valency. The constructions were Cayley graphs $X^{p,q}$ of group $SL_2(Z_q)$ with respect to special sets of $p + 1$ generators, p and q are primes congruent to 1 mod4. Then independently Margulis and Lubotsky, Phillips, and Sarnak [56] proved that for each p the constant γ for graphs $X^{p,q}$ with fixed p is $\geq 4/3$. Later on Biggs and Boshier showed that this γ is asymptotically $4/3$.

The family of $X^{p,q}$ is not a family of algebraic graphs because the neighbourhood of each vertex is not an algebraic variety over F_q. For each p, graphs $X^{p,q}$, where q is running via appropriate primes, form a family of small world graph of unbounded diameter. We give a brief outline of the explicit construction of of a class of Cayley graphs called *the Ramanujan Graph* $X(p, q)$ due to Lubotzky, Phillips and Sarnak [66].

Let p and q be primes, $p \equiv q \equiv 1(\text{mod } 4)$. Suppose that i be an integer so that $i^2 \equiv -1(\text{mod } q)$. By a classical formula of Jacobi, we know that there are $8 * (p+1)$ solutions $\alpha = (a_0, a_1, a_2, a_3)$ such that $a_0^2 + a_1^2 + a_2^2 + a_3^2 = p$. Among these, there are exactly $p+1$ with $a_0 > 0$ and a_0 odd and a_j even for $j \in \{1,\ 2,\ 3\}$, as is easily shown. To each such α we associate the matrix

$$\tilde{\alpha} = \begin{pmatrix} a_0 + ia_1 & a_2 + ia_3 \\ -a_2 + ia_3 & a_0 - ia_1 \end{pmatrix}$$

which gives us $p+1$ matrices in $\mathrm{PGL}_2(\mathbb{F}_q)$. We let S be the set generators of these matrices $\tilde{\alpha}$ and take $\mathrm{PGL}_2(\mathbb{F}_q)$. In [6], it is shown that the Cayley graphs $X(p,q)$ will be a $(p+1)$-regular graph, namely the Cayley graph of $\mathrm{PSL}_2(\mathbb{F}_q)$ if $\left(\frac{p}{q}\right) = 1$ and $\mathrm{PGL}_2(\mathbb{F}_q)$ if $\left(\frac{p}{q}\right) = -1$, (where $\left(\frac{p}{q}\right)$ is the Legendre symbol). As we vary q, we get an infinite family of such graphs, all $p+1$-regular.

Moreover, in papers written by Lubotzky, Phillips and Sarnak [6], an explicit formula for the girth $g(X(p,q))$ of $X(p,q)$ was found.

Corollary 7 (6)**.** Following cases draw ahead:

(1) If $\left(\frac{p}{q}\right) = -1$ then $X(p,q)$ is bipartitle of order $n = |X(p,q)| = q(q^2-1)$ and

$$g(X(p,q) \geq 4 \log_p q - \log_p 4$$

(2) If $\left(\frac{p}{q}\right) = 1$ then $X(p,q)$ is not bipartitle, $n = |X(p,q)| = q(q^2-1)/2$ and

$$g(X(p,q) \geq 4 \log_p q.$$

Above the Corollary 1 shows that the Ramanujan graph $X(p,q)$ of order n which asymptotically satisfies $g(X(p,q)) \geq 4 \log_{k-1} n/3$.

Here we can use the Ramanujan graphs to generating of matrices with large order. The algorithm is:

Algorithm 1. Let $X(p,q)$ be the Ramanujan graphs. Then

(1) Take the product

$$g = s_{i_1} s_{i_2} \dots s_{i_k}, \qquad (k - \text{small}, \quad s_{i_j} \in S, j = 1, \dots, k)$$

and $s_{i_1} \neq s_{i_k}^{-1}$. (We can chose the sequence $s_{i_1} s_{i_2} \dots s_{i_k}$ in a way that g is not similar to diagonal matrix or matrix of kind $\begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix}$)

(2) Then the order $|g|$ of g is such that

$$|g| \geq \frac{g\,(X\,(p,q))}{k},$$

where $(g\,(X\,(p,q))$-girth of the Ramanujan graph $X\,(p,q))$.

Remark. *The girth of the Ramanujan graph $X(p,q)$ is unbounded (p is fixed, $q \to \infty$). So we can choose a large q to make $|g|$ as large as we want.*

Troubleshooting for $X^{p,q}$:

1) For work with large plaintext we need large prime number, so real option to work with large texts is restricted.

2) As we write before, the important feature of general graph based encryption is the resistance to attacks, when adversary intercepts the pair plaintext - ciphertext, because the best algorithm of finding the pass between given vertices (by Dijkstra, see [26] and latest modifications) has complexity $v\ln v$ where v is the order of the graph, i.e. size of the plainspace. But in case of concrete family of graphs one can find a way to compute key faster. The class of Cayley graphs gives us clear example:

Let us assume that Γ is large Cayley graph corresponding to group G and the set of generators S. Let the password is determined by sequence of generators $s_1, s_2, \dots s_k$, the plaintext is $g \in G$ and the corresponding ciphertext is $h \in G$. Let us assume that adversary has the pair (g,h). The problem of getting the sequence $s_1, \dots, s_k$ can be difficult, but the adversary can compute element $l = g^{-1}h$ and use its inverse to control the communication channel under the condition that correspondents are not changing the password. Really, $c = l^{-1}p$.

To prevent such trouble the correspondences can modify such encryption scheme by "hiding the graphs up to isomorphism", i.e. instead of the sequence of k-regular (k is fixed) graphs Γ_i they can take the graphs of binary relations $\pi_i(\Gamma_i) = \{(u,v)|(\pi_i(v), \pi_i(u)) \in \Gamma\}$, where π_i be some bijection on $V(\Gamma_i)$. It is clear that parameters dk, mk) and the girth are same for graphs $\pi_i(\Gamma)_i)$ and Γ_i are different. It is important to take sparse π_i, which can be computed for linear time, do not slow down the computation seriously. Experiment show that the mixing properties of algorithms corresponding $\pi_i(\Gamma_i))$ and Γ_i can be very different.

P.Luks estimated that the complexity of mass problem of finding isomorphism between k-regular Γ_i and $\pi_i(\Gamma_i)$ is polynomial expression in variable $v_i = |V(\Gamma_i)|$ which is the size of the plainspace. So encryption algorithm for $\pi_i(X^{p,q})$ may have very good resistance to attack of type (ii). Notice that here we are treating π_i as a part of password.

Example 2: Parallelotopic graphs and latin squares

Let G be the graph with the colouring $\mu : V(G) \to C$ of the set of vertices $V(G)$ into colours from C such that the neighbourhood of each vertex looks like rainbow, i.e. consists of $|C|$ vertices of different colours. In case of pair

(G, μ) we shall refer to G as *parallelotopic graph* with the local projection μ.

It is obvious that parallelotopic graphs are k-regular with $k = |C|$. If C' is a subset of C, then induced subgraph $G^{C'}$ of G which consists of all vertices with colours from C' is also a parallelotopic graph. It is clear that connected component of the parallelotopic graph is also a parallelotopic graph.

The *arc* of the graph G is a sequence of vertices $v_1, \ldots, v_k$ such that $v_i I v_{i+1}$ for $i = 1, \ldots, k-1$ and $v_i \neq v_{i+2}$ for $i = 1, \ldots, k-2$. If $v_1, \ldots, v_k$ is an arc of the parallelotopic graph (G, μ) then $\mu(v_i) \neq \mu(v_{i+2})$ for $i = 1, \ldots, k-2$.

Various examples of simple parallelotopic graphs have been considered in previous section. The implementation for algorithm based on different from $\phi(n)$ family of graphs of large girth is discussed in [113].

Let $+$ be the latin square defined on the set of colours C. Let us assume $\rho(u, v) = \mu(u) - \mu(v)$. The operator $N_c(u)$ of taking the neighbour of the color is invertible, ${N_c}^{-1} = N_{-c}$, where $-c$ is the opposite for c element in the latin square. It means that ρ is invertible rainbow like colouring.

4.2. On extremal graph theory for balansed directed graphs

Recall that according to the Bourbaki the graph (or directed graph) is the pair V (vertex set) and subset ϕ in the Cartesian product $V \times V$. We refer to element $v \in V$ as vertex (state in automata theory).

We use term arc (or arrow as in automata theory) for the element $(a, b) \in \Phi$. We refer to $(a, b) \in \Phi$ as arc (arrow) from a to b, Element a and b are starting and ending vertex of the arc (a, b). We say that (a, b) is output of vertex a and b is input of b. As it follows from above definition graph has no multiple arcs.

The cardinalities of V and Φ are the order and size of the graph, respectively.

Graph is simple if Φ is symmetric and antireflexive relation. The information about simple graph can be given by edge i. e. set of kind $\{a, b\}$, where (a, b) is an arc. Graphically simple graph has no loops and multiple edges. In case of simple graph term size used for the number of edges within the graph.

The classical extremal graph theory studies extremal properties of simple graphs. Let F be family of graphs none of which is isomorphic to a subgraph of the graph Γ. In this case we say that Γ is F-free. Let P be certain graph theoretical property. By $\mathrm{ex}_P(v, F)$ we denote the greatest number of edges of F-free graph on v-vertices, which satisfies property P. If P is just a

property to be simple graph we omit index P and write $\mathrm{ex}(v, F)$. The missing definitions in extremal graph theory the reader can find in [11].

This theory contains several important results on $\mathrm{ex}(v, F)$, where F is a finite collection of cycles of different length [11], [28], [91]. The following statement had been formulated by P. Erdös'.

Let C_n denote the cycle of length n. Then

$$\mathrm{ex}(v, C_{2k}) \leq Cv^{1+1/k} \tag{4.1}$$

where C is independent positive constant.

For the proof of this result and its generalisations see [13], [29].

In [6] the upper bound

$$\mathrm{ex}(v, C_3, C_4, \ldots, C_{2k}, C_{2k+1}) \leq (1/2)^{1+1/k} v^{1+1/k} + O(V) \tag{4.2}$$

had been established for all integers $k \geq 1$.

Both bounds are known to be sharp for $k = 2, 3, 5$ in other cases the question on the sharpness is open (see [3], [11] and further references).

The girth of the simple graph is the minimal length of its cycle. So the above bound is the restriction on the size of the graph on v vertices of girth $\geq n$. Graphs of high girth, i.e. graphs which size is close to the above upper bounds can be used in Networking and Operation Research (see [11]) and Cryptography.

The generalisations of classical extremal graph theory on directed graphs require certain restrictions on inputs or outputs of the graph. Really, the graph: $P \cup L = V$, $|P \cap L| = 0$, $|V| = v$, $|P| = |L|$, $\Phi = P \times L$ of order $O(v^2)$ has no directed cycles or commutative diagrams.

In current section we generalise the above results on maximal size on the case of balanced graphs, when binary relation Φ is irreflexive and for each vertex $a \in V$ cardinalities of $\mathrm{id}(v) = \{x \in V | (a, x) \in \phi\}$ and $\mathrm{od}(v) = \{x \in V | (x, a) \in \phi\}$ are same. We refer to numbers $\mathrm{id}(v)$ and $\mathrm{od}(v)$ as input degree and output degree of vertex v in the graph, respectively (see section 7, where the pass, commutative diagram and the concept of girth indicator is defined0.

Let F be a list of directed graphs and P be some graph theoretical property. By $\mathrm{Ex}_P(v, F)$ we denote the greatest number of arrows of F-free directed graph on v vertices satisfying to property P (graph without subgraphs isomorphic to graph from F).

Let $\mathrm{E}_P = \mathrm{E}_P(d, v) = \mathrm{Ex}_P(v, O_{s,t}, s + t \geq 3 | 2 \leq s \leq d)$ be the maximal size (number of arrows) of the balanced binary relation graphs with the girth indicator $> d$.

The main result of [108] is the following statement. If B be the property to be the balanced directed graph, then

$$v^{1+1/d} - O(v) \le \mathrm{E}_B(d, v) \le v^{1+1/d} + O(v) \tag{4.3}$$

Notice, that the size of symmetric irreflexive relation is the double of the size of corresponding simple graph. because undirected edge of the simple graph corresponds to two arrows (arcs) of $O_{2,0}$. We will consider further the balanced graphs only and omit the index B.

If P is the property to be a graph of symmetric irreflexive relation then $\mathrm{Ex}_P(v, O_{s,t}, s + t \ge 3 | 2 \le s \le d) = 2\mathrm{ex}(v, C_3, \ldots, C_{2d-1}, C_{2d})$ because undirected edge of the simple graph corresponds to two arrows of $O_{2,0}$. So equality (1, 3) implies the following inequality

$$\mathrm{ex}(v, C_3, C_4, \ldots, C_{2k}) \le (1/2)v^{1+1/k} + O(V) \tag{4.4}$$

We evaluate the maximal size of the directed graph of order v with the girth indicator $> d$ which does not contain commutative diagrams $O_{d+1,d}$, as well. The inequality (4.2) is the corollary from such evaluation.

We can see that studies of extremal properties of balanced graphs with the high girth indicator and studies of $ex(v, C_3, \ldots, C_n)$ are far from being equivalent. Really, the sharpness of the Erdös' bound (4.1) and bounds (4.2) and (4.4) up to magnitude for $k = 8$ and $k \ge 12$ are old open questions (see [3], [11]).

The family of directed graphs G_i, $i = 1, \ldots$ with average output degree k_i and order v_i is the family of large girth if the girth indicator of G_i is $\ge \log_{k_i}(v)$. It agrees well with the standard definition for simple graphs. In case of balanced graphs of large girth their size is close to the upper bounds (4. 3).

4.2.1. On the upper bounds for size of the graphs with high girth indicator

Let Γ be the graph of irreflexive binary relation Φ on the vertex st V and the following property P holds:

for each vertex $v \in V$ the cardinalities $\{x|(x, v) \in \Phi\}$ and $\{x|(v, x) \in \Phi\}$ equals to same number k_v. As it follows from P the cardinality of $\{(x, y, z)|(x, y) \in \Phi \text{ and } (y, z) \in \Phi\}$ is $D = \sum_{v \in V} ({k_v}^2)$. So the number of random walk with two arrows from random vertex v is D/v. Any random walk in this graph can be viewed as the branching process with $\sqrt(D/v)$ branches from each node.

The bound $\mathrm{E}(d, v) \le v^{1+1/d} + O(v)$ is based on the studies of such branching process corresponding to the passes of length $\le d$ of the rooted tree. The definitions of such branching process, expectation operator and

the confidence interval the reader can find in the book [50] by Karlin and Taylor.

Theorem 20.

$$\mathrm{E}(d, v) \leq v^{1+1/d} + o(v^{1+1/d}) \tag{4.5}$$

$$\mathrm{Ex}(v, O_{d+1,d}, O_{s,t} | 3 \leq s \leq d) \leq (1/2)^{1/d} v^{1+1/d} + o(v^{1+1/d}) \tag{4.6}$$

For the demonstration of the justification technique of the upper bound for directed graphs we prove more general statement in the next section (upper bound for the size of directed graphs with large hooves)

In the next unit we show that the bounds of previous theorem are sharp.

It indicates that studies of extremal properties of graphs of binary relations with the high girth indicator and studies of $ex(v, C_3, \ldots, C_n)$ are far from being equivalent. Really, the sharpness of the Erdös' bound for $k = 4$ and $k \geq 6$ are old open questions.

4.2.2. On the sharpness of the bound

The diameter is the minimal length d of the shortest directed pass $a = x_0 \to x_1 \to x_2 \cdots \to x_d$ between two vertices a and b of the directed graph. We will say that graph is k-regular, if each vertex of G has exactly k outputs. Let F be the infinite family of k_i regular graphs G_i of order v_i and diameter d_i. We say, that F is a family of small world graphs if $d_i \leq C\log_{k_i}(v_i)$, $i = 1, \ldots$ for some independent on i constant C. The definition of simple small world graphs and related explicit constructions the reader can find in chapter 2, where some examples of small world graphs without small cycles are given.

Let M be a finite set, $m = |M| \geq 2$. We define M_k, $m \geq k + 2$ as the totality of tuples $(x_1, x_2, \ldots, x_k) \in M^k$, such that $x_i \neq x_j$ for each pair $(i, j) \in \{1, \ldots, k\}$. Let us consider the binary relation $\phi = \phi_k(m)$ on M_k consisting of all pairs of tuples $((x_1, \ldots, x_m), (y_1, \ldots, y_m))$, such that $y_i = x_{i+1}$ for $i = 1, \ldots, k-1$ and $y_m \neq x_i$ for each $i \in \{1, \ldots, k\}$. The corresponding directed graph $\Gamma = \Gamma_i(m)$ has order $m(m-1)\ldots(m-k+1)$, each vertex has $m - k$ input and output arrows.

Theorem 21. *The girth indicator and diameter of the graph $\Gamma_k(m)$ is $k+1$ and $2k$, respectively.*

Proof. Let us consider the $O_{s,t}$, $0 \leq t \leq s \leq k$, $s \geq 1$ of the graph $\Gamma_k(m)$ with the starting point $\mathrm{a} = (a_1, a_2, \ldots, a_k)$. Let $\mathrm{a}_x = (a_2, a_3, \ldots, a_k, x)$ be the neighbour of a within the pass P_x of the diagram of length s. Notice that x is different from a_i, $i = 1, 2, \ldots, k$. Let P be other pass of the

diagram. If length t of P is zero, we assume that P consist of one vertex a. The first component of ending point w of the P_x equals to x. But the first component of each vertex for each vertex of the pass P is either element of $\{a_1, a_2, \dots, a_k\}$ (case $t < s$ or element y, $y \neq x$ (case $t = s$). But w has to be the vertex of P as well. So we are getting a contradiction. Thus, we proved that the girth indicator of the graph is $> k$.

Notice that w $= (x, x_1, \dots, x_{k-1})$, where $x \neq a_i$, $i = 1, \dots, k$, $x_i \neq a_j$, $j = i+1, i+2, \dots, k$. We can add vertex $(x_1, x_2, \dots, x_{k_1}, x_k)$, consider the following specialization of variables $x_i = a_i$ for $i = 1, 2, \dots, k$ and obtain the diagram $O'_{0,k+1}$. So the girth indicator of the graph is $k+1$.

Let us consider the pass of length $2k$ starting from a and going throw w and $(x_1, x_2, \dots, x_k)$ as above. It contains the following tuples $(x_2, \dots, x_k, y_1)$, $(x_3, \dots, x_k, y_1, y_2), \dots, (x_k, y_1, \dots, y_{k-1}$. The only requirement on distinct elements $x_k, y_1, \dots, y_k$ is x_k *nex* and x can be arbitrarily element from the complement of $\{a_1, \dots, a_k\}$. If $m \geq k+2$, then arbitrary point of M_k can be reached from a via the pass as above and diameter of the graph is bounded by $2k$. It is clear that there is no pass of length $2k-1$ between a and element of kind $(z_1, \dots, z_{k-1}, a_k)$. So $\mathrm{diam}(\Gamma_k(m)) = 2m$.

□

Corollary 8. *Let F be the family of graphs $\Gamma_m(k)$, $m = k+2, k+3, \dots$. Then it is a family of directed small world graphs, the size of the members of this family is on the bound (4.1) of theorem 1.*

Really, $\Gamma_m(k)$ has degree $m-k$, order $v = m(m-1)\dots(m-k+1)$. So $\log_{m-k}(v)$ is some constant $> k$. So diameter of graphs from the family is bounded by $2\log_{m-k}(v)$. The size of $\Gamma_m(k)$ is $v(m-k)$. We have $(m)^k \geq v$. So $E(\Gamma_m(k)) \geq v[(v^{1/k}) - k] = v^{1+1/k} - kv$.

Let us consider the bipartite analog $\Gamma' = \Gamma'_k(m)$ of the graph $\Gamma = \Gamma_k(m)$ Let M be a finite set, $m = |M| \geq 2$. Let P (point set) and L (line set) are two copies of the vertex set M_k, $m \geq k+2$ of the graph Γ. We will use the brackets and parenthesis for the tuples from P and L, respectively.

Let$\Gamma' = \Gamma'_k(m)$ be the graph of binary relation on $P \cup L$ consisting of all pairs of tuples $((x_1, \dots, x_m), [y_1, \dots, y_m])$ or $([x_1, \dots, x_m], (y_1, \dots, y_m))$, such that $y_i = x_{i+1}$ for $i = 1, \dots, k-1$ and $y_m \neq x_i$ for each $i \in \{1, \dots, k\}$. The corresponding directed graph $\Gamma' = \Gamma'_k(m)$ has order $2m(m-1)\dots(m-k+1)$, each vertex has $m-k$ input and output arrows.

Theorem 22. *The girth indicator and diameter of the graph $\Gamma'_k(m)$ is $k+1$ and $2k = 1$, respectively. The graph does not contain commutative diagram $O_{k+1,k}$.*

Proof. The graph does not contain $O_{k+1,k}$ because of the ending point of the diagram can not be point and line at same time.The evaluation of the

girth indicator and diameter can be done similarly to the evaluation in the proof of proposition 1.

□

Corollary 9. *Let F' be the family of graphs $\Gamma'_m(k)$, $m = k+2, k+3, \ldots$. Then it is a family of directed small world graphs, the size of the members of this family is on the bound (4.2) of theorem 1.*

Really, $\Gamma'_m(k)$ has degree $m-k$ and order $v = 2m(m-1)\ldots(m-k+1)$. We have $(m-k)^k \leq m(m-1)\ldots(m-k+1)$. So $k \leq \log_{m-k}(m(m-1)\ldots(m-k+1))$. Thus $2k+1 < 3k \leq 3\log_{m-k}(m(m-1)\ldots(m-k+1)) < 3\log 2m(m-1)\ldots(m-k+1) = 3\log_{m-k}(v)$

The size of $\Gamma'_m(k)$ is $v(m-k)$. We have $(m)^k \geq m(m-1)\ldots(m-k+1) = v/2$. So $m > (1/2)^k v^{1/k}$. Thus $E(\Gamma'_m(k)) \geq v[(1/2)^{1/k} v^{1/k}) - k] = (1/2)^{1/k} v^{1+1/k} - kv$.

4.2.3. Remarks on the applications of graphs of large girth to Coding Theory

Low-density parity-check (LDPC) codes were originally introduced in doctoral thesis by Gallager [41] in 1961. The discovery of Turbo codes by Berrou, Glavieux, and Thitimajshima [5] in 1993, and the rediscovery of LDPC codes by Mackay and Neal [67] in 1995 renewed interest in Turbo codes and LDPC codes, because their error rate performance approaches asymptotically the Shannon limit. Much research is devoted to characterizing the performance of LDPC codes and designing codes that have good performance.

Commonly, the Tanner graph (see [94] and further references), is associated with the code and an important parameter affecting the performance of the code is the girth of corresponding Tanner graph. The design of structured regular LDPC codes whose Tanner graphs have large girth is considered in [45, 46, 77]. The regularity and structure of LDPC codes utilize memory more efficiently and simplify the implementation of LDPC coders. The Tanner graph is simple bipartite graph, in which the set of nodes is divided into two disjoint classes with edges only between nodes in the two different classes. The first such codes corresponds to finite generalised polygons) introduced by J. Tits [96] (see also well known paper [30]). Modern studies of generalised polygons are reflected in [95]. The graphs $D(n, F_q)$ have been used in [45], [46]. The impotence of the studies of undirected regular bipartite graphs with large girth for the design of turbo codes is discussed in [31].

Large girth speeds the convergence of iterative decoding and improves the performance of LDPC codes, at least in the high SNR range. Large size of such graphs implies fast convergence.

The concept of directed graph (finite automaton) could be used also in Quantum Coding Theory [1].

4.3. On directed graphs with large hooves

We will consider graphs of binary relations, for which the number of inputs (or outputs alternatively) for each vertex is ≥ 1. The commutative diagram is formed by two directed passes for which the same starting and ending points form the full list of common vertices. If the lengthes of these pases are equal t we use term *horseshoe* of size t and denote the diagram by Hs_t We define the *hoof size* $\mathrm{h} = \mathrm{h}(G) \geq 2$ of the directed graph G as the minimal size of its horseshoe. We compute the maximal size $Ex(\mathrm{Hs}_2, \mathrm{Hs}_3, \ldots, \mathrm{Hs}_d, v)$ (number of arrows) for the directed graph on v vertices with the size of hoof is $> d$ ($\mathrm{Hs}_2, \mathrm{Hs}_3, \ldots, \mathrm{Hs}_d$ is the list of prohibited subgraphs for the graph). The computation is based on the combinatorial upper bounds and explicit construction of the family of small world graphs with the increasing size of the hooves. We prove that $Ex(\mathrm{Hs}_2, \mathrm{Hs}_3, \ldots, \mathrm{Hs}_d, v) <=> v^{1+1/d}$. Notice that last inequality implies that $ex(C_4, C_6, \ldots, C_{2d}) \leq 1/2v^{1+1/d} + o(v^{1+1/d})$.

In the next section we observe classical results on Turan type problems on studies of the maximal size of simple graphs without prohibited cycles. Such problems were attractive for mathematicians because they are beautiful and difficult. Later the applications of these problems in Networking [6], Coding Theory and Cryptography were found.

We also discuss analogues of such problems for directed graphs without loops and multiple edges and formulate some results on the maximal size of digraphs without certain commutative diagrams, motivated by studies of turbo-codes and encryption algorithms. we introduce the concept of girth indicator here.

The definition of hoof size and the justification of equivalence

$$Ex(\mathrm{Hs}_2, \mathrm{Hs}_3, \ldots, \mathrm{Hs}_d, v) <=> v^{1+1/d}$$

the reader can find in section 4.3.2.

4.3.1. On the extremal graphs and digraphs without certain commutative diagrams

Classical Extremal Graph Theory developed by P. Erdös' and his school had been started with the following problem formulated by Turan.

What is the maximal value $ex(v, C_n)$ for the size (number of edges) of graph on v vertices without cycles C_n of length n? (see [4], [27] and further references). To discuss the behavior of function $ex(v, C_n)$ for large variable v we will use the following standard notations

Let f and g be two real valued functions on (a, ∞).

1. $f(x) <=> g(x)$, $x \to \infty$ if $f(x)/g(x) \to 1$ for $x \to \infty$;
2. $f(x) = o(g(x))$, $x \to \infty$ if $f(x)/g(x) \to 0$ for $x \to \infty$;
3. $f(x) = O(g(x))$, $x \to \infty$ if there exist C and x_0 such that $|f(x)| < C|g(x)|$ for all $x > x_0$;
4. $f(x) = \Omega(g(x))$, $x \to \infty$ if there exist a $c > 0$ and a sequence $x_1, x_2, \cdots \to \infty$ such that $|f(x_i)| > c|g(x_i)|$ for all $i \geq 1$.

If $n = 2k+1$ is odd we can assume that v is even and take the complete bipartite graph with the partition sets of same cardinality $v/2$. It contains $v^2/4$ vertices, so $ex(v, C_{2k+1}) = O(v^2)$.

If n is even, then according to famous Erdös' Even Circuit Theorem $ex(v, C_{2k}) = O(v^{1+1/k})$. This proof was obtained by famous Erdös' probabilistic method. Recall that the upper bound of the theorem is known to be sharp $ex(v, C_{2k}) = \Omega(v^{1+1/k})$ for $k = 2, 3$ and 5 only (see [28], [29] for $n = 2$ and [3] for $n = 3, 5$). The equivalence $ex(v, C_4) <=> 1/2v^{3/2}$ was obtained in [27] and [28]. The best lower bound $ex(v, C_6) \geq 1/2v^{4/3} + o(v^{4/3})$ was proved in [65]. The best known lower bound for the case $n = 5$ had been obtained in [66]: $ex(v, C_{10}) \geq 4/5^{6/5}v^{6/5}$.

The girth $g(G)$ of the simple graph G is the length of its smallest cycle.

The studies of maximal size $ex(v, C_3, C_4, \ldots, C_n)$ for graph on v vertices without cycles $C_3, C_4, \ldots, C_n$, i.e. graphs of girth $> n$ historically had been motivated by their applications to Telephone Networking. As it follows from Erdös' Even Circuit Theorem $ex(v, C_3, C_4, \ldots, C_{2n}) = O(v^{1+1/n})$.

More precise evaluations lead to the following bounds:

$$\mathrm{ex}(v, C_3, C_4, \ldots, C_{2k}, C_{2k+1}) \leq (1/2)^{1+1/k}v^{1+1/k} + o(v^{1+1/k}) \qquad (4.3.1)$$

$$\mathrm{ex}(v, C_3, C_4, \ldots, C_{2k}) \leq (1/2)v^{1+1/k} + o(v^{1+1/k}) \qquad (4.3.2)$$

The inequality (4.3.1) is established in [29] for all integers $k \geq 1$. The upper bound (1.2) can be obtained by similar probabilistic arguments (see, for instance, [119] and remarks after Theorem 1 below). Similar to the case of $ex(v, C_{2n})$ both bounds (4.3.1) and (4.3.2) are known to be sharp up to magnitude for $n = 2, 3$ and 5 only. The lower bound $ex(v, C_{10}) \geq 4/5^{6/5}v^{6/5}$ above and inequality (4.3.2) imply that $ex(v, C_{10}) \neq ex(v, C_3, C_4, \ldots, C_{10})$.

It is an interesting question whether or not

$$ex(v, C_6) \neq ex(v, C_3, C_4, C_5, C_6).$$

The first general lower bounds of kind

$$ex(v, C_3, C_4, \dots C_n) = \Omega(v^{1+c/n}) \tag{4.3.3}$$

where c is some constant $< 1/2$ was obtained in 50th by famous Erdös' via studies of *families of graphs of large girth*, i.e. infinite families of simple regular graphs Γ_i of degree k and order v_i such that the girth $g(\Gamma_i)$ is $\geq c\log_{k_i} v_i$, where c is the independent of i constant. Erdös' proved the existence of such a family with arbitrary large but bounded degree k with $c = 1/4$ by his famous probabilistic method.

Just several explicit families of graphs of large girth with unbounded girth and arbitrarily large k are known: the family of Cayley graphs had been defined in [69] and investigated in ([66], the family of algebraic graphs defined in [62] and its modifications suggested in [119].

Some of them can be easily converted in special finite automata and used for cryptographical purposes (see previous chapters on theoretical studies and software implementations). Graphs $D(n, q)$ and their directed analogues can be used in Coding Theory as so called Tanner graphs.

Notice that $ex(v, C_{2k}) \geq ex(v, C_3, C_4, \dots, C_{2k+1})$. The best known lower bound for $k \neq 2, 3, 5$ was obtained in [63]:

$$ex(v, C_3, C_4, \dots, C_{2k+1}) = \Omega(v^{1+2/(3k-3+e)}) \tag{4.3.4}$$

where $e = 0$ if k is odd, and $e = 1$ if k is even.

It is known that finite automaton roughly is a directed graph (or shortly digraph) with labels on arrows. So the Computer Science motivates the development of Extremal Graph Theory for Directed Graphs, which can be named alternatively as Extremal Digraph Theory. Last term is in the title of the current note.

Let us observe some analogues of $ex(v, C_3, C_4, \dots, C_n)$ for the special class of directed graphs.

Previously we consider *balanced* graphs for which the number i_v of inputs $x \to v$ and number o_v of outputs $v \to x$ are the same for each vertex v.

Recall that we use term *the family of graphs of large girth* for the family of balanced directed regular graphs Γ_i of degree k_i and order v_i such that $\mathrm{gi}(\Gamma_i)$ is $\geq c\log_{k_i} v_i$, where c' is the independent of i constant.

As it follows from the definition $g(\Gamma_i) \geq c'\log_{k_i}(v_i)$ for appropriate constant c'. So, it agrees with the well known definition for the case of simple graphs.

Let F be a list of directed graphs and P be some graph theoretical property. By $\mathrm{Ex}_P(v, F)$ we denote the greatest number of arrows of F-free directed graph on v vertices satisfying property P (graph without subgraphs isomorphic to graph from F). We will omit the index P in this section if it is just a property to be a balanced directed graph.

The maximal size $\mathrm{E}(d, v)$ (number of arrows) of the balanced binary relation graphs with the girth indicator $> d$ coincides with $\mathrm{Ex}(v, O_{s,t}, s+t \geq 2 \leq s \leq d)$.

Let $\mathrm{Ex}^{2d+1}(v))$ be the maximal size of the balanced directed graph of girth $> 2d+1$, then this number coincide with $\mathrm{Ex}(v, O_{d+1,d}, O_{s,t})$: $3 \leq s \leq d$).

The following analog of (4.3.1) has been stated previously.

$$E(d, v) <=> v^{1+1/d} \tag{4.3.5}$$

Remark 1. Let $E_P(d, v)$ be the maximal size (number of arrows) for the balanced graph on v vertices with the girth indicator $> d$ satisfying the graph theoretical property P. If P is the property to be a graph of symmetric irreflexive relation then $E_P(d, v) = 2\mathrm{ex}(v, C_3, \ldots, C_{2d-1}, C_{2d})$ because undirected edge of the simple graph corresponds to two arrows of symmetric balanced directed graph. So the bound (4.3.5) implies the inequality (4.3.2).

Remark 2. The precise computation of $E(d, v)$ does not provide the sharpness of (4.3.2). So the questions on the sharpness of (4.3.1) and (4.3.2) up to magnitude for $n \neq 3, 4$ and 5 are still open and the lower bound (4.3.5) is still the best known.

The above Theorem is analog of bound (4.3,2) for balanced directed graphs. The following analog of (4.3.1) was introduced previously.

$$\mathrm{Ex}^{2d+1}(v) <=> (1/2)^{1/d} v^{1+1/d} \tag{4.3.6}$$

Remarks:

(i) Let $E_P^{2d+1}(v)$ be the maximal size (number of arrows) for the balanced graph on v vertices with the girth $> 2d+1$ satisfying the graph theoretical property P. If P is the property to be a graph of symmetric irreflexive relation then $E_P^{2d+1}(v) = 2\mathrm{ex}(v, C_3, \ldots, C_{2d}, C_{2d+1})$ because undirected edge of the simple graph corresponds to two arrows of symmetric balanced directed graph. So the above Theorem implies the inequality (4.3.1).

(ii) The sharpness of the bound (4.3.1) does not follow from the above mentioned theorem. The function $\mathrm{ex}(v, C_3, \ldots, C_{2d}, C_{2d+1})$ is computed up to the magnitude for $d = 2, 3, 5$.

4.3.2. On the graphs with large hooves

In this section we will consider graphs of binary relations, for which the number of inputs (or outputs alternatively) for each vertex is ≥ 1. We refer to the commutative diagram $O_{s,t}$ with $s = t$ as *horseshoe* of size t and denote it by Hs_t We define the *hoof size* $\mathrm{h} = \mathrm{h}(G) \geq 2$ of the directed graph G as the minimal size of its horseshoe. We shall study the maximal size $Ex(\mathrm{Hs}_2, \mathrm{Hs}_3, \ldots, \mathrm{Hs}_d, v)$ (number of arrows) for the graph on v vertices with the size of hoof is $> d$. The sequence $\mathrm{Hs}_2, \mathrm{Hs}_3, \ldots, \mathrm{Hs}_d$ is the list of prohibited subgraphs for the graph

Theorem 23.

$$Ex(\mathrm{Hs}_2, \mathrm{Hs}_3, \ldots, \mathrm{Hs}_d, v) <=> v^{1+1/d} \qquad (4.3.7).$$

The equality (4.3.1) follows instantly from the above statement.

Proof. let us take care on lower bound for $E = Ex(\mathrm{Hs}_2, \mathrm{Hs}_3, \ldots, \mathrm{Hs}_d, v)$ first. Let us fixe the parameter v. Number E is a maximal size of the graph from the set T of all graphs with vertices of output degree ≥ 1, which do not contain commutative diagrams $O_{2,2}, O_{3,3}, \ldots, O_{d,d}$. Extremal balanced graph does not contain isolated points. So, $E(d, v$ is a maximal size of the graph from the set T' of balanced graphs with output degree ≥ 1, which do not contain commutative diagrams $O_{r,s}$ with $s \leq r \leq d$. It means that T' is a subset of T and $E \geq E(d, v) \geq v^{1+1/d} + o(v^{1+1/d})$ (look at theorem 1 from previous section) Let n_w be the output degree of extremal graph G from the set T. So E is the sum of all $n_w \in V(G)$. The ratio $a = E/v$ is an average output degree of vertex from $V(G)$. As it follows from our lower bound $a \geq v^{1/d}$. Let us consider the totality Δ of all directed passes of kind $w_1 \to w_2 \to \cdots \to w_d$. Let $\Delta(w)$ be the totality of all passes with the starting point w. The property "output degree of w is ≥ 1 implies that $\Delta(w$ is not an empty set. This property allow us to estimate size of Δ from below: $|\Delta| \geq va^d - o(va^d)$. There is a special vertex $w*$, such that $|\Delta(w*)|$ is $\geq a^{1/d}$ because absence of such a vertex contradicts to our lower bound on Δ. Let us consider two different passes P_w and P'_w from $\Delta(w*)$ with ending points w and w' . Let $S(P_w)$ and $S(P_{w'}$ be sets of internal vertices for each pass (without w, w' and v). If $|S(P_w) \cap S(P_{w'})| = 0$ then $w \neq w'$ because of absence of diagrams Hs_d. If $|S(P_w) \cap S(P_{w'})| \neq 0$ and $w = w'$ then the induced subgraph on the set of vertices $S(P_w) \cup S(P_{w'}) \cup \{w^*, w\}$ is union of several commutative diagram O_{r_i,s_i}, $i = 1, \ldots, l$ with $r_i \neq s_i$ and several common arrows $a_1, a_2, \ldots a_t$. We assume that arrows of length r_i belong to one pass of length d and remaining arrows belong to other pass. We have $s_1 + s_2 + \ldots, +s_l = r_1 + r_2 +, \cdots + r_l$, $l \geq 2$. We can construct the pass P from w^* to w containing passes of length $\min(r_i, s_i)$, $i = 1, \ldots, l$ and arrows

$a_1, a_2, \dots a_t$ The length of P is $< d$. The number of vertices at the distance $< d$ from w^* is $o(|\Delta|)$. It means that all ending points at distance d from w^* are different, the number $N = N(v)$ of such points can be estimated as $|\Delta(w^*)|$ up to equivalence <=>. So, $a^d - o(v^{1/d}) \leq \Delta(w^*) \leq v$, $a <=> v^{1/d}$ and $E <=> v^{1+1/d}$. □

Remark. The condition that output (or input, alternatively) degrees are ≥ 1 is important. Really, the graph with the vertex set $V = P \cup L$ with the subdivision into point set P and line set L of same cardinality, $|P \cap L| = 0$, $|V|$ is even number v, formed by all arrows from point to line has order $O(v^2)$ and does not contain commutative diagrams.

Recall that $ex(C_4, C_6, \dots, C_{2d})$ is the maximal size of simple graph on v vertices without cycles of length $4, 6, \dots, 2d$. It is well known that extremal simple graph G is connected. The corresponding directed graph of symmetric binary relation contains twice more vertices and does not contain $Hs_2, Hs_4, \dots, Hs_{2d}$. So, directly from previous theorem we get the following known statement.

Corollary 10.

$$ex(C_4, C_6, \dots, C_{2d}) \leq 1/2v^{1+1/d} + o(v^{1+1/d})$$

Remark We are not able to deduce the sharpness of the above upper bound from theorem 3 because the explicit construction supporting sharpness of the upper bound of Theorem 23 had been obtained via the family of graphs of asymmetrical relations (see previous sections).

Proposition 26. $ex(C_4, C_6, \dots, C_{2d}) <=> 1/2v^{1+1/d}$ *for* $d = 2, 3$ *and* 5.

Proof. The generalised m-gons had been defined by J. Tits as biregular bipartite graphs of diameter m and girth $2m$. So such graph does not contain cycles of length $4, 6, \dots, 2m-2$. Let us assume that r -regular generalised m-gon $GP_r(m)$ has a polarity automorphism π i.e.the symmetry of order 2 which maps set of points P onto set of lines L. Let I be the symmetric binary relation (incidence) corresponding to such a graph. Then we may consider the binary relation ϕ on the set P: $(p_1, p_2) \in \phi$ if and only if $(p_1, \pi(p_2)) \in I$. The new symmetric directed graph contains loops. Vertices with loops corresponded to so called *absolute points*. We delete loops, consider new symmetric binary relation I' and corresponding simple graph $GH(m)$ to I' (the polarity graph for the generalised m-gon). The graph $GH(m)$ contains $1+[(r-1)+(r-1)^2+\dots+(r-1)^{m-1}$ vertices The degree of absolute point is $r-1$ and degrees of remaining points are equal r. As it follows directly from the definition even cycles for $GH(m)$ and $GP_r(m)$ are same. The examples of regular generalised m-gons are known for $m = 3, 4$ and 6 only.In fact,

geometries of simple group of Lie type $A_2(q)$, $B_2(q)$ and $G_2(q)$ defined over finite field F_q are generalised $2m$-gons, respectively. Geometry $\Gamma(A_2(q))$ of group $A_2(q)$ (classical projective plane) admits the polarity for each finite field F_q. Geometries $\Gamma(B_2(q))$ and $\Gamma(G_2(q))$ of simple groups $B_2(q)$ and $G_2(q)$ have a polarity if and only if $q = 2^{2t+1}$ and $q = 3^{2t+1}$, respectively. Let $\Gamma'(A_2(q))$, $\Gamma'(B_2(q))$, $q = 2^{2t+1}$ and $\Gamma'(G_2(q))$,$q = 3^{2t+1}$ be polarity graphs for geometries of groups $A_2(q)$, $B_2(2^{2t+1})$ and $G_2(3^{2t+1})$,respectively. We have $v = |V(\Gamma'(A_2(q))| = 1 + q + q^2$, degree of each vertex of $\Gamma'(A_2(q))$ is $\geq q$ and this graph does not contain C_4. It means that the size of the family $(\Gamma'(A_2(q))$ is $1/2(v^{1+1/2} + o(v^{1+1/2})$. This value is on the bound of theorem 3 for $d = 2$. Similarly, we have $v = |V(\Gamma'(B_2(q))| = 1 + q + q^2 + q^3$, $q = 2^{2t+1}$ degree of each vertex of $\Gamma'(B_2(q))$ is $\geq q$ and this graph does not contain C_4 and C_6. So the size is on the bound of Theorem 3 for $d = 3$. In case of polarity graph $\Gamma'(G_2(q))$, $q = 3^{2t+1}$ the order v equals to $1 + q + q^2 + q^3 + q^4 + q^5$, each vertex of this graph has degree $\geq q$ and the graph does not contain C_4, C_6, C_8, C_{10}. It means that the size is on the upper bound of theorem 23 for $d = 5$.

□

Other examples of graphs (affine subgraphs of generalised m-gons, $m = 3, 4, 6$) with the size on the upper bound of Theorem 3 ($d = 3, 4, 6$) the reader can find in [123].

The following general lower bound for $k \neq 2, 3, 5$ can be obtained from the studies [64] of polarity graphs for the family of graphs $D(n, q)$.

Proposition 27.

$$ex(v, C_4, C_6, \ldots, C_{2k}) \geq 1/2(v^{1+2/(3k-3+e)}) - o(v^{1+2/(3k-3+e)})$$

where $e = 0$ if k is odd, and $e = 1$ if k is even.

4.4. On the directed graphs without commutative diagrams of rank $< d$ of minimal order

Recall that (k, g)-cage is a simple graph of degree k, girth g of minimal order $v(k, g)$. The following objects are similar to classical cages.

Definition 1. *We refer to the directed graph with the girth g, output degree k and minimal order $u(k, g)$ as directed (k, g)-cage.*

As it follows from the definition of directed (k, g)-cage

Theorem 24. *The following inequalities hold:*

$$(k + d)(k + d - 1) \ldots (k + 1) \geq u(k, 2d + 1) \geq 1 + k(k - 1) + \ldots k(k - 1)^{d-1},$$

$$2[(k+d)(k+d-1)\dots(k+1)] \geq u(k,2d+2) \geq (1+(k-1)+\dots(k-1)^d)+(k-1)^d$$

Proof. Let Γ be directed graph with k-outputs for each vertex and girth indicator d, then the branching process starting with the chosen vertex a gives $s = 1 + k + k(k-1) + \dots k(k-1)^d$ different vertices. So we prove (i).

Let us consider the arc $a \to b$. We have $k-1$ output arcs (a,x) from a, which are different from (a,b). The branching process starting from each element $x \neq b$ gives at least $(k-1) + \dots (k-1)^{d-1}$ passes of length $\leq d-1$. This way we get set T of elements of distance $(d-1)$ from a. Let us consider arcs of kind (b,y), $y \neq a$. The branching process from y gives us $(q-1)+(q-1)^{d-1}$ at distance $d-2$ from y. Together with b we have $1 + (q-1) + \dots (q-1)^{d-1}$ elements at distance $\leq d-1$ from b. This set has empty intersection with T because of absence of commutative diagrams $O_{d+1,d}$. So we have at least $(1 + (k-1) + \dots (k-1)^d) + (k-1)^d$ different vertices of the graph. □

Proposition 28. *Let Γ be directed cage with the output degree ≥ 3 of order v and girth indicator d.*

(i) If its girth is $2d+1$, then the size E of the graph satisfies the following inequality

$$v^{1+1/d} - dv \leq E \leq v^{1+1/d} + v$$

(ii) if its girth is $2d+2$, then the size E of the graph satisfies the following inequality

$$(1/2)^{1/d} v^{1+1/d} - dv \leq E \leq (1/2)^{1/d} v^{1+1/d} + v$$

Proof. The proof of the theorem 1 establishes in fact the upper bound on E. Let us consider the case of odd girth. We have $(k+d)^d > (K+d)(k+d-1)\dots(K+d+1) \geq v$. Thus $k+d > v^{1/d}$ and $k > v^{1/d} - d$. So $E = Vk > v(v^{1/d} - d)$.

In case of even girth we have $2(k+d)^d > 2(K+d)(k+d-1)\dots(K+d+1) \geq v$, which leads to $(v/2) < (k+d)^d$, $(k+d) > (1/2)^{1/d} v^{1/d}$ and $k > (1/2)^{1/d} v^{1/d} - d$. So $E = vk > (1/2)^{1/d} v^{1+1/d} - dv$ in this case.

□

Let P be some property of directed regular graphs and $u_P(k,g)$ be the minimal order of graph with the output degree K and the girth indicator g. It is clear that $u_P(k,g) \geq u(k,g)$. So $v(m,g) \geq u(m,g)$, in particular. The following statement follows immediately from the above inequalities.

Corollary 11. *Let s be the property to be simple graph. Then*

$$a) \quad v(k,2d+1) = u_S(k,2d+1) \geq u(k,2d+1) \geq 1 + k + k(k-1) + \cdots + k(k-1)^{d-1},$$

$$b) \quad v(k,2d+2) = u_s(k,2d+2) \geq u(k,2d+2) \geq (1 + (k-1) + \cdots + (k-1)^d) + (k-1)^d$$

The above lower bound for $g = 2d+2$ can be improved by Tutte inequality $v(k,2d+2 \geq 2(1+(k-1)+\ldots(k-1)^d)$ (see [15]). The Tutte's lower bound for $v(k,2d+2)$ is the same with (b). The upper and lower bound for $U(k,g)$ are quite tight, both of them are given by polynomial expression in variable k of kind $k^d + f(k)$, where $d = [(q-1)/2]$ and $\deg f(x) \leq d-1$. The situation with the known upper bound on the order of cages is different, such bound is quite far from the known lower bound.

Cages of odd girth with the order on the Tutte's bound are known as Moore graphs. There are only finite examples of Moore graphs. Well-known finite generalized m-gons are examples of cages of even girth (see next section of the paper).

From the existence of the k-regular Moore graph of girth $2d+1$ $(2d+2)$ follows $U(k,d) = v(k,2d+1) = 1 + k(k-1) + \ldots k(k-1)^{d-1}$ $(u(k,d) = v(k,2d+1) = 2(1+(k-1)+\ldots(k-1)^d)$, respectively.

There is a finite number of Moore graphs of order v of odd girth. Some infinite families of Moore graphs of even girth are known (see [17] or next section).

Proposition 29. *Let A be the property to be the graph of antisymmetric relation Φ i.e. $(a,b) \in \Phi$ implies that $(b.a)$ is not in Φ. Then*

(i) $(k+d)(k+d-1)\ldots(k+1) \geq u_A(k,2d+1) \geq 1 + k + k^2 + \ldots k^d$,

(ii) $2[(k+d)(k+d-1)\ldots(k+1)] \geq u_A(k,2d+2) \geq [1+k+k^2+\ldots k^d] + (k-1)k^{d-1}$.

Proof. The graphs $\Gamma_k(d)$ and $\Gamma'_k(d)$ are antisymmetric graphs. So their size establishes the upper bounds for $u_A(2k+2)$ and $u_A(2k+1)$, respectively. The antisymmetric graph does not contain diagram $O_{1,1}$. So, the branching process starting from chosen vertex a produces $1 + k + k^2 + \ldots k^d$ different vertexes and we are getting lower bound of the inequality (i).

Let b satisfy $a \to b$. The branching process with $d-1$ steps starting from b forms set B with $1 + k + k^2 + \ldots k^d$ distinct elements. We can consider $k-1$ output arcs (a,x) from a, which are different from (a,b). The branching processes with $d-1$ steps starting from each x bring $(k-1)k^{d-1}$ different elements at distance d from the vertex a. They are different from elements of B because of the absence of diagrams $O_{d+1,d}$ and diagrams $O_{d,s}$, $1 \leq s \leq d$. So we get the lower bound of inequality (ii).

□

The bounds (i) and (ii) are valid for balanced antisymmetric regular graphs because of $\Gamma_k(d)$ and $\Gamma'_k(d)$ are balanced graphs.

4.5. Algebraic explicit constructions of extremal regular directed graphs with the fixed girth indicator

We shall use the term of *algebraic graph* for the of graph $\Gamma(K)$ of binary relation Φ, such that the vertex set $V(\Gamma) = V(K)$ is an algebraic variety over commutative ring K of dimension ≥ 1 and for each vertex $v \in V$ the neighborhoods $\mathrm{In}(v) = \{x|(x,v) \in V\}$ and $\mathrm{Ou}(v) = \{x|(v,x) \in V\}$ are algebraic varieties over K of dimension ≥ 1 as well (see [7] for the case of simple graphs). We shall use the term *the family of directed graphs of large girth* for the family of regular graphs Γ_i with output degree k_i and order v_i such that their girth indicator $d_i = \mathrm{gi}(\Gamma_i)$ are $\geq c \log_{k_i}(v_i)$, where $c > 0$ does not dependent on i. So the size of such graphs is quite close to the above bounds.

We say that Γ_i form a family of asymptotical directed cages of odd (even) girth if $v_i = {k_i}^{d_i-1} + o({k_i}^{d_i-1})$ ($v_i = 2{k_i}^{d_i-1} + o({k_i}^{d_i-1})$. It is clear that asymptotical cages of even or odd girth are families of graphs of large girth.

In this section we consider examples of families of algebraic graphs of large girth with fixed girth indicator, asymptotical directed cages of odd and even girth, in particular.

Recall that we use used term *tactical configuration* of order (s,t) for biregular bipartite simple graphs with bidegrees $s+1$ and $r+1$. It corresponds to incidence structure with the point set P, line set L and symmetric incidence relation I. Its size can be computed as $|P|(s+1)$ or $|L|(t+1)$.

Let $F = \{(p,l)|p \in P, l \in L, pIl\}$ be the totality of flags for the tactical configuration with partition sets P (point set) and L (line set) and incidence relation I. We define the following irreflexive binary relation ϕ on the set F:

$((l_1,p_1),(l_2,p_2)) \in \phi$ if and only if $p_1 I l_2$, $p_1 \neq p_2$ and $l_1 \neq l_2$. Let $F(I)$ be the binary relation graph corresponding to ϕ. The order of $F(I)$ is $|P|(s+1)$ (or $|L|(t+1)$ We refer to it as *directed flag graph* of I.

Lemma 17. *Let (P,L,I) be a tactical configuration with bi-degrees $s+1$ and $t+1$ of girth $g \geq 4k$. Then the girth indicator of directed graph $F(I)$ with the output and input degree st is $> k$.*

Proof. The absence of even cycles C_{2s}, $2 < s < 2k-2$ in the graph I insure the absence of commutative diagrams $O_{r,s}$, $1 \leq s \leq r \leq k$ in the directed graph $F(I)$. □

Let (P, L, I) be the incidence structure corresponding to regular tactical configuration of order t.

Let $F_1 = \{(l,p)|l \in L, p \in P, lIp\}$ and $F_2 = \{[l,p]|l \in L, p \in P, lIp\}$ be two copies of the totality of flags for (P, L, I). Brackets and parenthesis allow us to distinguish elements from F_1 and F_2. Let $DF(I)$ be the directed graph (double directed flag graph) on the disjoint union of F_1 with F_2 defined by the following rules

$(l_1, p_1) \to [l_2, p_2]$ if and only if $p_1 = p_2$ and $l_1 \neq l_2$,

$[l_2, p_2] \to (l_1, p_1)$ if and only if $l_1 = l_2$ and $p_1 \neq p_2$.

Lemma 18. *Let (P, L, I) be a regular tactical configuration of order s with the girth $g \geq 2m$. Then the girth indicator of double directed graph $DF(I)$ with the output and input degree s is $> m$.*

Proof. The absence of even cycles C_{2s}, $2 < s < m-1$ in the bipartite graph I insure the absence of commutative diagrams $O_{r,s}$, $1 \leq s \leq r \leq m$ in the double directed graph $DF(I)$.

□

Generalized m-gons $GP_m(r,s)$ of order (r,s) were defined by J. Tits in 1959 as tactical configurations of order (s,t) of girth $2m$ and diameter m.

According to well known Feit - Higman theorem a finite generalized m-gon of order (s,t) has $m \in \{3,4,6,8,12\}$ unless $s = t = 1$.

The known examples of generalized m-gons of bidegrees ≥ 3 and $m \in \{3,4,6,8\}$ include rank 2 incidence graphs of finite simple groups of Lie type (see [18]). The regular incidence structures are $I_{1,1}(3,q)$ for $m = 3$ (groups $A_2(q)$), $I_{1,1}(4,q)$, $m = 4$ (groups $B_2(q)$) and $I_{1,1}(6,q)$, $m = 6$ (group $G_2(q)$). In all such cases $s = t = q$, where q is prime power.

The biregular but not regular generalized m-gons have parameters $s = q^\alpha$, $t = q^\beta$, where q is a prime power. The list is below: relation $I_{2,1}(4,q)$, $s = q^2, t = q$, q is arbitrary large prime power for $m = 4$; $I_{3,2}(6,q)$, $s = q^3, t = q^2$, where $q = 3^{2k+1}$, $k > 1$ for $m = 6$; $I_{2,1}(8,q)$, $s = q^2, t = q$, $q = 2^{2k+1}$ for $m = 8$. For each triple of parameters (m,s,t) listed above there is an edge transitive generalized m-gon of order (s,t) related to certain finite rank 2 simple group of Lie type. in the cases of $m = 3$ (projective planes. in particular) and $m = 4$ (generalized quadrangles) some infinite families of graphs without edge transitive automorphism group are known.

The following 2 lemmas can be obtained immediately from the axioms of generalized polygon.

Lemma 19. *Let (P, L, I) be the generalized $2k$-gon of order (r,s). Then*
$|P| = \sum_{t=0,k-1} (r^t s^t + r^{t+1} s^t)$, $|L| = \sum_{t=0,k-1} (s^t r^t + s^{t+1} r^s)$.

Lemma 20. *Let (P, L, I) be regular generalized m-gon of degree $q+1$. Then $|P| = |L| = 1 + q + \cdots + q^{m-1}$.*

Corollary 12. *For each $m = 3, 4, 6$ and prime p the family $F_m(q)$, $q = p^n$, $n = 1, \ldots$ of edge transitive polygons is an algebraic family over F_p of cages of girth $2m$ of degree $q + 1$ with the order on the Tutte's lower bound.*

Let (P, L, I) be generalized m-gon of order (s, t), $s \geq 2$, $t \geq 2$ and $e = \{(p, l\}$, $(p \in P, l \in L, pIl)$ be chosen edge of this simple graph.

Let $\mathrm{S}_e = \mathrm{Sch}_e(I)$ be the restriction of incidence relation I onto $P' \cup L'$ where P' (L') is the totality of points (lines) at maximal distance from p (l, respectively). It can be shown that (P', L', S_e) is a tactical configuration of degree $(s - 1, t - 1)$. Let us refer to (P', L', S_e) as Schubert graph. If the generalized polygon is edge-transitive its Schubert graph is unique up to isomorphism. In this case Schubert graph corresponds to the restriction of incidence relation onto the union of two of the largest "large Schubert cells", i. e. orbits of standard Borel subgroups of the highest dimension.

The following statement immediately follows from the definitions of graphs $S_m(q)$.

Proposition 30. *For each $S_m(p)$ $m = 3, 4, 6$ and prime p the family of Schubert graphs $S_m(p)$ of regular generalized m-gons $F_m(q)$ is algebraic over F_p family of asymptotical cages of even girth with the order $2q^{m-1}$ and degree q.*

The extremal properties of finite generalized polygons, their Schubert graphs and some of their induced subgraphs have been considered in [123].

Remark. The girth of $S_m(q)$ is $2m$ for "sufficiently large" parameter q.

Let (P, L, I) be a regular tactical configuration of order (t, t). The *double configuration* $I' = DT(I)$ is the incidence graph of the following incidence structure $(P', L', I') :$ $P' = F(I) = \{(p, l)|p \in P, l \in L, pIl\}$, $L' = P \cup L$, $f = (p, l)Ix$, $x \in L'$ if $p = x$ or $l = x$. It is clear that the order of tactical configuration I' is $(1, t)$. If (P, L, I) is a generalized m-gon, then (P', L', I') is a generalized $2m$-gon.

Proposition 31. *(i) If the girth of regular tactical configuration (P, L, I) of degree $s + 1$ is $2t$, then the girth of $DT(I)$ is $4t$. The order of $DT(I)$ is $(s, 1)$.*

(ii) Let (P, L, I) be regular generalized m-gon, then $DT(I)$ is generalized $2m$-gon.

Proof. It is clear that cycle C_l of length $2l$ in the simple graph $DT(I)$ corresponds to the cycle C_l of original tactical configuration. Notice that bipartite graphs does not contain odd cycles. So equality $g(I) = 2t$ implies $g(DT(I)) = 4t$.

Let I be generalised m-gon. Then the girth and diameter of m-gon are $g(I) = 2m$ and $d(I) = m$ respectively. As it follows from the definition the diameter of $DT(I)$ is twice large than $d(I)$. So the girth and diameter of $DT(I)$ are $4m$ and $2m$, respectively.

□

Corollary 13. *The configurations $DT(I) = I^2(m, q)$ for known regular m-gons, $m = 3, 4, 6$ of degree $q + 1$, q is a prime power, are generalized $2m$-gons of order $(1, q)$.*

Theorem 25. *(i) Let $I_{s.t}(m, q)$, $m \geq 4$ be the incidence relation of one of the known edge transitive m-gons defined over the field F_q, $q = p^n$, where p is a prime number.*
Then for each tuple (m, s, t, p) the family of directed flag-graphs $F^n = F^n(m, s, t, p)$, $n = 1, \ldots$ for generalized m-gon of order (q^s, q^t) is an algebraic over F_p family of asymptotic cages of odd girth. The girth indicator of each graph from the family is $m/2+1$ and the girth is $m+1$ $(5, 7, 9)$.

(ii) Let $S_{s.t}(m, q)$, $m \geq 4$ be the Schubert graph of the incidence relation $I_{s.t}(m, q)$ of one of the known edge transitive m-gons defined over the field F_q, $q = p^n$, where p is a prime number.
Then for each tuple (m, s, t, p) the family of directed flag-graphs $SF^n(m, s, t, p)$ for $S_{s.t}(m, q)$ is an algebraic family of asymptotic cages of odd girth defined over F_p. The girth indicator of graph from the family is $m/2 + 1$ and the girth is $m + 1$ if parameter q is sufficiently large.

(iii) Let $I_{1.1}(m, q)$ be the incidence relation of one of the known edge transitive regular m-gons defined over the field F_q, $q = p^n$, where p is a prime number. Then for each pair (m, p) the family $DF(m, p)$ of double flag graphs $DF(m, i) = DF(I_{1.1}(m, p^i))$, $i = 1, 2, \ldots$ is an algebraic over family of directed asymptotic cages of even girth defined over F_p . The girth indicator of each $DF(m, i)$ is $m + 1$ and the girth is $2m + 2$ (possible values are $(8, 10.14)$). Double flag graphs of Schubert subgraphs for $I_{1.1}(m, p^n)$, $n = 1, \ldots$ form the family of asymptotical directed cages as well.

(iv) Let $I^2(m, q)$, $m \geq 3$ be the incidence relation of double tactical configuration of regular generalized m-gon defined over F_q, $q = p^n$, where p is a prime. Then for each pair (m, p) the family $F(m, p)$ of directed flag-graphs $F^n(m, p)$ of $I^2(m, p^n)$, $n = 1, \ldots$ is an algebraic family of directed graphs of large girth over F_p. The girth indicator of each graph is $m + 1$ and girth is $2m + 1$ (possible values are 7, 9, 13).

Proof. As it follows from lemma 11 the girth indicator of each directed graph F^n is $> m/2$. The existence of cycles C_{2m} in the corresponding generalised m-gon leads to the existence of commutative diagrams $O_{m/2+1,m}$. So the girth indicator of each graph is $m/2+1$ and the girth is $2(m/2)+1$.

The order of each directed graph F^n coinsides with the cardinality of the flag set of the correspondent generalised m-gon or its size and can be given by polynomial expession $f(q)$ in single variable q (see lemmas 13 and 14 for the close formulae for the order). The degree of the balanced graph F^n is q^{s+t}. The highest term for the polynomial $F(q)$ is $q^{(s+t)m/2}$.

So for each prime p the family F^n is the family of asymptotical cages of odd girth and we proved statement (i) of the theorem.

The Schubert subgraphs SF^n is the induced subgraps of F^n. So for the the girth indicator and the girth of the Schubert subgraph we have $\mathrm{gi}(SF^n) \geq \mathrm{gi}(F^n) \geq m/2+1$ and $\mathrm{g}(SF^n) \geq \mathrm{g}(F^n) \geq m+1$. Notice that the order of SF^n is exactly $q^{(s+t)m/2}$. The assumption that $\mathrm{gi}(SF^n) > \mathrm{gi}(F^n) \geq m/2+1$ for sufficiently large q contrudict to previously proven statement (i) (or established upper bound for directed cages). So graphs (SF^n), $n = 1, \ldots$ form the family of asymptotical cages and we proved (ii).

The graphs $DF(m,i)$. $i = 1, \ldots$ are graphs of order $2f(q)$ where $f(q)$ is the order of corresponding directed flag graph F^i. As it follows from lemma 12 the girth indicator of each double directed graph $DF(m,i)$ is $> m$. The bipartite structure of the graph corresponding to the partition which formed by 2 copies of $F(I)$ insures the absence of commutative diagrams $O_{m+1,m}$ The existence of cycles C_{2m} in the corresponding generalised m-gon leads to the existence of commutative diagrams $O_{m+1,m+1}$. So the girth indicator of each graph is $m+1$ and the girth is $2m+2$. The highest term of polyniomial expression $2f(q)$ is $2q^m$. So the graphs form the family of asymptotical directed cages. Double flag graphs of Schubert subgraphs for $I_{1.1}(m, p^n)$, $n = 1, \ldots$ have order $2q^m$, $q = p^n$. So if n is sufficiently the girth indicator and gith of such graph is $m+1$ and $2m+2$, respectively. Thus we show that they form the family of asymptotical cages as well. So we proved point (iii).

Acording to proposition 17 the double tactical configuration $I^2(m,q)$, $q = p^n$, p is prime is generalised $2m$-gon. Similarly to part (i) of the proof we can show that the girth indicator of directed flag graph of $I^2(m,q)$ is $m+1$ and its girth is $2m+1$ (7, 9, 13). The order $v = v^n(m,p)$ of the graph $F^n(m,p)$ can be computed as the size of generalised $2m$-gon of order $(q,1)$. It is polynomial expression in variable q of degree m. So these graphs form the family of graphs of large girth.

□

Regular finite generalized polygons have been used in works of R. Tanner

on Coding Theory. The applications of biregular generalized polygons and their Schubert graphs to Cryptography the reader can find in [110].

Paper [114] is devoted to cryptographical algorithms based on nonsymmetric directed asymptotical cages as above. Let $DE_n(K)$ ($DE(K)$) be the double directed graph of the bipartite graph $D(n,K)$ ($D(K)$, respectively). Remember, that we have the arc e of kind $(l^1,p^1) \to [l^2,p^2]$ if and only if $p^1 = p^2$ and $l^1 \neq l^2$. Let us assume that the colour $\rho(e)$ of arc e is $l^1_{1,0} - l^2_{1,0}$.

Recall, that we have the arc e' of kind $[l^2,p^2] \to (l^1,p^1)$ if and only if $l^1 = l^2$ and $p^1 \neq p^2$. Let us assume that the colour $\rho(e')$ of arc e' is $p^1_{1,0} - p^2_{1,0}$. It is easy to see that ρ is a perfect algebraic colouring.

If K is finite, then the cardinality of the colour set is $(|K|-1)$. Let Reg$K$ be the totality of regular elements, i.e. not zero divisors. Let us delete all arrows with colour, which is a zero divisor. We will show that a new graph $RE_n(K)$ ($RE(K)$) with the induced colouring into colours from the alphabet $\mathrm{Reg}(K)$ is vertex transitive. Really, according to [59] graph $D(n,K)$ is an edge transitive. This fact had been established via the description of regular on the edge set subgroup $U(n,K)$ of the automorphisms group $\mathrm{Aut}(G)$. The vertex set for the graph $DE_n(K)$ consists of two copies F_1 and F_2 of the edge set for $D(n,K)$. It means that Group $U(n,K)$ acts regularly on each set F_i, $i = 1,2$. Explicit description of generators for $U(n,K)$ implicates that this group is a colour preserving group for the graph $DE_n(K)$ with the above colouring.

If K is finite, then the cardinality of the colour set is $(|K|-1)$. Let Reg$K$ be the totality of regular elements, i.e. non-zero divisors. Let us delete all arrows with colour, which is a zero divisor. We will show that a new affine graph $RE_n(K)$ ($RE(K)$) with the induced colouring into colours from the alphabet $\mathrm{Reg}(K)$ is vertex transitive. Really, the group $U(n,K)$ is an edge transitive on the graph $D(n,K)$. The vertex set for the graph $DE_n(K)$ consists of two copies F_1 and F_2 of the edge set for $D(n,K)$. It means that the group $U(n,K)$ induced on vertices of $DE_n(K)$ acts regularly on each set F_i, $i = 1,2$. The above explicit description of elements for $U(n,K)$ implicate that this group is a colour preserving group for the graph $DE_n(K)$ with the above colouring. The vertex set of $DE_n(K)$ coincides with the set of vertices of $RE_n(K)$, we just delete the arcs with colours from $\{K - \mathrm{Reg}(K)\}$. The permutation group of $(U(n,K), F_1 \cup F_2)$ is an automorphism group of the graph $RE_n(K)$. Let $n = 2s$ be even number then the polarity π of $D(2n,K)$ maps F_1 onto F_2 and F_2 on F_1. It means that group $\tilde{U}(2n,K)$ generated by elements of $U(2n,K)$ and π acts transitively on the vertex set of $RE_{2n}(K)$. So all connected components of the graph $RE_{2n}(K)$ are isomorphic. Let $CRE_{2n}(K)$ be such a connected component. Let us set $t = 2n$. We show that the girth indicator gi_t for the family of k-regular graphs $CRE_t(K)$ can

be evaluated via inequality $\mathrm{gi}_t \geq 2/(3log_k|K|)\log_k(v) + C$ where C is some independent on v constant.

We use the restrictions of the relations $DE_t(K)$ and $RE_t(K)$ on the vertices of the double flag graph for $CD_t(K)$. As it follows from the above discussion D_t is a union of connected components of graphs $CD_t(K)$. Each connected component of $CD_t(K)$ is a disjoint union of appropriate connected components of $CRE_t(K)$. Let v_t be the order of $CRE_t(K) = G_n(K)$. We set $|K| = k^\alpha$. The parameter v_t is $\leq 2|CD_t(K)||K|$ (the order of $DE_t(k)$). Instantly from the definition of $CD(t, K)$ we get $|DE_t(k)| = 2|K|^{3/4t+c}$, where c is some independent constant. So we get the following 3 equivalent inequalities:

$$v_t \leq 2k^{\alpha(3/4t+c_1)},$$
$$v_t/2 \leq k^{\alpha(3/4t+c_1)}$$
$$\log_k(v_t/2) \leq \alpha(3/4t + c_1).$$

From the last inequality we get

$$t \geq 4/(3\alpha)\log_k(v_t) + c_2$$

(c_1 and c_2 are independent on t constants).

According to [114] the family $RE_t(k)$ is a family of graphs with increasing girth indicator $\mathrm{gi}(RE_t(K)) \geq [(t + 4)/2]$. Notice that the girth indicators of $RE_t(K)$ and its connected component $CDE_t(K)$ coincide. It means that we prove

Theorem 26. *For each finite commutative ring K with at least 3 regular elements the family $CRE_{2n}(K)$, $n = 1, 2, \ldots$ is a family of directed graphs of large girth and the following lower bound foe the girth indicator holds*

$$\mathrm{gi}_t \geq (t + 2)/2 \geq 2/(3log_k|K|)\log_k(v) + C$$

The order of the graph $G_n(K) = CRE_2n(K)$ is less that the order of flag graph for $CD_n(K)$. We have $g(G_n(K)) \geq n+6$. According to [114] the r-regular graph has at least $2r^{(n+4)/2}$ vertices.

4.6. Simple homogeneous algebraic graphs over infinite field: two optimisation problems

Families of finite graphs of large girth were introduced in classical extremal graph theory. One important theoretical result here is the upper bound on the maximal size of the graph with girth $\geq 2d$ established in Even Circuit Theorem by P. Erdös. We consider some results on such algebraic

graphs over any field. The upper bound on the dimension of variety of edges for algebraic graphs of girth $\geq 2d$ is established. Getting the lower bound, we use the family of bipartite graphs $D(n, K)$ with $n \geq 2$ over a field K, whose partition sets are two copies of the vector space K^n. We consider the problem of constructing homogeneous algebraic graphs with a prescribed girth and formulate some problems motivated by classical extremal graph theory. Finally, we present a very short survey on applications of finite homogeneous algebraic graphs to coding theory and cryptography.

We study extremal graphs and their applications to coding theory, cryptography, and quantum computations. The main object of consideration is a homogeneous algebraic graph defined in terms of algebraic geometry in the following way.

Let F be a field. Recall that a *projective space* over F is a set of elements constructed from a vector space over F such that a distinct element of the projective space consists of all non-zero vectors which are equal up to a multiplication by a non-zero scalar. Its subset is called a *quasiprojective variety* if it is the set of all solutions of some system of homogeneous polynomial equations and inequalities.

An *algebraic graph* ϕ over F consists of two things: the *vertex set* Q being a quasiprojective variety over F of nonzero dimension, and the *edge set* being a quasiprojective variety ϕ in $Q \times Q$ such that $(x, x) \notin \phi$ for each $x \in Q$ and $x\phi y$ implies $y\phi x$ ($x\phi y$ means $(x, y) \in \phi$). The graph ϕ is *homogeneous* (or *M-homogeneous*) if for each vertex $v \in Q$ the set $\{x \mid v\phi x\}$ is isomorphic to some quasiprojective variety M over F of nonzero dimension. The reader can find the general conception of algebraic graphs [3].

We assume that the field F contains at least 5 elements. If F is finite then the vertex set and the edge set are finite and we get a usual finite graph.

The *cycle* C_t in ϕ is a sequence $x_1, x_2, \ldots, x_t$ of distinct elements of Q such that $x_1\phi x_2,\ x_2\phi x_3,\ \ldots,\ x_{t-1}\phi x_t,\ x_1\phi x_t$ are edges of the graph.

We define the *girth* $g = g(\phi)$ of a graph ϕ as the length of its minimal cycle. If ϕ is without cycles then $g(\phi) = \infty$.

The paper is devoted to the following two optimization problems:

(A) Let Q be a M-homogeneous graph such that $\dim M = k$ over F and its girth is a finite number g. What is the minimal possible dimension $v_a(k, g)$ for the variety of vertices?

(B) Let ϕ be a homogeneous graph of girth $g \geq t$ and $\dim M = k$. What is the maximal possible dimension of ϕ?

Problems (A) and (B) are related to each other, in case of finite field we can change the dimension of Q and ϕ on their cardinalities and get classical problems on minimal order of regular simple graph of given degree and given

girth (analogue of A) and maximal size (number of edges) of the graph with girth $\geq t$ (analogue of B).

So (A) and (B) are motivated by the branch of extremal graph theory which studies order of cages, related bounds, cages itself, bounds on maximal number of edges of the graph of given order and girth, and families of graphs of large girth (see Section 2).

we propose an analogue of Tutte's bound and variants of Erdös' Even Circuit Theory for homogeneous graphs, and define the family of algebraic graphs of large girth over an arbitrary field. Examples of extremal algebraic graphs of bounded dimension are presented. We formulate some open problems for general homogeneous graphs motivated by classical extremal graph theory.

we construct examples of families of algebraic graphs of large girth over fields and establish the upper bound on the minimal dimension of the vertex set for the graph of prescribed girth g.

We hope that the construction of homogeneous algebraic graphs over $\mathbb{C}$ and over the ring of Gaussian numbers can be used in quantum coding theory [1] and quantum cryptography [63].

4.6.1. Dense finite graphs of large girth and of large size

All graphs that 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 a finite graph G, respectively. The number of vertices $|V(G)|$ is called the *order* of G, and $|E(G)|$ is called the *size* of G. A path in G is called *simple* if all its vertices are distinct. When it is convenient, we identify G with the corresponding symmetric antireflexive binary relation Φ on $V(G)$, i.e. Φ is a subset of $V(G) \times V(G)$. The *length* of a path is the number of its edges.

The *girth* of a graph G, denoted by $g = g(G)$, is the length of a shortest cycle in G. Let $k \geq 3$ and $g \geq 3$ be integers. A (k, g)-*graph* is a k-regular graph with girth exactly g. A (k, g)-*cage* is a (k, g)-graph of minimal order. The problem of determining $v(k, g)$ of a (k, g)-cage is unsolved for most pairs (k, g) and is extremely hard in general case. By counting the number of vertices in the breadth-first-search tree of a (k, g)-graph, Tutte [17] established the following classical lower bounds for $v(k, g)$:

$v(k, g) \geq k(k-1)^{(g-1)/2}/(k-2)$ for g odd, $k \geq 4$,

$v(k, g) \geq 2(k-1)^{g/2-2}/(k-2)$ for g even, $k \geq 4$.

The graphs of odd girth for which equality holds are called *Moore graphs*. Each Moore graph with valency $k = 2$ is a polygon and each $(2d+1)$-gon is a Moore graph. Damerell proved that Moore graph with valency $k \geq 3$ has a diameter 2 and $k \in \{3, 7, 57\}$. There are unique examples for $k = 3$ (the

Petersen graph) and $k = 7$ (Hoffman-Singleton graph). No example with $k = 57$ is known, see [17].

The problem of determining $v(k, g)$ was posed in 1959 by F. Kartesi who observed that $v(3, 5) = 10$ was realized by the Petersen graph (see [17]). The classical extremal graph theory studies extremal properties of simple graphs. Let F be family of graphs none of which is isomorphic to a subgraph of the graph Γ. In this case we say that Γ is F-free. Let P be certain graph theoretical property. By $\mathrm{ex}_P(v, F)$ we denote the greatest number of edges of F-free graph on v-vertices that satisfies property P. If P is just a property to be simple graph we omit index P and write $\mathrm{ex}(v, F)$. The reader can find the missing definitions in extremal graph theory in previous units of this textbook.

Theorem 27. *Let G be quasi homogeneous algebraic graph over a field F of girth g such that the dimension of neighbourhood for each vertex is N, $N \geq 1$. Then*

$$[(g-1)/2] \leq \dim(V)/N.$$

Proof. Assume that $[(g-1)/2] = k > \frac{\dim V}{\dim N(F)}$. Let v be a vertex and M be the variety of elements at distance k from v. The absence of cycles C_s, $1 \leq s \leq 2k$, means that each element from M is connected with v by the unique pass. Elements of M are in one to one correspondence with such passes. Let $N_v(F)$ be a neighbourhood of v. A *pass* is a sequence $v, u_1, u_2, \ldots, u_k$, where $u_1 \in N_v(F)$, $u_2 \in N_{u_1}(F) - \{u_1\}, \ldots, u_k \in N_{u_{k-1}}(F) - \{u_{k-1}\}$. So the dimension of M is $N \times k$. But $N \times k > \dim V$ by our assumption, so we get a contradiction. □

We can rewrite the above statement in the form similar to Tutte's inequalities as follows.

Corollary 14. *Let G be an algebraic (k, g)-graph, i.e. a homogeneous algebraic graph over a field F of girth g such that the neighbourhood of each vertex is isomorphic to variety $N(F)$ of dimension k. Then* $\dim V(G) \geq [(g-1)/2]k$.

The following form of Theorem 2 is an analogue of inequality (1).

Corollary 15. *Let G be a homogeneous graph over a field F and $E(G)$ be the variety of its edges. Then* $\dim(E(G)) \leq \dim V(G)(1 + [(g-1)/2]^{-1})$.

Indeed, $\dim(E(G)) = \dim(V) \times \dim(N(F))$. From the previous inequality we have $\dim(N(F)) \leq \dim V(G)[2/(g-1)/2]$. So $\dim(E(G)) \leq \dim(V) \times \dim V(G)[(g-1)/2] = \dim V(G)((1 + [(g-1)/2]^{-1})$.

Let $v(k, g, F)$ be the minimal dimension of the variety of vertices for algebraic (k, g)-graph defined over F. Let $v_a(k, g)$ be the minimal dimension

of the variety of vertices for algebraic (k, g)-graph defined over some field F. We have

$$v_a(k, g) \geq [(g-1)/2]k. \tag{4.1}$$

The bi-homogeneous incident structure is a bipartite graph with partition sets P and L containing points and lines, respectively, such that there is a field F such that $P \cup L$ is an algebraic variety over F and neighbourhoods of each pair of points (lines) are isomorphic algebraic varieties over F as well. J. Tits [96] defined generalized m-gon as a graph of diameter m and girth $2m$, see also [95, 97].

We use the term *bi-homogeneous generalized polygon* for a bi-homogeneous incident structure which is a generalized polygon.

Theorem 28. *The equalities* $v_a(n, 6) = 2n$, $v_a(n, 8) = 3n$ *and* $v_a(n, 12) = 5n$, $n \geq 1$ *hold.*

Proof. From the previous theorem we have $v_a(n, 6) \leq 2n$, $v_a(n, 8) \leq 3n$ and $v_a(n, 12) \leq 5n$. Let G be Shevalley group of rank 2 defined over a field F which is an n-dimensional extension of field K. In particular, we can take $K = \mathbb{Q}$ or consider finite field $K = \mathbb{F}_p$, where p is prime, and define the extension via an irreducible polynomial of degree n. Let B be the standard Borel subgroup of G, P_1 and P_2 are standard parabolic subgroups of G, i.e. proper subgroups containing B. The geometry of G is the incidence structure with the point set $(G : P_1)$ and the line set $(G : P_2)$ consisting of left cosets gP_i, $i = 1, 2$. A point $p \in (G : P_1)$ and a line $l \in (G : P_2)$ are *incident* (pIl) if and only if the set theoretical intersection of cosets p and l is not empty. The simple graph of binary relation I is a homogeneous algebraic graph over $F = K^n$, the neighbourhood of each vertex is isomorphic to projective line over F. So the dimension of neighbourhood over K is n. It is well known that graph I is an algebraic generalized m-gon. For each field F we have the following options:

(i) $G = A_2(F)$ (classical linear group $PSL_2(F)$), $m = 3$.

(2i) $G = B_2(F)$ (classical projective symplectic group $PSp_4(F)$), $m = 4$.

(3i) $G = G_2(F)$ (well known Dixon group over F), $m = 6$.

The projective variety $(G : P_1) \cup (G : P_2)$ has dimension $m - 1$ over the field F. So for each $m \in \{3, 4, 6\}$ we have an example of algebraic (n, g)-graph of dimension $n(m-1)$. So $v_a(n, 2m) = n(m-1)$ for $m \in \{3, 4, 6\}$. □

Let $e(g, n)$ be the maximal dimension of the variety of edges of homogeneous algebraic graph of girth g with the n-dimensional variety of vertices. The following equalities are dual to equalities of the above theorem.

Corollary 16. *The following equalities hold:* $e(6,n) = n + n/2$ *for even* n, $n \geq 2$, $e(8,n) = n + n/3$ *for* $n = 3s$, $s = 1, 2, \ldots$, *and* $e(12,n) = n + n/5$ *for* $n = 5s$, $s = 1, 2, \ldots$

The following open problems are interesting:

(i) Find all values of girth g for which the lower bound (4.1) is sharp.

(ii) Find all values m for which there exist homogeneous algebraic generalized polygons. The word 'algebraic' is strict here, the polygon has to be a homogeneous algebraic graph in a sense of the above definition, i.e. neighbourhoods of each two vertices are isomorphic.

From the existence of homogeneous generalized m-gon follows that the bound (4.1) is sharp in case of girth $2m$.

As follows from [30], finite bi-homogeneous generalized m-gons are exist if $m \in \{3, 4, 6, 8\}$ (see [17]). Recall the assumption that each vertex of the graph contains at least 3 neighbours. J. Tits and R. Weiss (see [20]) classified all bi-homogeneous generalized m-gons with Moufang property.

Conjecture 1. *Equality* $v_a(k,g) = [(g-1)/2]k$ *implies* $g \in \{6, 8, 12\}$

We define the family of algebraic graphs of large girth G_i, $i = 1, 2, \ldots$ over the field F if the $\dim V(G_i)$ is growing and the girth of $G_i \geq c \cdot \frac{\dim V(G_i)}{\dim N(G_i)}$, where $c > 0$ is a constant independent of i. From Theorem 1 we get $c \leq 2$.

In the next section we prove the following upper bound on $v_a(k, 2s)$.

Theorem 29. *For each even* g, $g \geq 6$, *we have* $v_a(k,g) \leq k((3/4)g - \alpha)$, *where* $\alpha = 3, 5/2$ *for* $g = 0, 2 \mod 4$, *respectively.*

The problem of finding the good upper bound for $v_a(k, 2s+1)$ is very interesting. Algebraic $(k, 2s+1)$-graphs such that the dimension of variety of vertices is $v_a(k, 2s+1)$ are analogues of well known Moore graphs or cages with odd girth.

Bibliography

[1] Andris Ambainis, Ashwin Nayak, Amnon Ta-Shma, Umesh Vazirani, *Dense quantum coding and quantum finite automata*, Journal of the ACM (JACM), Volume 49 , Issue 4 (July 2002), 496 - 511.

[2] D. Atkins, M. Graff, A. K. Lenstra, P. C. Leyland, *The magic words are aquemish ossifrage*, ASIACRYPT-94, Lect. Notes in Comp. Sci, v. 917, 1995.

[3] C.T. Benson, *Minimal regular graphs of girth eight and twelve*, Canadien Journal of Mathematics, (18):1091- 1094, 1966.

[4] A. Beutelspachera, *Enciphered Geometry.* Some Applications of Geometry To Cryptography, Annals of Discrete Mathematics, V.37, 1988, 59-68.

[5] C. Berrou, A. Glavieux and P. Thitimajshima, *Near Shannon limit error-correcting coding and decoding: turbo-codes*, ICC 1993, Geneva, Switzerland, pp. 1064-1070, May 1993.

[6] F. Bien, *Constructions of telephone networks by group representations*, Notices Amer. Mah. Soc., 36 (1989), 5-22.

[7] N. Biggs, *Algebraic Graph Theory* (2nd ed), Cambridge, University Press, 1993.

[8] N.L. Biggs, *Graphs with large girth*, Ars Combinatoria, 25C (1988), 73–80.

[9] N.L. Biggs and A.G. Boshier, *Note on the Girth of Ramanujan Graphs*, Journal of Combinatorial Theory, Series **B 49**, pp. 190–194 (1990).

[10] N.L. Biggs and M.J. Hoare, *The sextet construction for cubic graphs*, Combinatorica **3** (1983), 153–165.

[11] B. Bollobás, *Extremal Graph Theory*, Academic Press, London, 1978.

[12] B. Bollobás, *Random Graphs*, Academic Press, London, 1985.

[13] J.A. Bondy and M.Simonovits, *Cycles of even length in graphs*, J. Combin.Theory, Ser. B, 16 (1974) 87-105.

[14] A. Borovik, *Matroids and Coxeter groups*, In: Survey in Combinatorics 2003, London Math Soc. Lect. Notes Ser., vol 307, Cambridge University Press, 2003, 79-114.

[15] A. Borovik, I. Gelfand, N. White, *Combinatorial flag varieties*, J. Comb. Theory (A), 2000, v. 91, 111-136.

[16] N. Bourbaki, *Lie Groups and Lie Algebras*, Chapters 1 - 9, Springer, 1998-2008.

[17] A. Brower, A. Cohen, A. Nuemaier, *Distance regular graphs*, Springer, Berlin, 1989.

[18] W. G. Brown, *On graphs that do not contain Thomsen graph*, Canad. Math. Bull. 9, No. 3 (1966), 281-285.

[19] A. A. Bruen D. L. Wehlau, *Error-Correcting Codes, Finite Geometries and Cryptography*, AMS, 2010.

[20] F. Buekenhout (Editor), *Handbook on Incidence Geometry*, North Holland, Amsterdam, 1995.

[21] P. J. Cameron and J.H. van Lint, *Graphs, Codes and Designs*, London. Math. Soc. Lecture Notes, 43, Cambridge (1980).

[22] R. W. Carter, *Simple Groups of Lie Type*, Wiley, New York (1972).

[23] A. Cossidente, M. J. de Ressmine, *Remarks on Singer Cycle Groups and Their Normalizers*, Desighns, Codes and Cryptography, 32, 97-102, 2004.

[24] P. Dembovski, *Finite Geometries*, Springer, Berlin, 1968.

[25] W. Diffie and M. E. Hellman *New directions in cryptography*, IEEE Transactions on Information Theory, vol. IT-22, Nov. 1976, 644-654.

[26] E. Dijkstra, *A note on two problems in connection with graphs*, Num. Math., 1 (1959), 269-271.

[27] P. Erdös', A. R'enyi and V. T. S'oc, *On a problem of graph theory*, Studia. Sci. Math. Hungar. 1 (1966), 215-235.

[28] P. Erdös', M. Simonovits, *Compactness results in extremal graph theory*, Combinatorica 2 (3), 1982, 275-288.

[29] W. Faudree, M. Simonovits, *On a class of degenerate extremal graph problems*, Combinatorica 3 (1), 1983, 83-93.

[30] W. Feit, D. Higman *The nonexistence of certain generalised polygons*, J. of Algebra 1 (1964), 114-131.

[31] V. Futorny, V. Ustimenko, *On Small World Semiplanes with Generalised Schubert Cells*, Acta Applicandae Mathematicae, N4, 2007.

[32] I. Gelfand, R. MacPherson, *Geometry in Grassmanians and generalisation of the dilogarithm*, Adv. in Math., 44 (1982), 279-312.

[33] I. Gelfand, V. Serganova, *Combinatorial geometries and torus strata on homogeneous compact manifolds*, Russ. Math. Surv. 42 (1987), 133-168.

[34] W. C. Huffman and V. Pless, *Fundamentals of Error-Correcting Codes*, Cambridge University Press, 2003.

[35] V. Kac. *Infinite dimensional Lie algebras*, Birkhauser, Boston, 1983.

[36] W. Kantor, *Linear groups containing a Singer cycle*, J. of Algebra 62, 1982, 232-234.

[37] Yu Khmelevsky, V. Ustimenko, *Practical aspects of the Informational Systems reengineering*, The South Pacific Journal of Natural Science, volume 21, 2003, p.75-21 (www.usp.ac.fj/spjns/volume21)

[38] Yu.Khmelevsky, M. Govorov, V. Ustimenko, P. Sharma, S. Dhanjal, *Security Solutions for Spatial Data in Storage (Implementation Case within Oracle 9iAS)*, Proceedings of 8th World Multiconference on Systemics, Cybernetics and Informatics (SCI 2004) Orlando, USA, in July 18-21, 2004, pp 318-323.

[39] Jon-Lark Kim, U. N. Peled, I. Perepelitsa, V. Pless, S. Friedland, *Explicit construction of families of LDPC codes with no 4-cycles* , Information Theory, IEEE Transactions, 2004, v. 50, Issue 10, 2378 - 2388.

[40] M. Klisowski, V. Ustimenko, *On the public keys based on the extremal graphs and digraphs*, International Multiconference on Computer Science and Informational Technology, October 2010, Wisla, Poland, CANA Proceedings.

[41] R. G. Gallager, *Low-density parity-check codes*, IRE Transactions on Information Theory, vol. IT-8, pp. 21-28, Jan. 1962.

[42] M. Gari, D. Johnson, *Computers and Intractability, A Guide to the Theory of NP-Completeness*, Freeman, 1979.

[43] M. Govorov, Yu Khmelevsky, V. Ustimenko, A. Khorev, *Security Control for Spatial Warehouses*, Proceedings ofthe 21th International Cartographic Conference (ICC), Durban, South Africa, 2003, 1784-1794.

[44] M. Govorov, Yu Khmelevsky, V. Ustimenko, A. Khorev, *Security for GIS N-tier Architecture. Developments in Spatial Data Handling*, in 11thInternationalSymposium on Spatial Data Handling (editor P. Fisher), Springer, 2005, pp 71-85.

[45] , P. Guinand and J. Lodge, *Tanner Type Codes Arising from Large Girth Graphs*, Proceedings of the 1997 Canadian Workshop on Information Theory (CWIT '97), Toronto, Ontario, Canada, pp. 5-7, June 3-6, 1997.

[46] P. Guinand and J. Lodge, *Graph Theoretic Construction of Generalized Product Codes*, Proceedings of the 1997 IEEE International Symposium on Information Theory (ISIT '97), Ulm, Germany, p. 111, June 29-July 4, 1997.

[47] S. Hoory, N. Linial, and A.Wigderson, *Expander graphs and their applications*, Bulletin (New Series) of AMS, volume 43, N4, 439-461,

[48] W. Imrich, *Explicit construction of graphs without small cycles*, Combinatorica **2** (1984) 53–59.

[49] J. P. Jones, D. Sato, H. Wada and D. Wiens, *Diophantine representation of the set of prime numbers*, Amer. Math. Monthly, 83 (1976) 449-464.

[50] S. Karlin, H.M. Taylor, *A first course in stochastic processes*, Academic Press, New York, 1975.

[51] Yu Khmelevsky, V Ustimenko, *Practical aspects of the Informational Systems reengineering*, The South Pacific Journal of Natural Science, volume 21, 2003, p.75-21 (www.usp.ac.fj/spjns/volume21).

[52] N. Koblitz, *A Course in Number Theory and Cryptography*, Second Edition, Springer, 1994, 237 p.

[53] N. Koblitz, *Algebraic aspects of Cryptography*, in Algorithms and Computations in Mathematics, v. 3, Springer, 1998.

[54] J. Kotorowicz, V. A. 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.

[55] I. Kovalenko, *The survey of my scientific works, Teachers and colleagues*, Cybernetics and systems analysis, Springer, vol. 3, 2010, 3-27.

[56] F. Lazebnik, V. A. Ustimenko, *New Examples of graphs without small cycles and of large size*, Europ. J. of Combinatorics, 14 (1993) 445-460.

[57] F. Lazebnik and V. Ustimenko, *Some Algebraic Constructions of Dense Graphs of Large Girth and of Large Size*, DIMACS series in Discrete Mathematics and Theoretical Computer Science, V. 10 (1993), 75-93.

[58] F. Lazebnik, V. A. Ustimenko and A. J. Woldar, *Properties of certain families of $2k$-cycle free graphs*, J. Combin. Theory, ser B, 60, No. 2 (1994), 293-298.

[59] F. Lazebnik, V. Ustimenko, *Explicit construction of graphs with an arbitrary large girth and of large size*, Discrete Appl. Math. , 60, (1995), 275 - 284.

[60] F. Lazebnik, V. A. Ustimenko and A. J. Woldar, *Properties of certain families of $2k$-cycle free graphs*, J. Combin. Theory, ser B, 60, No. 2 (1994), 293-298.

[61] F. Lazebnik, V. A. Ustimenko and A. J. Woldar, *A New Series of Dense Graphs of High Girth*, Bull (New Series) of AMS, v.32, N1, (1995), 73-79.

[62] Lazebnik, F., Ustimenko, V.A. and A.J. Woldar, *A characterisation of the components of the graph $D(k, q)$*, Discrete Mathematics, 157 (1996), 271-283.

[63] Lazebnik, F., Ustimenko, V.A. and A.J. Woldar, *New upper bounds on the order of cages*, Electronic J. Combin. 14 R13 (1997), 1–11.

[64] F. Lazebnik, V. A. Ustimenko and A. J. Woldar, *Polarities and $2k$-cycle-free graphs*, Discrete Mathematics, 197/198, (1999), 503–513.

[65] Arijen K. Lenstra, *Factoring multivariable polynomials over finite field*, Proceedings of the fifteenth annual ACM symposium on Theory of Computing, 1983, 189-192.

[66] A. Lubotsky, R. Philips, P. Sarnak, *Ramanujan graphs*, J. Comb. Theory., 115, N 2., (1989), 62-89.

[67] D. J. C. MacKay and R. N. Neal, *Good Codes based on very sparse matrices*, In Cryptography and Coding, 5th IMA Conference, Lecture Notes in Computer Science, v. 1025, 1995, pp. 110-111.

[68] W. Magnus, A. Karras, D. Solitar, *Combinatorial group theory*, Interscience publ., 1966.

[69] G. A. Margulis, *Explicit construction of graphs without short cycles and low density codes*, Combinatorica, 2, (1982), 71-78.

[70] G. Margulis, *Explicit group-theoretical constructions of combinatorial schemes and their application to desighn of expanders and concentrators*, Probl. Peredachi Informatsii., 24, N1, 51-60. English translation publ. Journal of Problems of Information transmission (1988), 39-46.

[71] M. Margulis, *Arithmetic groups and graphs without short cycles*, 6th Intern. Symp. on Information Theory, Tashkent, abstracts, vol. 1, 1984, pp. 123-125 (in Russion).

[72] Y. V. Matijasevic, *A Diophantine representation of the set of prime numbers* (in Russian), Dokl. Akad. Nauk SSSR, 196 (1971) 770–773. English translation by R. N. Goss, in Soviet Math. Dokl., 12, 1971, 249-254.

[73] Y. V. Matijasevic, *Primes are enumerated by a polynomial in 10 variables* (in Russian), Zapiski Sem. Leningrad Mat. Inst. Steklov, 68 (1977) 62–82, 144–145. English translation by L. Guy and J. P. Jones, J. Soviet Math., 15, 1981, 33–44.

[74] H. L. Montgomery, *Topics in Multiplicative Number Theory*, Lecture Notes in Mathematics 227, Springer Verlag, New York, 1971.

[75] E. H. Moore, *Tactical Memoranda*, Amer. J. Math., v.18, 1886, 264-303.

[76] B. Mortimer, *Permutation groups containing affine of the same degree*, J. London Math. Soc., 1971, 15, N3, 445-455.

[77] Jose M. F. Moura, Jin Lu, and Haotian Zhang, *Structured LDPC Codes with Large Girth*, IEEE Signal Processing Magazine, vol. 21:1, pp.42-55,

January 2004. Included in Special Issue on Iterative Signal Processing for Communications.

[78] H. Niederreiter, Chaoping Xing, *Algebraic Geometry in Coding Theory and Cryptography*, Princeton University Press, 2009.

[79] R. Ore, *Graph Theory*, Wiley, London, 1971.

[80] J. Patarin, *Cryptoanalysis of the Matsumoto and Imai public key scheme of the Eurocrypt '88*, Advances in Cryptology, Eurocrypt '96, Springer Verlag, 43-56.

[81] P. Ribenboim, *The new book of prime number records*, 3rd edition, Springer-Verlag, New York, NY, 541 p., ISBN 0-387-94457-5. 1995.

[82] T. Richardson, R. Urbanke, *Modern Coding Theory*, Cambridge University Press, 2008.

[83] U. Romanczuk, V. Ustimenko, *On the key exchange with matrices of large order and graph based nonlinear maps*, Proceedings of the conference "Applications of Computer Algebra", Vlora, 2010 (to appear)

[84] H. Sachs, Regular graphs with given girth and restricted circuits, J. London. Math. Soc. 38 (1963), 423-429.

[85] N. Sauer. *Extermaleigenschaften regularer Graphen gegebener Taillenweite*, 1, 2, Osterreich. Acad. Wiss. Math. Natur. Kl. S. -B 2, 176 (1967), 9-25, 27-43.

[86] J. Seberry, J. Pieprzyk, *Cryptography: An Introducion to Computer Security*, Prentice Hall 1989, 379 p.

[87] J. P. Serre, *Lie Algebras and Lie groups*, N. Y., Lectures in Math., Springer, Berlin, 1974.

[88] T. Shaska , W C Huffman, D. Joyner, V Ustimenko (Editors), *Advances in Coding Theory and Crytography*, (Series on Coding Theory and Cryptology), World Scientific Publishing Company, 2007.

[89] T. Shaska, V. Ustimenko, *On the homogeneous algebraic graphs of large girth and their applications*, Linear Algebra and its Applications Article, Volume 430, Issue 7, 1 April 2009, Special Issue in Honor of Thomas J. Laffey.

[90] T. Shaska and V. Ustimenko, *On some applications of graph theory to cryptography and turbocoding*, Special issue of Albanian Journal of Mathematics:Proceedings of the NATO Advanced Studies Institute "New challenges in digital communications", May 2008, University of Vlora, 2008, v.2, issue 3, 249-255.

[91] M. Simonovitz, *Extermal Graph Theory* , In "Selected Topics in Graph Theory", 2, edited by L. W. Beineke and R. J. Wilson, Academic Press, London, 1983, pp. 161-200.

[92] V, Suschanskij, F. Lazebnik, V. Ustimenko, M. Klin, R. Poschel and V. Vyshenskij, *Lev Arkad'evich Kaluznin (1914 - 1990)*, Acta Applicandae Mathematicae, vol. 52, 5-18.

[93] Takashi Soma, V. Ustimenko, *Graph Laplacians and Fourier Transformations on Boolean*, CITR TR - 82, Tech Reports of the Center of Informational Tech. and Robotics (electronic journal), The University of Auckland, 2001, 8p.

[94] R. Michiel Tanner, *A recursive approach to low density codes*, IEEE Trans. on Info Th., IT, 27(5):533-547, Sept.1984.

[95] J. A. Thas, *Generalised polygons*, in F. Buekenhout (ed), Handbook in Incidence Geometry, Ch. 9, North Holland, Amsterdam, 1995.

[96] J. Tits, *Sur la trialite at certains groupes qui s'en deduicent*, Publ. Math. I.H.E.S. 2 (1959), 15-20.

[97] J. Tits, *Les groupes simples de Suzuki et de Ree*, Seminaire Bourbaki 13 (210), 1960/1961, 1-18.

[98] J. Tits, *Buildings of spherical type and Finite BN-pairs, Lecture Notes in Math*, Springer Verlag, 1074.

[99] A. Touzene, V. Ustimenko, *Graph Based Private Key Crypto System*, International Journal on Computer Research, Nova Science Publisher, volume 13 (2006), issue 4, 12p.

[100] A. Touzene, V. Ustimenko,*Private and Public Key Systems Using Graphs of High Girth*,In "Cryptography Research Perspectives", Nova Publishers, Ronald E. Chen (the editor), 2008, pp. 205-216.

[101] W. Tutte, *A family of cubical graphs*, Proc. Cambridge Philos. Soc. 43 (1945).

[102] V. A. Ustimenko, *On some properties of Chevalley groups and their generalisations*, In: Investigations in Algebraic Theory of Combinatorial objects, Moskow, Institute of System Studies, 1985, 134 - 138 (in Russian), Engl.trans.: Kluwer, Dordrecht, 1992, pp. 112-119

[103] V. Ustimenko, *On the embeddings of some geometries and flag systems in Lie algebras and superalgebras*, in "Root systems, representations and geometries", Kiev, IM AN UkrSSR, pp. 3–16, 1990.

[104] V. A. Ustimenko, *Geometries of twisted simple groups of Lie type as objects of linear algebra*, in Questions of Group Theory and Homological Algebra, University of Jaroslavl, Jaroslavl, 1990, 33-56 (in Russian).

[105] V. A. Ustimenko, *Linear interpretation of Chevalley group flag geometries*, Ukraine Math. J. 43, Nos. 7,8 (1991), pp. 1055–1060 (in Russian).

[106] V. Ustimenko, *Small Schubert cells as subsets in Lie algebras*, Functional Analysis and Applications, v. 25, no. 4, 1991, pp. 81–83.

[107] V. A. Ustimenko, *Coordinatisation of regular tree and its quotients*, in "Voronoi's impact on modern science", eds P. Engel and H. Syta, book 2, National Acad. of Sci, Institute of Matematics, 1998, 228p.

[108] V. A. Ustimenko, *On the Varieties of Parabolic Subgroups, their Generalizations and Combinatorial Applications*, Acta Applicandae Mathematicae 52 (1998): pp. 223–238.

[109] V. Ustimenko, *CRYPTIM: Graphs as Tools for Symmetric Encryption*, in Lecture Notes in Computer Science, Springer, 2001, v. 2227, 278-287.

[110] V. A. Ustimenko, *Graphs with Special Arcs and Cryptography*, Acta Applicandae Mathematicae, vol. 71, N2, November 2002, 117-153.

[111] V. Ustimenko, *Maximality of affine group and hidden graph cryptsystems*, Journal of Algebra and Discrete Mathematics, October, 2004, v.10, pp. 51-65.

[112] V. Ustimenko, *Small world graphs with memory and Coxeter groups*, technical report 110/05 of the Centre of Mathematical Sciences, Madeira University, Portugal, July, 2005, 12 p.

[113] V. A. Ustimenko, *On the graph based cryptography and symbolic computa-*

tions, Serdica Journal of Computing, Proceedings of International Conference on Application of Computer Algebra, ACA-2006, Varna, N1 (2007).

[114] 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, vol. 3, 181-200 (2007).

[115] V. Ustimenko, *On the Cryptography with "Mathematica package"*, Proceedings of the conference - Leaning Mathematics and Technology Middle East Conference, University of Arizona and Sultan Qaboos University, Oman, March, 2007, 11 p.

[116] V. A. Ustimenko, *On the extremal regular directed graphs without commutative diagrams and their applications in coding theory and cryptography*, Albanian. J. of Mathematics, Special Issue "Algebra and Computational Algebraic Geometry", vol. 1, N4, 387-400, 2007.

[117] V. A. Ustimenko, *On the hidden discrete logarithm for some polynomial stream ciphers*, International Multiconference on Computer Science and Informational Technology, 20-22 October 2008, Wisla, Poland, CANA Proceedings.

[118] V. A. Ustimenko, *Linguistic Dynamical Systems, Graphs of Large Girth and Cryptography*, Journal of Mathematical Sciences, Springer, vol.140, N3 (2007) pp. 412-434.

[119] V. A. Ustimenko, *On the cryptographical properties of extremal 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.

[120] V. Ustimenko, Yu. Khmelevsky, *Walks on graphs as symmetric and asymmetric tools for encryption*, 2002, South Pacific Journal of Natural Studies, 2002, vol. 20, 23-41 (www.usp.ac.fj/spjns).

[121] V. A. Ustimenko, J. Kotorowicz, *On the properties of Stream Ciphers Based on Extremal Directed graphs*, In "Cryptography Research Perspectives", Nova Publishers, Ronald E. Chen (the editor), 2008.

[122] V. Ustimenko, D. Sharma, *CRYPTIM: system to encrypt text and image data*, Proceedings of International ICSC Congress on Intelligent Systems, Wollongong, 2001, 11pp.

[123] V. Ustimenko, A. Woldar, *Extremal properties of regular and affine generalised polygons of tactical configurations*, European Journal of Combinatorics, 24 (2003) 99–111.

[124] Gilles Van Assche, *Quantum Cryptography and Secret-Key Distillation*, Cambridge University Press, 2006.

[125] H. Walther, Uber regulare Graphen gegebener Taillenweite und inimaler Knotenzahl, *Wiss. Z. Techn Hochsch. Ilmenau* 11 (1965) 93–96.

[126] A. L. Weiss, Girth of bipartite sextet graphs, *Combinatorika* 4 (no. 2-3) (1984) 241–245.

[127] R. Wenger, *Extremal graphs with no C^4, C^6 and $C^{10}s$*, 1991, J. Comb. Theory, Ser B, 52, 113-116.

[128] A. Wroblewska *On some properties of graph based public keys* , Albanian Journal of Mathematics, Volume 2, Number 3, 2008, 229-234.