On new results on extremal graph theory, theory of algebraic graphs, and their applications

Presented by Corresponding Member of the NAS of Ukraine O.M. Trofimchuk

New explicit constructions of infinite families of finite small world graphs of large girth with well-defined projective limits which is an infinite tree are described. The applications of these objects to constructions of LDPC codes and cryptographic algorithms are shortly observed. We define families of homogeneous algebraic graphs of large girth over the commutative ring $K$. For each commutative integrity ring $K$ with $|K| > 2$, we introduce a family of bipartite homogeneous algebraic graphs of large girth over $K$ formed by graphs with sets of points and lines isomorphic to $K^n$, $n > 1$, and cycle indicator $\geq 2n + 2$ such that their projective limit is well defined and isomorphic to an infinite forest.

Keywords: family of graphs of large girth, small world graphs, cryptographic algorithms, LDPC codes.

The girth and diameter of a graph are the minimal length of its cycle and the maximal distance of the graph. We can consider the girth indicator $\text{Cind}(v)$ of a vertex $v$ of the graph $\Gamma$ as the minimal length of the cycle through $v$ and introduce a cycle indicator $\text{Cind}(\Gamma)$ of the graph as the maximal value of $\text{Cind}(v)$ for its vertices.

The constructions of finite or infinite graphs with prescribed girth and diameter is an important and difficult task of the graph theory. Noteworthy that the incidence of the classical projective geometry over various fields is a graph of girth 6 and diameter 3. J. Tits defined generalized $m$-gons as bipartite graphs of girth $2m$ and diameter $m$. Feit and Higman proved that finite generalized $m$-gons with bi-degrees $> 2$ exist only in the cases of $m = 3, 4, 6, 8,$ and 12. Geometries of finite simple groups of rank 2 are natural examples of generalized $m$-gons for $m = 3, 4, 6, 8$. Classification of flag transitive generalized $m$-gons of the Moufang type were obtained by J. Tits and R. Weiss.

Infinite families of graphs of large girth of bounded degree are important objects of extremal graph theory which were introduced by P. Erdős'. He proved the existence of such families via his well-known probabilistic method. Nowadays, a few explicit constructions of such families are known. The concept of an infinite family of small world graphs of bounded degree turns out to be very important for various applications of graph theory.
Noteworthy that only one family of small world graphs of large girth is known. This is the family $X(p, q)$ of Ramanujan graphs introduced by G. Margulis [1] and investigated via the computation of their girth, diameter, and the second largest eigenvalue by A. Lubotsky, R. Phillips, and P. Sarnak [2].

We have to admit that studies of families of graphs $\Gamma_i$ with well-defined projective limit $\Gamma$, which is isomorphic to an infinite tree, is well motivated.

We refer to such family as a tree approximation. There is only one tree approximation by finite graphs which is a family of large girth. This is the family of $CD(n, q)$ defined by F. Lazebnik, V. Ustimenko, and A. Woldar [3]. The question whether or not $CD(n, q)$ form a family of small world graphs has been still open since 1995.

In 2013, the tree approximation by finite graphs $A(n, q)$ which is a family of small world graphs was presented (see [4]). It was proven that the graph from the family has maximal possible cycle indicator (in fact, $\text{Cind}(A(n, q)) = 2n + 2$).

One of the main statements of this paper is $A(n, q)$, where $n = 2, 3, \ldots$, is a family of large girth. We generalize these results in terms of the theory of algebraic graphs defined over an arbitrary field and consider properties and applications of the above-mentioned graphs.

1. Case of finite simple graphs. All graphs we consider are simple, i. e., undirected without loops and multiple edges. Let $V(\Gamma)$ and $E(\Gamma)$ denote the set of vertices and the set of edges of $\Gamma$, respectively. The parameter $|V(\Gamma)|$ is called the order of $\Gamma$, and $|E(\Gamma)|$ is called the size of $\Gamma$. A path in $\Gamma$ is called simple, if all its vertices are distinct. When it’s convenient, we shall identify $\Gamma$ with the corresponding antireflexive binary relation on $V(\Gamma)$, i. e., $E(\Gamma)$ is a subset of $V(\Gamma) \times V(\Gamma)$. The length of a path is a number of its edges. The girth of a graph $\Gamma$, denoted by $g(\Gamma)$, is the length of the shortest cycle in $\Gamma$. Let $k \geq 3$ and $g \geq 3$ be integers. The distance between vertices $v$ and $u$ of the graph $\Gamma$ is a minimal length of the path between them. The diameter of the graph is maximal distance between its vertices.

The graph is connected, if its diameter is finite. The graph is $k$-regular, if each vertex of the graph is incident exactly to $k$ other vertices. The tree is a connected graph which does not contain a cycles.

1. An infinite family of simple regular graphs $\Gamma_i$ of constant degree $k$ and order $v_i$ such that $\text{diam}(\Gamma_i) \leq c \log_{k-1}(v_i)$, where $c$ is the constant independent of $i$ and $\text{diam}(\Gamma_i)$ is the diameter of $\Gamma_i$, is called a family of small world graphs.

2. Recall that the infinite families of simple regular graphs $\Gamma_i$ of constant degree $k$ and order $v_i$ such that $g(\Gamma_i) \geq c \log_{k-1}(v_i)$, where $c$ is the constant independent of $i$ and $g(\Gamma_i)$ is a girth of $\Gamma_i$, are called families of graphs of large girth.

Let $\Gamma$ be a simple graph. Assume that $\text{Cind}(x)$ is the minimal length of a cycle through the vertex $x$ of the graph $\Gamma$. Let $\text{Cind}(\Gamma)$ stand for the maximal value of $\text{Cind}(x)$ via all vertices $x$ of $\Gamma$. We refer to parameter $\text{Cind}(\Gamma)$ as a cycle indicator of $\Gamma$.

One of the main purposes of the paper is to present a special interpretations of $q$-regular tree ($q$-regular simple graph without cycles) in terms of the algebraic geometry over a finite field $F_q$.

**Theorem 1.** For each prime power $q$, $q > 2$ there is a family of $q$-regular graphs $\Gamma_i$ satisfying the following properties:

(i) $\Gamma_i$ is a family of small world graphs;

(ii) $\Gamma_i$ is a family of large girth;
(iii) Projective limit of graphs $\Gamma_i$ is well defined and coincides with the $q$-regular tree $T_q$; 
(iv) $\text{Cind} \Gamma_i = 2 \log_q(v_i/2) + 2$.

We refer to the family of graphs $\Gamma_i$ satisfying condition (iii) as a tree approximation.

The proof of Theorem 1 is given via the explicit construction of graphs $\Gamma_i = A(i, q), i \geq 2$ satisfying requirements of the statement. Noteworthy that $A(i, q)$ is the unique known example of the family satisfying conditions (i), (ii), an (iii).

In fact, there is exactly one other known construction of the $q$ regular family satisfying (i) and (ii), i.e., the explicit construction of the family of regular simple small world graphs of large girth and with an arbitrarily large degree $q$.

This family $X(p, q)$ formed Cayley graphs for $\text{PSL}_2(p)$, where $p$ and $q$ are primes, had been defined by G. Margulis [1] and investigated by A. Lubotzky, Sarnak and Phillips [2]. As it is easy to see, the projective limit of $X(p, q)$ does not exist.

**The construction of $A(n, q)$.** Let $K$ be a finite field $F_q$. We define $A(n, K) = A(n, q)$ as a bipartite graph with the point set $P = K^n$ and line set $L = K^n$ (two copies of a Cartesian power of $K$ are used). We will use brackets and parenthesis to distinguish tuples from elements constructed from a vector space over $K$.

So $(p_i) = (p_1, p_2, ..., p_n)$ $\in P$, and $[l_i] = [l_1, l_2, ..., l_n] \in L$.

The incidence relation $I = A(n, K)$ (or the corresponding bipartite graph $I$) is given by condition $p$ and $l$, if and only if the equations of the following kind hold:

$$p_2 - l_2 = l_1 p_1,$$
$$p_3 - l_3 = l_1 l_2,$$
$$p_4 - l_4 = l_1 p_3,$$
$$p_5 - l_5 = l_1 l_4,$$
$$...,$$
$$p_n - l_n = l_1 l_{n-1} \text{ for odd } n \text{ and } p_n - l_n = l_1 p_{n-1} \text{ for even } n.$$

We can consider an infinite bipartite graph $A(K)$ with points $(p_1, p_2, ..., p_n, ...)$ and lines $[l_1, l_2, ..., l_n, ...]$.

**Proposition 1** [4]. If $K = F_q$, $q > 2$, then $A(n, F_q)$ is a family of small world graphs and a tree approximation with $\text{Cind}(A(n, F_q)) = 2n + 2$.

Let $K$ be an arbitrary field. We define $A(n, K)$ via a simple change of $F_q$ on $K$ and announce the following statement.

**Proposition 2.** Let $K$ be a field. Then the girth of $A(n, K)$ is $\geq 2 \lceil n/2 \rceil + 2$.

Symbol $\lceil x \rceil$ stands for the flow function from $x$. Theorem 1 follows from Propositions 1 and 2.

2. **Case of algebraic graphs.** 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 $\phi$ over $F$ consists of two things: the vertex set $Q$ being a quasiprojective variety over $F$ of non-zero dimension and the edge set being a quasiprojective variety $\phi$ in $Q \times Q$ such that $(x, x)$ is not element of $\phi$ for each $x \in Q$ and $x\phi y$ implies $y\phi x$ ($x\phi y$ means $(x, y) \in \phi$). The graph $\phi$ 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 a non-zero dimension. We further assume that $M$ contains at least 3 elements.
Theorem [5]. Let \( \Gamma \) be the homogeneous algebraic graph over a field \( F \) of girth \( g \) such that the dimension of a neighborhood for each vertex is \( N, N \geq 1 \). Then \( [(g - 1)/2] \leq \text{dim}(V)/N \).

The following corollary is an analog of Even Circuit Theorem by Erdős for finite simple graphs.

**Corollary.** Let \( \Gamma \) be a homogeneous graph over a field \( F \), and let \( E(\Gamma) \) be a variety of its edges. Then \( \text{dim}(E(\Gamma)) \leq \text{dim}(V)(1 + [(g - 1)/2]^{-1} \).

We announce a stronger statement.

**Theorem 2.** Let \( \Gamma \) be a homogeneous algebraic graph over the field \( F \) with cycle indicator \( z \) such that the dimension of a neighborhood for each vertex is \( N, N \geq 1 \). Then \( [(z - 1)/2] \leq \text{dim}(V)/N \).

We refer to the family of homogeneous algebraic graphs \( \varphi_n \) for which the dimension of a neighborhood for each vertex is an independent constant \( N, N \geq 1 \) as a family of *small world graphs*, if the diameter of each graph \( \varphi_n \) is bounded from above by the linear function \( \alpha n + \beta \) defined by constants \( \alpha \) and \( \beta \).

We refer to the family of homogeneous algebraic graphs \( \varphi_n \) for which the dimension of a neighborhood for each vertex is an independent constant \( N, N \geq 1 \) as a *family of large girth*, if the girth of each graph \( \varphi_n \) is bounded from below by the linear function \( \alpha n + \beta \) defined by constants \( \alpha \) and \( \beta \).

We refer to the homogeneous algebraic graph as an algebraic forest, if it does not contain cycles. The term an algebraic tree stands for the connected algebraic forest.

We say that the family of homogeneous algebraic graphs \( \varphi_n \) is a forest (tree) approximation, if the projective limit of \( \varphi_n \) is an algebraic forest (tree) and formulate the following statement.

**Theorem 3.** For each field \( F, F \neq F_2 \) there exists a tree approximation which is a family \( \varphi_n \) of small world algebraic graphs of large girth with the vertex set of the dimension \( n \) and cycle indicator \( 2n + 2 \).

The family of graphs \( \varphi_n = A(n, F) \) provides the explicit construction of objects described in the theorem. As it follows from Theorem 2, the homogeneous algebraic graphs \( A(n, F) \) form a family with maximal possible girth indicator.

**Remark 1.** The graphs \( A(n, F_2) \) are disconnected. So, they are a disjoint union of cycles. The graph \( A(F_2) \) is a 2-regular forests with trees presented on the following diagram: …. ----*-----* -----*--- .... .

The girth indicator of \( A(n, F_2) \) coincides with its girth. So, \( A(n, 2) \) formally are algebraic graphs of large girth.

Noteworthy that cycles can be defined via a system of equations.

3. **Some properties of \( A(n,K) \) and some their applications.** The graphs \( A(n, q) \) obtained as special homomorphic images (see [9]) of graphs \( D(n, q) \) (see [7]) which defines the projective limit \( D(q) \) with points

\[
(p) = (p_{01}, p_{11}, p_{12}, p_{21}, p_{22}, p'_{22}, ..., p'_{ii}, p_{i,i+1}, p_{i+1,i}, p_{i+1,i+1}, ...)
\]

lines

\[
[l] = [l_{19}, l_{11}, l_{12}, l_{21}, l_{22}, l'_{22}, ..., l'_{ii}, l_{i,i+1}, l_{i+1,i}, l_{i+1,i+1}, ...]
\]
and the incidence relation given by the equations

\begin{align*}
    l_{ii} - p_{ii} &= l_{i0} p_{i=1,j}; \\
    l'_{ii} - p'_{ii} &= l'_{i,j-1} p_{01}; \\
    l_{i,i} + p_{i,i+1} &= l_{ii} p_{01}; \\
    l_{i+1,i} - p_{i+1,j} &= l_{10} p'_{ii}.
\end{align*}

This four relations are defined for $I \geq 1$ ($p'_{11} = p_{11}, l'_{11} = l_{11}$).

Remark 2. You can see that the indices of vectors correspond to the coordinates of positive roots of the root system $A_1$ with a wave.

Historically, the graph $D(q)$ is not the first example of the description of a $q$-regular forest in terms of Algebraic Geometry. The geometries of buildings corresponding to the extended Dynkin diagram $A_1$ as incidence structures are $q + 1$-regular trees or $q + 1$-regular forests (see [6]). As a result, we get the description of a tree in group theoretical terms.

In [6], it was noticed that the restriction of this incidence relation on orbits of the Borel subgroup $B'$ acting on maximal parabolas are $q$-regular bipartite graphs. So, we get the description of a $q$-regular tree in terms of positive roots of $A_1$ with a wave.

In [7], the authors proved that $D(n, q)$ defined via the first $n - 1$ equations of $D(q)$ form a family of graphs of large girth.

Unexpectedly, we discover that these graphs are disconnected, if $n \geq 6$. So, the forest $D(q)$ contains infinitely many trees, and the diameter is infinite. F. Lazebnik conjectured that the connected components of graphs $D(n, q)$, $n = 3, 4, \ldots$ form a family of small world graphs. This conjecture is still open.

In 1994, were discovered how to define the connected components $CD(n, q)$ of graphs $D(n, q)$ in terms of equations (see [8]).

The graphs $A(n, q)$ were obtained in 2007 as the homomorphic images of graphs $D(n, q)$ ([9]). The corresponding homomorphism $\eta$ is a procedure to delete the coordinates of points and lines with indices $(i + 1, i)$ and $(i, i)'$.

The self importance of these graphs has been justified in my joint research with U. Romanczuk (see [10] and further references) and M. Polak [11] via applications to Cryptography and constructions of Low Density Parity Check (LDPC) Codes (see [12]).

In the case of the families of graphs of large girth, we would like to have “the rate of growth” $c$ of the girth “as high as possible”. P. Erdos' proved the existence of such a family with arbitrary large but bounded degree $k$ with $c = 1/4$ by his probabilistic method.

In the case of families $X(p, q)$ and $CD(n, q)$, the constant $c$ is $4/3$. In the case of $A(n, q)$, we just get the inequality $1 \leq c < 2$. So, the exact computation of the girth is the area of the future research.

There are essential differences between the family of graphs $X(p, q)$ and tree approximations. Recall that the projective limit of $X(p, q)$ does not exist.

We prove that the bipartite graphs $A(n, q)$ are not edge-transitive and not vertex-transitive (transitivity on points and intransitivity on lines). Noteworthy that their the projective limit $T$ (the tree) is obviously an edge-transitive infinite graph.
The usage of generalizations and modifications of the graphs $A(n, q)$ allows us to construct postquantum cryptosystem of the El Gamal type with encryption procedure for a potentially infinite vector from $F_q$ with the execution speed $O(n^{1+2/n})$ (see [13] and further references).

In fact, the diameter of $A(n, q)$ is growing slower than the diameter of $X(p, q)$. So, $A(n, q)$ are the best known small world graphs among the known families of large girth. Recall that the girth of $A(n, q)$ is not yet computed precisely.

So, the comparison of the growth of the girth for $A(n, q)$ and $X(p, q)$ is the interesting task for the future research.

In the case of finite fields, both families are expanding graphs, the second largest eigenvalue of $A(n, q)$ tends to $2q^{1/2}$, they are not Ramanujan graphs for which the second largest eigenvalue has to be bounded above by $2(q - 1)^{1/2}$.

The family $X(p, q)$ is formed by Ramanujan graphs; so, they are better expanding graphs than $A(n, K)$.

The families $X(p, q)$, $CD(n, q)$, and $A(n, q)$ can be used for the constructions of LDPC codes for the noise protection in satellite communications. D. MacKay and M. Postol [12] proved that $CD(n, q)$ based LDPC codes have better properties than those from $X(p, q)$ for the constructions of LDPC codes.

Together with Monika Polak, we proved that $A(n, q)$ based LDPC codes even better than those from $CD(n, q)$ (see [11]).

The Cayley nature of $X(p, q)$ does not allow one to use these graphs in cryptography. Various applications of graphs $D(n, q)$, $CD(n, q)$, and $A(n, q)$ have been known since 1998.

The most recent postquantum cryptosystem based on noncommutative multivariate group associated with $A(n, q)$ was described in [13], IACR e-print Archive 2021/1466. This cryptosystem will be presented at ICM-2022 conference “Mathematical Aspects of Post-quantum Cryptography”.

This algorithm can be used for the encryption of potentially infinite vectors from $(F_q)^n$ in time $O(n^{1+2/n})$. So, it can work with Big Data files.

4. The case of integrity rings. Let $K$ stand for an arbitrary commutative ring. Noteworthy that the graphs $A(n, K)$ over an arbitrary commutative ring $K$ have been already defined. The graphs $D(k, K)$ over $K$ were considered in [14], where the graphs $CD(k, K)$ with $k \geq 6$ were introduced as induced subgraphs of $D(k, K)$ with vertices $u$ satisfying special equations

$$a_2(u) = 0, a_3(u) = 0, ..., a_t(u) = 0, t = [(k + 2)/4],$$

where $u = (u_{a_1}, u_{a_2}, u_{a_3}, ..., u_{a_t}, u'_{r_1}, u'_{r_2}, u'_{r_3}, u'_{r_4}, u'_{r_5}, u'_{r_6}, ...)$, $2 \leq r \leq t, a \in \{(1, 0), (0, 1)\}$ is a vertex of $D(k, K)$, and

$$a_r = a_r(u) = \sum_{i=0, r} (u_{a_1} - u_{a_2} + u_{a_3} - u_{a_4} + u_{a_5} - u_{a_6} + ...),$$

for every $r$ from the interval $[2, t]$.

We set $a = a(u) = (a_2, a_3, ..., a_t)$ and assume that $D(k, K) = CD(k, K)$ if $k = 2, 3, 4, 5$.

As was proven in [9], the graphs $D(n, K)$ are edge-transitive. So, their connected components are isomorphic graphs. Let $\gamma CD(k, K)$ be a solution set of the system of equations

$$a_2(u) = 0, a_3(u) = 0, ..., a_t(u) = 0, t = [(k + 2)/4],$$

where $u = (u_{a_1}, u_{a_2}, u_{a_3}, ..., u_{a_t}, u'_{r_1}, u'_{r_2}, u'_{r_3}, u'_{r_4}, u'_{r_5}, u'_{r_6}, ...)$, $2 \leq r \leq t, a \in \{(1, 0), (0, 1)\}$ is a vertex of $D(k, K)$, and

$$a_r = a_r(u) = \sum_{i=0, r} (u_{a_1} - u_{a_2} + u_{a_3} - u_{a_4} + u_{a_5} - u_{a_6} + ...),$$

for every $r$ from the interval $[2, t]$.

We set $a = a(u) = (a_2, a_3, ..., a_t)$ and assume that $D(k, K) = CD(k, K)$ if $k = 2, 3, 4, 5$.
On new results on extremal graph theory, theory of algebraic graphs, and their applications

$a(u) = (v_2, v_3, ..., v_t) = v$ for certain $v \in K^{t-1}$. It is proven that each $vCD(k, K)$ is the disjoint union of some connected components of the graph $D(n, K)$.

It is easy to see that the sets of vertices of $vCD(k, K), v \in K^{t-1}$ form a partition of the vertex set of $D(n, K)$.

The concept of quasiprojective variety over the commutative ring $K$ can be introduced via the simple substitution of $K$ instead of the field $F$. It leads to concepts of homogeneous algebraic graphs over $K$, forest and tree approximations, and families of graphs of large girth over $K$. It was proven that, for the case of commutative ring $K$ with unity of odd characteristic, the graphs $CD(n, K)$ are connected (see [15]). So the graph $CD(n, q) = CD(n, F_q)$ for odd $q$ is a connected component of $D(n, q)$.

As it follows from the definitions, the image of a restriction of the homomorphism $\eta$ from $D(n, K)$ onto $CD(n, K)$ coincides with $A(n, K)$.

So, the graphs $A(n, K)$ are connected for the case of $K$ with unity of an odd characteristic.

**Theorem 4.** For each commutative integrity ring $K$, the families of graphs $CD(n, K), n = 2, 3, ...$ and $A(n, K), n = 2, 3, ...$ are forest approximations and the families of graphs of large girth.

The content of the paper was present as a talk at the international conference at the end of the year 2021 in Kyiv. The talk was dedicated to the memory of V. V. Sergeichuk (1949-2021), who was a prominent member of the algebraic community of Ukraine.

This research is supported by Fellowship of British Academy for Researchers at Risk.

REFERENCES

В.О. Устименко, https://orcid.org/0000-0002-2138-2357

Інститут телекомунікацій і глобального інформаційного простору НАН України, Київ
Університет Марії Кюри-Склодовської, Люблін, Польща
E-mail: vasulustimenko@yahoo.pl

ПРО НОВІ РЕЗУЛЬТАТИ ЕКСТРЕМАЛЬНОЇ ТЕОРІЇ ГРАФІВ, ТЕОРІЇ АЛГЕБРАІЧНИХ ГРАФІВ ТА ЇХ ЗАСТОСУВАННЯ

Описано нові конструктивні приклади нескінчених сімейств графів малого світу та великого обхвату. Коротко оглянуто застосування цих об’єктів для побудови LDPC кодів та криптографічних алгоритмів. Визначено сімейства однорідних алгебраїчних графів великого обхвату над довільним комутативним кільцем $K$. Для кожного комутативного кільця цілісності $K$, |$K$| > 2, наведено сімейство дводольних однорідних алгебраїчних графів великого обхвату над $K$, утворене графами з многовидами точок і прямих, ізоморфними $K^n$, та цикловим показником $\geq 2n + 2$. З цим сімейством пов’язано проєктивну границю графів, що є нескінченим лісом.

Ключові слова: сімейство графів великого обхвату, графи малого світу, криптографічні алгоритми, LDPC коди.