Journal of Algebra 569 (2021) 195–219

Contents lists available at ScienceDirect

Journal of Algebra

www.elsevier.com/locate/jalgebra

Diameter of Cayley graphs of $SL(n,p)$ with generating sets containing a transvection ☆

Zoltán Halasi [a,b,*]

[a] *Department of Algebra and Number Theory, Eötvös University, Pázmány Péter sétány 1/c, H-1117, Budapest, Hungary*
[b] *Alfréd Rényi Institute of Mathematics, Reáltanoda utca 13-15, H-1053, Budapest, Hungary*

ARTICLE INFO

Article history:
Received 15 March 2020
Available online 24 November 2020
Communicated by Martin Liebeck

Keywords:
Cayley graph
Babai conjecture
Special linear group
Transvection

ABSTRACT

A well-known conjecture of Babai states that if G is a finite simple group and X is a generating set of G, then the diameter of the Cayley graph $\mathrm{Cay}(G,X)$ is bounded above by $(\log|G|)^c$ for some absolute constant c. The goal of this paper is to prove such a bound for the diameter of $\mathrm{Cay}(G,X)$ whenever $G=SL(n,p)$ and X is a generating set of G which contains a transvection. A natural analogue of this result is also proved for $G=SL(n,K)$, where K can be any field.

☆ This work on the project leading to this application has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 741420). The author was also supported by the National Research Development and Innovation Office (NKFIH) Grant No. K115799 and by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

* Correspondence to: Department of Algebra and Number Theory, Eötvös University, Pázmány Péter sétány 1/c, H-1117, Budapest, Hungary.
E-mail addresses: zhalasi@caesar.elte.hu, halasi.zoltan@renyi.hu.

https://doi.org/10.1016/j.jalgebra.2020.10.031

1. Introduction

For a finite group G and a generator set $X \subset G$ the (undirected) Cayley graph $\mathrm{Cay}(G, X)$ is the connected graph with vertex set G and with edge set $\{(g, gh) \mid g \in G, h \in X\}$. The diameter of this graph is the smallest k such that every element of G can be written as a product of at most k elements from $X \cup X^{-1}$. The diameter of a group G, denoted by $\mathrm{diam}(G)$, is the maximum of the diameters of all Cayley graphs $\mathrm{Cay}(G, X)$ where X runs through all the generating sets of G. The following conjecture was formalised by Babai [2, Conjecture 1.7].

Conjecture 1.1. *If G is a non-Abelian finite simple group, then* $\mathrm{diam}(G) < (\log |G|)^c$ *for some absolute constant c.*

Remark 1.2. One can similarly define the Cayley digraph $\overrightarrow{\mathrm{Cay}}\,(G, X)$ with set of directed edges $\{(g, gh) \mid g \in G, h \in X\}$ and the directed diameters $\mathrm{diam}(\overrightarrow{\mathrm{Cay}}\,(G, X))$ and $\overrightarrow{\mathrm{diam}}\,(G)$. (Note that if X is symmetric, i.e. if $X = X^{-1}$, then $\overrightarrow{\mathrm{Cay}}\,(G, X)$ can be identified with $\mathrm{Cay}(G, X)$ in a natural way.) Clearly, $\mathrm{diam}(\mathrm{Cay}(G, X)) \leq \mathrm{diam}(\overrightarrow{\mathrm{Cay}}\,(G, X))$. On the other hand, a result of Babai [1, Theorem 1.4] states that $\mathrm{diam}(\overrightarrow{\mathrm{Cay}}\,(G, X)) \leq O(\mathrm{diam}(\mathrm{Cay}(G, X))^2(\log |G|)^3)$ also holds. As a consequence, a positive answer to Conjecture 1.1 implies the same result for $\overrightarrow{\mathrm{diam}}\,(G)$ (with a possibly larger c). Therefore, in the following we only consider undirected Cayley graphs $\mathrm{Cay}(G, X)$ with symmetric generating set X.

Babai's conjecture was proved by Helfgott [8] for the case $G = SL(2, p)$. Later, this conjecture was verified for finite simple groups of Lie type of bounded rank independently by Pyber and Szabó [15] and Breuillard, Green and Tao [5] In view of the classification theorem, it remains to prove Babai's conjecture for alternating groups and for classical groups of unbounded rank. Despite much serious effort, Babai's conjecture is still unsolved for both of these classes. So far, the best known general upper bounds for $\mathrm{diam}(G)$ are the following.

On the one hand, let G be an alternating group of degree n. Then a quasipolynomial bound $\mathrm{diam}(G) \leq \exp(O(\log n)^4 \log\log(n))$ has been proved by Helfgott and Seress [10]. Later, their argument has been greatly simplified in [9]. On the other hand, let G be a classical group of rank n over the q-element field. Then $\mathrm{diam}\, G \leq q^{O(n(\log n)^2)}$ by the main result of [7].

In case of $G = A_n$, an upper bound $\mathrm{diam}(\mathrm{Cay}(G, X)) = O(n^C)$ has been proved by Babai, Beals and Seress [3] under the restriction that X contains an element of degree $< n/(3 + \varepsilon)$. (The degree of a permutation is the number of elements moved by it.) Under this assumption, the authors managed to show that there is a 3-cycle which can be written as a product of $O(n^c)$ many elements from X. Then an upper bound $\mathrm{diam}(\mathrm{Cay}(G, X)) = O(n^C)$ follows trivially from this, since there are only $O(n^3)$ many 3-cycles in A_n and every element of $A(n)$ can be written as a product of at most $O(n)$

3-cycles. This result has been improved by Bamberg et al. [4] who proved the same result under the weaker assumption that X contains a permutation fixing at least 37% of the elements.

This, and similar results motivated Pyber to suggest a split of Babai's conjecture into three subproblems in the classical case. If G is a classical group with natural KG-module V, and $g \in G$, then the *support* of g can be defined as the codimension of the eigenspace of g corresponding to the eigenvalue 1 of g. (Intuitively, small support of g means that g is close to being the identity map.) Given a finite classical group G of rank n over the q-element field and a generator set X for G, a proof for Babai's conjecture might be found by solving each of the following subproblems:

- Find an element $1 \neq g \in G$ whose length over X is polynomial in $n(\log q)$ and whose support is at most cn for some $c < 1$.
- Starting with the assumption of the existence $g \in X$ with support $< cn$, find an element $1 \neq t \in G$ whose length over X is polynomial in $n(\log q)$ and whose support is minimal in G.
- Starting with the assumption of the existence $1 \neq t \in X$ whose support is minimal in G, finish the proof of Babai's conjecture.

The goal of this paper is to manage the third subproblem from this list for the case $G = SL(V)$. To achieve this goal we need to consider transvections.

Let V be an $n > 2$-dimensional vector space over an arbitrary field K and $G := SL(V)$. A *transvection* $t \in SL(V)$ is an element of the form $t = 1 + x$ where $x \in \mathrm{End}(V)$ has the property that $\mathrm{Im}(x)$ is a one-dimensional subspace in $\ker(x)$. Thus, $x^2 = 0$ and $\dim(\ker(x)) = n - 1$. (Throughout this paper, $1 \in SL(V)$ denotes the identity map on V.)

Note that an element $1 \neq t \in SL(V)$ has smallest support in $SL(V) \setminus \{1\}$ if and only if t is a transvection.

Theorem 1.3. *Let V be an n-dimensional vector space over the finite field $\mathbb{F}_p$ where p is a prime and let $X \subset G = SL(V)$ be a generating set of $SL(V)$ which contains a transvection. Then* $\mathrm{diam}(\mathrm{Cay}(G, X)) = O((\log p)^c n^{13})$ *for some absolute constant c.*

Remark 1.4. In this theorem, the constant c is the same as in [8, Main Theorem]. In fact, the constant c can be chosen to be 3323 thanks to an explicit version of Helfgott's theorem given by Kowalski [12, Corollary 1.3]. We note that the factor $(\log p)^c$ can be changed to a bound for the diameter of Cayley graphs of $SL(2, \mathbb{F}_p)$ corresponding to generating sets $\{r, s\}$ where r, s is chosen to be two arbitrary non-commuting transvection in $SL(2, \mathbb{F}_p)$.

For any transvection $t = 1 + x \in SL(V)$, and for every field element $\lambda \in K^\times$, the element $t^\lambda := 1 + \lambda x$ is also a transvection. Now, there is a unique *transvection group* t^K

containing t, which is defined as $t^K = \{t^\lambda \mid \lambda \in K\}$. During the proof of Theorem 1.3, we also prove the following, which holds for any field K.

Theorem 1.5. *Let V be an n-dimensional vector space over an arbitrary field K and let $X \subset G = SL(V)$ be a generating set of $SL(V)$ that contains a whole transvection group. Then* $\mathrm{diam}(\mathrm{Cay}(G, X)) = O(n^{11})$.

In [11], Humphries gave conditions when a set of n many transvections generate $SL(n, p)$. (Note that n is the minimal possible size of such a generating set for $SL(n, p)$.) Although we do not use Humphries result directly, his condition was very useful to find a proof for Theorem 1.5. In fact, our proof for Theorem 1.5 also provides a generalisation and extension of Humphries' theorem. For details, see Section 5.

2. Notation and some basic tools

The purpose of this section is to introduce some terminology and to explain some very basic ideas used in the rest of the paper. Through this section let K be any field and let V be an n-dimensional vector space over K.

If $1 \neq t \in SL(V)$ is any transvection, then it can be parametrised by $(u, \phi) \in V \times V^*$, where $V^* = \mathrm{Hom}(V, K)$ is the dual space of V and $\phi(u) = 0$ such that the transvection $t = t_{u,\phi}$ satisfies $t(v) = v + \phi(v)u$ for every $v \in V$. Note that this parametrisation is just almost unique, namely, $t_{\lambda \cdot u,\phi} = t_{u,\lambda \cdot \phi}$ holds for every $\lambda \in K$. Therefore, the set of transvections can be identified with the elements

$$\{1 + u \otimes \phi \mid u \otimes \phi \in V \otimes_K V^*, 0 \neq u \in V, 0 \neq \phi \in V^* \text{ and } \phi(u) = 0\}.$$

By fixing a basis $e_1, \ldots, e_n$ of V, we use the notation $e_1^*, \ldots, e_n^*$ for the dual basis of V^* satisfying $e_i^*(e_j) = \delta_{ij}$. With help of these bases, elements of V and V^* can be identified with the set of column vectors and row vectors over K (each of length n), respectively. We use the notation $[u]$ and $[\phi]^T$ for the corresponding column and row vectors, respectively.

Under this identification, $u \otimes \phi$ is identified with the usual matrix product $[u] \cdot [\phi]^T$ and $V \otimes V^*$ is identified with $K^{n \times n}$, the algebra of all $n \times n$ matrices over K. Furthermore, $E_{ij} := e_i \otimes e_j^*$ becomes the usual basis of $K^{n \times n}$. Now, $0 = \phi(u) = [\phi]^T[u] = \mathrm{Tr}([u] \cdot [\varphi]^T)$, which means that $1 + u \otimes \phi$ is a transvection for some $u \in V$, $\phi \in V^*$ if and only if $u \otimes \varphi$ corresponds to a matrix M satisfying $\mathrm{rank}(M) = 1$ and $\mathrm{Tr}(M) = 0$. Since any matrix of trace zero can be written as a linear combination of the set $\{M \in K^{n \times n} \mid \mathrm{rank}(M) = 1 \text{ and } \mathrm{Tr}(M) = 0\}$, the subspace of $V \otimes V^*$ generated by the set $\{x = u \otimes \phi \mid 1 + x \text{ is a transvection}\}$ corresponds to the subspace $\{M \in K^{n \times n} \mid \mathrm{Tr}(M) = 0\}$, so it has dimension $n^2 - 1$.

Consider the usual action of $SL(V)$ on V and its dual action on V^*, that is, $g \cdot \varphi(v) := \varphi(g^{-1} \cdot v)$ for every $g \in SL(V)$, $\varphi \in V^*$ and $v \in V$. An easy calculation shows that if $t = 1 + u \otimes \varphi$ is a transvection and $g \in SL(V)$, then

$$gtg^{-1} = 1 + (g \cdot u) \otimes (g \cdot \varphi). \qquad \text{(Eq. 1)}$$

This equation and the following three ones will be frequently used in this paper.

Lemma 2.1. *Let $r = 1 + u \otimes \phi$ and $s = 1 + v \otimes \psi$ be two transvections. Then we have*

(a) $srs^{-1} = 1 + (u + \psi(u) \cdot v) \otimes (\phi - \phi(v) \cdot \psi)$. *In particular, if $\phi(v) = 0$ then $srs^{-1} = 1 + (u + \psi(u) \cdot v) \otimes \phi$.*
(b) *If $\phi(v) = 0$, then $[s, r] = srs^{-1}r^{-1} = 1 + \psi(u)v \otimes \phi$.*
(c) *If $\phi = \psi$, then $rs = 1 + (u + v) \otimes \phi$.*

Proof. All of these equations can be easily checked by direct calculations. For the reader's convenience, we give a proof for (b) based on a general identity.

Let x, y be two elements in an associative ring with unity satisfying $x^2 = y^2 = yx = 0$. Then $1 + x,\ 1 + y$ are invertible, and $(1 + x)^{-1} = 1 - x,\ (1 + y)^{-1} = 1 - y$. Thus, we have

$$\begin{aligned}[1 + x, 1 + y] &= (1 + x)(1 + y)(1 - x)(1 - y) = (1 + x + y + xy)(1 - x - y + xy) \\ &= 1 + 2xy + (x + y + xy)(-x - y + xy) = 1 + xy.\end{aligned}$$

In particular, if $r = 1 + u \otimes \phi$ and $s = 1 + v \otimes \psi$ are two transvections satisfying $\phi(v) = 0$, then the above identity may be applied for $x = v \otimes \psi,\ y = u \otimes \phi$. Thus, we get

$$[s, r] = 1 + (v \otimes \psi)(u \otimes \phi) = 1 + \psi(u) \cdot v \otimes \phi. \quad \square$$

We next introduce the concept of *transvection graphs*, which represents the relationship between any pairs of transvections. This tool will be crucial in our argument.

For the remainder of this paper, let $\mathcal{T} = \{1 + u \otimes \phi \mid u \in V,\ \phi \in V^*,\ \phi(u) = 0\}$ denote the set of all transvections. Occasionally, we allow ourselves to consider 1 as an element of $\mathcal{T}$, and we think of 1 as the trivial or the non-proper transvection.

Let $Y \subset \mathcal{T}$ be a set of transvections. Then the directed graph $\Gamma(Y)$ (called the transvection graph on Y) is defined as follows. Its vertex set is $Y \setminus \{1\}$ and for two (proper) transvections $t_i = 1 + x_i = 1 + u_i \otimes \phi_i,\ t_j = 1 + x_j = 1 + u_j \otimes \phi_j$ there is a directed edge from t_i into t_j if $x_j x_i \neq 0$, i.e. if $\phi_j(u_i) \neq 0$.

In particular, if Y is the set of all transvections, we get the full transvection graph $\Gamma(\mathcal{T})$. Clearly, for every set Y of transvections, $\Gamma(Y) = \Gamma(Y \setminus \{1\})$ is just the subgraph of $\Gamma(\mathcal{T})$ induced by $Y \setminus \{1\}$.

For two transvections r, s, we say that (r, s) is an edge if (r, s) is a directed edge in $\Gamma(\mathcal{T})$. The set of all edges (in $\Gamma(\mathcal{T})$) is denoted by $E(\mathcal{T})$. If (r, s) is an edge, then we say that (r, s) is *one-way directed* (resp. *two-way directed*) if (s, r) is not an edge (resp. if (s, r) is also an edge). Note that an edge of some kind between s and t exists in $\Gamma(\mathcal{T})$

if and only if s and t fail to commute. In particular, the transvection graph is some refinement of the non-commuting graph on $\mathcal{T}$.

For the reader's convenience we give an algebraic description of the various types of edges in $\Gamma(\mathcal{T})$. Given two transvections $s,t \in \mathcal{T} \setminus \{1\}$, there is a strong connection between the type of the edge (s,t) and the isomorphism type of the subgroups $\langle s^K, t^K\rangle$. More exactly, the following holds. (In the following proposition $[s]_B$ denotes the matrix form of s with respect to a basis $B \subset V$ and $(K,+)$ denotes the additive group of K.)

Proposition 2.2. *Let $s = 1 + u \otimes \phi, t = 1 + v \otimes \psi$ be two (proper) transvections.*

(1) *If $(s,t) \notin E(\mathcal{T})$ and $(t,s) \notin E(\mathcal{T})$, then there are four possibilities:*
   (a) $\langle u\rangle = \langle v\rangle$, $\langle \phi\rangle = \langle \psi\rangle \iff s^K = t^K \iff \langle s^K, t^K\rangle \simeq (K,+) \iff [s]_B = 1 + E_{12}$, $[t]_B = 1 + \lambda E_{12}$ *for a suitable basis B of V and for some $\lambda \in K^\times$.*
   (b) $\langle u\rangle = \langle v\rangle$, $\langle \phi\rangle \neq \langle \psi\rangle \iff [s]_B = 1 + E_{12}$, $[t]_B = 1 + E_{13}$ *for a suitable basis B of V. Then $\langle s^K, t^K\rangle \simeq (K,+) \times (K,+)$.*
   (c) $\langle u\rangle \neq \langle v\rangle$, $\langle \phi\rangle = \langle \psi\rangle \iff [s]_B = 1 + E_{21}$, $[t]_B = 1 + E_{31}$ *for a suitable basis B of V. Then $\langle s^K, t^K\rangle \simeq (K,+) \times (K,+)$.*
   (d) $\langle u\rangle \neq \langle v\rangle$, $\langle \phi\rangle \neq \langle \psi\rangle \iff [s]_B = 1 + E_{12}$, $[t]_B = 1 + E_{34}$ *for a suitable basis B of V. Then $\langle s^K, t^K\rangle \simeq (K,+) \times (K,+)$.*
(2) *$(s,t) \in E(\mathcal{T})$ is a one-way directed edge $\iff [s]_B = 1 + E_{23}$, $[t]_B = 1 + E_{12}$ for a suitable basis B of V. In that case $\langle s^K, t^K\rangle$ is isomorphic to the Heisenberg group over K.*
(3) *$(s,t) \in E(\mathcal{T})$ is a two-way directed edge $\iff [s]_B = 1 + E_{12}$, $[t]_B = 1 + \lambda E_{21}$ for a suitable basis B of V and for some $\lambda \in K^\times$. In that case $\langle s^K, t^K\rangle \simeq SL(2,K)$.*

Remark 2.3. If $s \in SL(2,K)$ is any transvection, then

$$\langle s\rangle = \{s^\lambda \mid \lambda \in \mathbb{Z}\} \leq \{s^\lambda \mid \lambda \in K\} = s^K.$$

Therefore, $\langle s\rangle = s^K$ if and only if $K = \mathbb{F}_p$ for some prime p. It follows that if $K = \mathbb{F}_p$ and (s,t) is a two-way directed edge in $\Gamma(\mathcal{T})$, then $\langle s,t\rangle = \langle s^K, t^K\rangle = SL(2,K)$.

On the other hand, if $K = \mathbb{F}_q$ is a finite field, where q is a proper prime power and (s,t) is a two-way directed edge, then $\langle s,t\rangle$ can be strictly smaller than $SL(2,K)$. In fact, by a Lemma of Dickson (see [6, Chapter 2, Theorem 8.4]) we have $\langle s,t\rangle \simeq SL(2,L)$ for some subfield $L \leq K$ unless $\operatorname{char} K = 2$ (when $\langle s,t\rangle$ is a dihedral group) or $K \geq \mathbb{F}_9$ (in that case it can happen that $\langle s,t\rangle$ is a central extension of A_5 by C_2).

Now, Lemma 2.1 easily implies

Lemma 2.4. *Let $r,s,t \in \mathcal{T}$ be three proper transvections.*

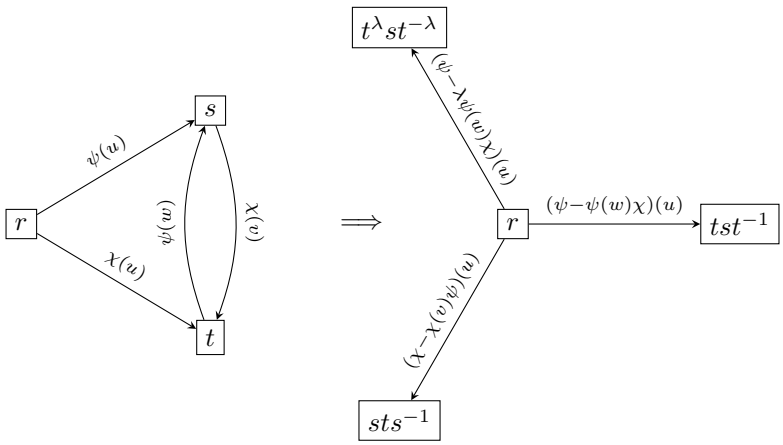
(a) *If* $(r,s),(r,t) \notin E(\mathcal{T})$, *then* $(r,tst^{-1}) \notin E(\mathcal{T})$;
(b) *If* $(r,s) \in E(\mathcal{T})$, *but* $(r,t) \notin E(\mathcal{T})$ *or* $(t,s) \notin E(\mathcal{T})$, *then* $(r,tst^{-1}) \in E(\mathcal{T})$.
(c) *If* (r,s), $(s,t) \in E(\mathcal{T})$ *but* $(r,t) \notin E(\mathcal{T})$, *then* $(r,sts^{-1}) \in E(\mathcal{T})$.
(d) *If* $(r,t),(t,s) \in E(\mathcal{T})$, *then* $(r,t^{\lambda}st^{-\lambda}) \notin E(\mathcal{T})$ *for a suitable* $\lambda \in K$.
(e) *If* (r,s) *is a one-way edge, then* $[s,r]$ *is also a transvection. In that case* $([s,r],t) \in E(\mathcal{T}) \iff (r,t) \in E(\mathcal{T})$ *and* $(t,[s,r]) \in E(\mathcal{T}) \iff (t,s) \in E(\mathcal{T})$.

In (a)-(d), dual statements can be obtained by reversing the direction of the edges. As an example, the dual statement of (a) says that

(a') *If* $(s,r),(t,r) \notin E(\mathcal{T})$, *then* $(tst^{-1},r) \notin E(\mathcal{T})$.

Proof. Instead of a detailed proof we only give a picture to help the reader to better visualize these claims. In fact, instead of simple directed edges we will use labelled edges: For two transvections of the form $r = 1+u\otimes\phi,\ s = 1+v\otimes\psi$ we label the directed edge (r,s) with the field element $l(r,s) = \psi(u)$. Then $(r,s) \notin E(\mathcal{T})$ if and only if $l(r,s) = 0$, so non-edges in $\Gamma(\mathcal{T})$ correspond to edges of this picture with label 0.

Now, let $r = 1+u\otimes\phi,\ s = 1+v\otimes\psi,\ t = 1+w\otimes\chi$ be three (proper) transvections. Then



(Only those edges appear on this picture, which have roles in the above claims.) Now, claims (a)-(d) can be easily checked. For example, claim (d) only says that if $\chi(u) \neq 0$ and $\psi(w) \neq 0$, then one can choose a field element $\lambda \in K$ such that $(\psi - \lambda\psi(w)\chi)(u) = 0$. □

In what follows, we will freely use Lemmas 2.1 and 2.4 without referring to them.

For a set $X \subset SL(V)$, we define $X^k := \{g_1 g_2 \cdots g_k \,|\, g_1, \dots, g_k \in X\}$ for every $k \in \mathbb{N}$. In the following we assume that $1 \in X$. This can clearly be assumed without loss of generality and it implies that $X^k \subset X^l$ whenever $k < l$, which results in some simplification in the notation.

For any two sets $X, Y \subset SL(V)$ we denote by $\ell_X(Y)$ the length of Y over X, i.e. $\ell_X(Y) := \min\{k \in \mathbb{N} \,|\, Y \subset X^k\}$ is the smallest number k such that every element of Y can be written as a product of at most k many elements from X. (If there is no such k, then $\ell_X(Y) := \infty$ or undefined.)

Note that using this notation, the conclusion of Theorem 1.5 can be rewritten as $\ell_X(SL(V)) = O(n^{11})$. Furthermore, ℓ has the property that

$$\ell_X(Z) \leq \ell_X(Y) \cdot \ell_Y(Z) \tag{Eq. 2}$$

for any three sets $X, Y, Z \subset SL(V)$. This property can be used to split the proof of Theorem 1.5 into several steps by providing a chain of sets of transvections with stronger and stronger properties such that each one has a sufficiently small length over the previous one. The main goal of the proof of Theorem 1.5 is to construct all elements of $\mathcal{T}$ as a short product of elements from X. In order to achieve our goal, we need to ensure that the above mentioned sets of transvections have a property, which we call K-closed.

We remind the reader that for any (proper) transvection $t = 1 + x$ the transvection subgroup t^K is the subgroup defined as $t^K = \{t^\lambda := 1 + \lambda x \,|\, \lambda \in K\}$. We say that a subset $Y \subseteq \mathcal{T}$ is *K-closed* if $t^K \subseteq Y$ holds for every $t \in Y$. If Y is any set of transvections, then its K-closure is defined as

$$Y^K = \{s \in \mathcal{T} \,|\, s \in t^K \text{ for some } t \in Y\}.$$

Clearly $Y \subseteq \mathcal{T}$ is K-closed $\iff Y = Y^K$.

During our argument, we always generate new transvections in a way as in Equation (Eq. 1) and in Lemma 2.1. More precisely, we start with $t_0^K \subset X$. In a general step, we have an already constructed K-closed $Y \subset \mathcal{T}$ whose length is known to be short enough in X and we construct a new element $t \in \mathcal{T}$ by one of the following ways:

- $t = grg^{-1}$ where $r \in Y$ and $g \in X^k \cup Y$ (for some short enough k); then $t^K = gr^K g^{-1}$.
- $t = [s, r]$ where $r, s \in Y$ and (r, s) is a one-way directed edge; if $r = 1 + y$, $s = 1 + x$, then by using the identity in the proof of Lemma 2.1/(b), for any $\lambda \in K$ we obtain

$$[s^\lambda, r] = [1 + \lambda x, 1 + y] = 1 + \lambda xy = (1 + xy)^\lambda = t^\lambda,$$

so $t^K = [s^K, r]$.

- $t = rs$ where $r, s \in Y$ and $\ker(r-1) = \ker(s-1)$; then r and s commute. Moreover, we are in the situation of Lemma 2.1/(c), so $t^\lambda = r^\lambda s^\lambda$ for every $\lambda \in K$.

This implies that the upper bound we give for $\ell_{X\cup Y}(t)$ is also an upper bound for $\ell_{X\cup Y}(t^K)$. As a consequence, along with t we can add the whole t^K to Y. In this way, we can guarantee that the set of already constructed transvections remains K-closed through the whole proof.

3. The reduction of Theorem 1.3 to Theorem 1.5

In this section let V be an n-dimensional vector space over $\mathbb{F}_p$. If X is a symmetric generating set of $SL(V)$ and $t \in X$ is a transvection, then we can form many transvections by repeatedly conjugating with the elements of X. In that way we get sets of transvections

$$\mathcal{C}_k = \mathcal{C}(t, X, k) := \{g_k \cdots g_1 t g_1^{-1} \cdots g_k^{-1} \mid g_1, \dots g_k \in X\} \subset X^{2k+1}$$

and corresponding directed graphs $\Gamma(\mathcal{C}_k)$ for every k.

Theorem 3.1. *Let X be a symmetric generating set of $SL(V)$, which contains a transvection t and let*

$$\mathcal{C}_n = \mathcal{C}(t, X, n) = \{g_n \cdots g_1 t g_1^{-1} \cdots g_n^{-1} \mid g_1, \dots g_n \in X\} \subset X^{2n+1}.$$

Then $\Gamma(\mathcal{C}_n)$ contains a directed cycle.

Proof. Let $t = 1 + u \otimes \phi$ for some $u \in U$, $\phi \in V^*$. Since $SL(V)$ acts irreducibly on V and X generates $SL(V)$, we get that $\langle X^i(u) \mid i \in \mathbb{N}\rangle = V$. Therefore, if $\langle X^i(u)\rangle = \langle X^{i+1}(u)\rangle$ for some i, then $\langle X^i(u)\rangle = V$ must hold. Thus, $\langle X(u)\rangle < \langle X^2(u)\rangle < \dots$ is a chain of subspaces, which is strictly increasing until it reaches V. Since $\dim V = n$, it follows that $X^n(u)$ generates V. Let $m = |\mathcal{C}_n|$ and let $t_1 = t, t_2, \dots, t_m$ be the elements of $\mathcal{C}_n$ with $t_i = 1 + u_i \otimes \phi_i$ for every $1 \le i \le m$. Using Equality (Eq. 1), we get that $\langle u_1, \dots, u_m\rangle = \langle X^n(u)\rangle = V$. In a similar way, one can also show that $\langle \phi_1, \dots, \phi_m\rangle = \langle X^n(\phi)\rangle = V^*$.

It follows from the definition of $\Gamma(\mathcal{C}_n)$ that there is no sink vertex in $\Gamma(\mathcal{C}_n)$. Recall that a directed graph is acyclic, if it does not have a directed cycle. As any finite acyclic graph has a sink vertex, we get that $\Gamma(\mathcal{C}_n)$ has a directed cycle, which proves our claim. □

Now, we can reduce Theorem 1.3 to Theorem 1.5.

Theorem 3.2. *Let V be an n-dimensional vector space over $\mathbb{F}_p$ where p is a prime and let $X \subset SL(V)$ be a symmetric generating set containing a transvection t. Then X^l contains a full transvection group for $l = O((\log p)^c \cdot n^2)$.*

Proof. By Theorem 3.1, $\Gamma(\mathcal{C}_n)$ contains a directed cycle. Choosing a directed cycle of minimal length we get $r_1, \dots, r_k \in \mathcal{C}_n$ such that (r_i, r_j) is a directed edge in $\Gamma_{\mathcal{C}_n}$ if and only if $j - i \equiv 1 \pmod{k}$. (In the following, when we consider a directed cycle of length k we think of the indices as elements of $\mathbb{Z}_k$.) Let $r_i = 1 + v_i \otimes \phi_i$ for every $1 \leq i \leq k$, so $\phi_j(v_i) \neq 0 \iff j - i \equiv 1 \pmod{k}$. Let $\alpha_1, \alpha_2 \dots, \alpha_k \in \mathbb{F}_p$. Using the equality $\phi_{i+1}\left(\sum_{s=1}^{k} \alpha_s v_s\right) = \alpha_i \cdot \phi_{i+1}(v_i)$ for every $1 \leq i \leq k$, we get that $v_1, \dots, v_k \in V$ are linearly dependent, so $k \leq n$.

Let $s_i := r_i r_{i+1} \dots r_{k-1} r_k r_{k-1}^{-1} \dots r_{i+1}^{-1} r_i^{-1}$ for every $2 \leq i \leq k-1$. Using Lemma 2.4/(a), (b), (c) or their duals repeatedly for $i = k-1, \dots, 2$, we get that (r_{i-1}, s_i) and (s_i, r_1) are one way edges for $i > 2$ and (r_1, s_2) is a two-way edge. Clearly, $\ell_{\mathcal{C}_n}(s_2) \leq 2k - 3 < 2n$, so $\ell_X(r_1, s_2) \leq 4n^2$. Now, $\langle r_1, s_2 \rangle \simeq SL(2, \mathbb{F}_p)$ by Remark 2.3. (Note that this is the only point where we use that $K = \mathbb{F}_p$. If $|K|$ is a proper prime power, then it can happen that $\langle r_1, s_2 \rangle$ does not contain a full transvection subgroup t^K at all, so our argument does not work in this case.) By [8, Main Theorem], the diameter of $SL(2, \mathbb{F}_p)$ is $O((\log p)^c)$, so $X^{O((\log p)^c n^2)}$ contains a full transvection group. □

Corollary 3.3. *Theorem 1.5 implies Theorem 1.3.*

Proof. Let V be an n-dimensional vector space over $\mathbb{F}_p$ and let $X \subset SL(V)$ be a symmetric generating set containing a transvection. By Theorem 3.2, $X' := X^k$ contains a transvection group for $k = O((\log p)^c \cdot n^2)$. Thus, Theorem 1.5 says that $\operatorname{diam}(\operatorname{Cay}(G, X')) = O(n^{11})$. Therefore,

$$\operatorname{diam}(\operatorname{Cay}(G, X)) \leq k \cdot \operatorname{diam}(\operatorname{Cay}(G, X')) = O((\log p)^c \cdot n^{13}). \quad \square$$

4. Proof of Theorem 1.5

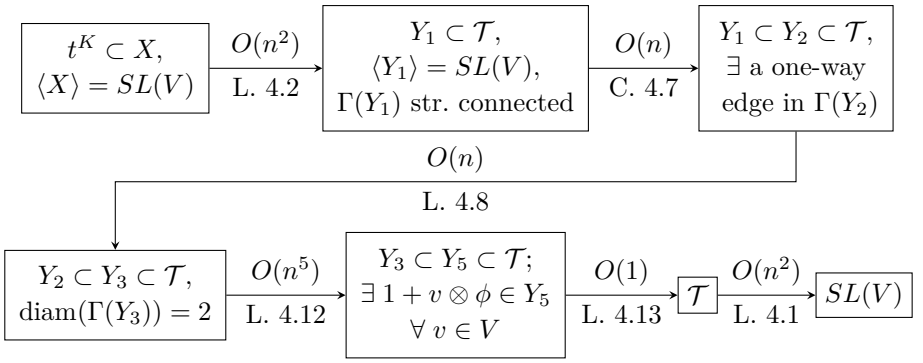
For the remainder, we assume that V is an n-dimensional vector space over an arbitrary field K and X is a symmetric generating set of $SL(V)$ which contains a whole transvection group t^K.

First, we claim that it is enough to prove that the length of $\mathcal{T}$ over X is bounded by a polynomial in n. In fact, we prove this for a relatively small subset of transvections instead of $\mathcal{T}$.

Lemma 4.1. *Let $e_1, \dots, e_n$ be a basis of V and let $Y = \{1 + K \cdot e_i \otimes e_j^* \mid 1 \leq i, j \leq n,\ i \neq j\}$. Then we have* $\operatorname{diam}(\operatorname{Cay}(G, Y)) = O(n^2)$.

Proof. Identify $G = SL(V)$ with $SL(n, K)$ and Y with the set $\{1 + K \cdot E_{ij} \mid i \neq j\}$ as in Section 2. As is well-known, every row operation on a matrix can be given by multiplying the matrix from the left with a product of constantly many elements from Y. The claim follows from the Gaussian elimination process. □

Starting from $t^K \subset \mathcal{T}$, we will construct new transvections by a repeated use of one of the three ways described at the end of Section 2. In that way we construct a chain of K-closed sets of transvections $Y_1 \subset Y_2 \subset \dots Y_5 \subset \mathcal{T}$ until we reach $\mathcal{T}$ such that the length of each of these subsets over the preceding one is bounded by a polynomial of n. The following diagram collects the most important properties of these subsets along with their length over the preceding one and where their existence is proved.



We remind the reader that a directed graph with vertex set S is called *strongly connected* if for every two distinct vertices $a, b \in S$ there is a directed path from a to b. One can easily see that this property is equivalent with the property that for every proper subset $\emptyset \neq S' \subsetneq S$ of S there is an outgoing edge from S', that is, an edge (a, b) with $a \in S',\ b \notin S'$.

Lemma 4.2. *Let X be a symmetric generating set of $SL(V) \simeq SL(n, K)$ which contains a transvection subgroup t^K for some transvection $t \in SL(V)$. Let Y_1 be the K-closed set of transvections defined as*

$$Y_1 = \mathcal{C}(t^K, X, n^2) = \{g_{n^2} \cdots g_1 t^\lambda g_1^{-1} \cdots g_{n^2}^{-1} \mid \lambda \in K,\ g_1, \dots, g_{n^2} \in X\} \subset X^{2n^2+1}.$$

Then

(1) *There are sets of transvections $\{s_1, s_2, \dots, s_n\} \subset Y_1$ and $\{t_1, \dots, t_n\} \subset Y_1$ with $s_i = 1 + u_i \otimes \phi_i$ and $t_i = 1 + v_i \otimes \psi_i$ for each i such that $u_1, \dots, u_n$ is a basis of V and $\psi_1, \dots, \psi_n$ is a basis of V^*.*
(2) *For every transvection $r \in \mathcal{T}$, there are $1 \leq i, j \leq n$ such that (s_i, r) and (r, t_j) are edges in $\Gamma(Y_1 \cup r)$.*
(3) *Y_1 generates $SL(V)$.*
(4) *$\Gamma(Y)$ is strongly connected for any set of transvections Y containing Y_1.*

Proof. It follows from the first paragraph of the proof of Theorem 3.1 that

$$\mathcal{C}(t,X,n)=\{g_n\cdots g_1tg_1^{-1}\cdots g_n^{-1}\mid g_1,\dots,g_n\in X\}$$

contains a set of transvections $s_1=1+u_1\otimes\phi_1,\dots,s_n=1+u_n\otimes\phi_n$ such that $u_1,\dots,u_n$ is a basis of V. The existence of a set of transvections $t_1=1+v_1\otimes\psi_1,\dots,t_n=1+v_n\otimes\psi_n\in\mathcal{C}(t,X,n)$ such that $\psi_1,\dots,\psi_n$ is a basis of V^* can be proved in a similar way. Since $\mathcal{C}(t,X,n)\subset\mathcal{C}(t^K,X,n^2)=Y_1$, we get (1). Now, (2) easily follows from (1).

In order to prove that Y_1 generates $SL(V)$, we consider the case $K\neq\mathbb{F}_2$ first. Let $W=\langle v\otimes\phi\in V\otimes V^*\mid\phi(v)=0\rangle$, so $\dim(W)=n^2-1$. For every i, we consider the subspace $W_i=\langle\nu\in V\otimes V^*\mid 1+\nu\in\mathcal{C}(t^K,X,i)\rangle\leq W$. Since X generates $SL(V)$ and the transvection subgroups are all conjugate to each other, we get that $\cup_{i=1}^{\infty}W_i=W$. As in the proof of Theorem 3.1, we get that $W_1<W_2<\dots$ is a strictly increasing chain of subspaces until it reaches W. This means that $W_{n^2-1}=W$. Identifying $V\otimes V^*$ with $\mathrm{End}(V)$, one can see that there is no proper $W_{n^2-1}=W$-invariant subspace of V, which implies the same for the subgroup $H=\langle\mathcal{C}(t^K,X,n^2-1)\rangle\leq SL(V)$. In other words, H is an irreducible subgroup of $SL(V)$, which is generated by transvection groups. As a special case of the main result of [13], it follows that $H\geq Sp(V)$. Since X generates $SL(V)$ and $Sp(V)$ is not normal in $SL(V)$ (unless $Sp(V)=SL(V)$), we have $\langle xSp(V)x^{-1}\mid x\in X\rangle=SL(V)$. Thus, $\langle Y_1\rangle\geq\langle xHx^{-1}\mid x\in X\rangle\geq SL(V)$ as claimed.

The case when $K=\mathbb{F}_2$ is similar, but we take the strictly increasing chain of subgroups $H_0=t^K<H_1<H_2<\dots$ where $H_i=\langle\mathcal{C}(t^K,X,i)\rangle$ for each i. Now, the length of this chain can be bounded above by $\log_2(|SL(V)|)\leq n^2$.

Now, let us assume that $\Gamma(Y)$ is not strongly connected for some set of transvections $Y\supset Y_1$. Then there is a $\emptyset\neq Z\subsetneq Y$ such that there is no outgoing edge from Z in $\Gamma(Y)$. This means that the subspace $V_Z=\langle v\mid\exists\phi\in V^*:\ 1+v\otimes\phi\in Z\rangle$ is a proper subspace of V fixed by each element of Y. This contradicts with the irreducibility of $SL(V)=\langle Y\rangle$ on V. □

Remark 4.3. For $K=\mathbb{F}_2$, the statement analogous to the result of [13] does not hold. For this case, there are several other types of irreducible subgroups of $SL(V)$ which are generated by transvection groups. For a complete list, see [14].

Our next goal is to produce a K-closed set of transvections $Y_2\supset Y_1$ such that Y_2 contains a one-way directed edge and $\ell_{Y_1}(Y_2)\leq O(n)$. Of course, if Y_1 itself contains a one-way directed edge, then we can choose $Y_2=Y_1$. Therefore, we assume that every edge is two-way directed in Y_1.

To achieve our goal, we consider cycles in $\Gamma(Y_1)$. Let $(r_1,r_2,\dots,r_k)$ be a (two-way directed) cycle in $\Gamma(Y_1)$ with $k\geq 3$ and $r_i=1+v_i\otimes\phi_i$ for each i.

We say that this cycle is *non-singular* if

$$\det_c(v_1, \phi_1, \dots, v_k, \phi_k) := \prod_{i=1}^{k} \phi_i(v_{i+1}) + (-1)^{k-1} \cdot \prod_{i=1}^{k} \phi_{i+1}(v_i) \neq 0. \tag{Eq. 3}$$

Remark 4.4.

(1) Note that the non-singularity of a cycle $(r_1, r_2, \dots, r_k)$ only depends on the transvection groups $r_1^K, \dots, r_k^K$. Indeed, by changing $(1 + v_i \otimes \phi_i)$ to $(1 + v_i \otimes \phi)^\lambda = 1 + (\lambda v_i) \otimes \phi_i = 1 + v_i \otimes (\lambda \phi_i)$ for some $\lambda \in K^\times$, the value of $\det_c(v_1, \phi_1, \dots, v_k, \phi_k)$ is multiplied by λ.

(2) The formula in (Eq. 3) can also be defined for any set of transvections $r_1, \dots, r_k$. Clearly, its value is zero unless at least one of $(r_1, \dots, r_k)$ and $(r_k, \dots, r_1)$ is a directed cycle in $\Gamma(\mathcal{T})$, while it is non-zero if exactly one of $(r_1, \dots, r_k)$ and $(r_k, \dots, r_1)$ is a directed cycle in $\Gamma(\mathcal{T})$.

(3) In the particular case when k is odd and $(r_1, \dots, r_k)$ is a chordless (or induced) cycle, $\det_c(v_1, \phi_1, \dots, v_k, \phi_k)$ has a special meaning: It is just the determinant of the $k \times k$-matrix $(\phi_i(v_j))$.

Lemma 4.5. *Let $C = (r_1, \dots, r_k)$ be a chordless non-singular cycle in $\Gamma(Y_1)$ (with $k \geq 3$). Then $\Gamma(C^{2n+1})$ contains a one-way directed edge.*

Proof. Let $r_i = 1 + v_i \otimes \phi_i$ for each $1 \leq i \leq k$. By our assumption, $\phi_{i+1}(v_i) \neq 0$ and $\phi_i(v_{i+1}) \neq 0$ for each $1 \leq i \leq k$, while $\phi_i(v_j) = 0$ if $i - j \not\equiv \pm 1 \pmod{k}$.

First, let us assume that $k = 3$. We calculate the conjugate

$$r_1'(\lambda) := r_2^\lambda r_1 r_2^{-\lambda} = 1 + (v_1 + \lambda \phi_2(v_1) v_2) \otimes (\phi_1 - \lambda \phi_1(v_2) \phi_2).$$

Now, $(r_1'(\lambda), r_3)$ is a one-way directed edge in $\Gamma(C \cup \{r_1'(\lambda)\})$ if and only if both of the following inequality and equality hold:

$$\phi_3(v_1 + \lambda \phi_2(v_1) v_2) = \phi_3(v_1) + \lambda \phi_2(v_1) \phi_3(v_2) \neq 0,$$
$$(\phi_1 - \lambda \phi_1(v_2) \phi_2)(v_3) = \phi_1(v_3) - \lambda \phi_1(v_2) \phi_2(v_3) = 0.$$

Since each $\phi_i(v_j) \neq 0$, there is such a $\lambda \in K^\times$ if and only if

$$\begin{vmatrix} \phi_3(v_1) & \phi_2(v_1)\phi_3(v_2) \\ \phi_1(v_3) & -\phi_1(v_2)\phi_2(v_3) \end{vmatrix} = -\det_c(v_1, \phi_1, v_2, \phi_2, v_3, \phi_3) \neq 0,$$

which exactly means that the cycle (r_1, r_2, r_3) is non-singular. Thus, there is a one-way directed edge in $\Gamma(C^3)$.

Now, we turn to the case $k \geq 4$, and we use induction on k. Using the same argument as in the proof of Theorem 3.2, we get that $v_1, \dots, v_{k-2} \in V$ is linearly independent, so $k \leq n+2$ holds. Let r'_{k-1} be the conjugate of r_k by r_{k-1}, so

$$r'_{k-1} = r_{k-1} \cdot r_k \cdot r_{k-1}^{-1} = 1 + (v_k + \phi_{k-1}(v_k)v_{k-1}) \otimes (\phi_k - \phi_k(v_{k-1})\phi_{k-1}),$$

that is, $r'_{k-1} = 1 + (v'_{k-1}) \otimes (\phi'_{k-1})$ with $v'_{k-1} = v_k + \phi_{k-1}(v_k)v_{k-1}$ and $\phi'_{k-1} = \phi_k - \phi_k(v_{k-1})\phi_{k-1}$ Thus, for every $1 \leq i \leq k-2$, we have (r'_{k-1}, r_i) is an edge in $\Gamma(C \cup \{r'_{k-1}\})$ if and only if

$$\phi_i(v'_{k-1}) = \phi_i(v_k) + \phi_{k-1}(v_k)\phi_i(v_{k-1}) \neq 0 \iff i \in \{1, k-2\}.$$

Similarly, (r_i, r'_{k-1}) is an edge if and only if $i \in \{1, k-2\}$. Thus, $(r_1, \dots, r_{k-2}, r'_{k-1})$ is a chordless cycle of length $k-1$. Furthermore,

$$\begin{aligned} \phi_1(v'_{k-1}) &= \phi_1(v_k), & \phi_{k-2}(v'_{k-1}) &= \phi_{k-1}(v_k)\phi_{k-2}(v_{k-1}), \\ \phi'_{k-1}(v_1) &= \phi_k(v_1), & \phi'_{k-1}(v_{k-2}) &= -\phi_k(v_{k-1})\phi_{k-1}(v_{k-2}), \end{aligned}$$

which implies that

$$\det_c(v_1, \phi_1, \dots, v_{k-2}, \phi_{k-2}, v'_{k-1}, \phi'_{k-1}) = \det_c(v_1, \phi_1, \dots, v_{k-1}, \phi_{k-1}, v_k, \phi_k) \neq 0.$$

Using this process repeatedly, we find shorter and shorter non-singular chordless cycles $(r_1, \dots, r_{i-1}, r'_i)$ where $r'_i = r_i r'_{i+1} r_i^{-1}$ for $i = k-1, k-2, \dots, 3$. In that way we can find a non-singular two-way directed cycle (r_1, r_2, r'_3), where

$$r'_3 = r_3 r_4 \cdots r_{k-1} r_k r_{k-1}^{-1} \cdots r_3^{-1} \in C^{2n-1}.$$

By the first part of the proof, $\Gamma(C^{2n+1})$ contains a one-way directed edge. □

Lemma 4.6. $\Gamma(Y_1)$ *contains a non-singular chordless cycle.*

Proof. First, we reformulate the concept of non-singularity given by (Eq. 3) by introducing the *potential* for two-way directed cycles. First, for any two transvections, $1 + c \otimes \gamma$, $1 + d \otimes \delta$ connected by a two-way directed edge let

$$r(c, \gamma, d, \delta) := \frac{\delta(c)}{\gamma(d)} \qquad \text{(Eq. 4)}$$

Now, let $(r_1, \dots, r_k)$ be a two-way directed cycle where $r_i = 1 + v_i \otimes \phi_i$ for each i. Then its potential is defined as

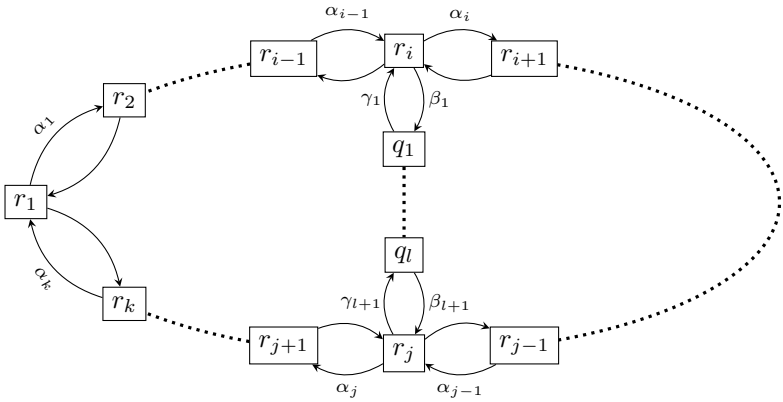
$$\mathrm{Pot}(r_1, \dots, r_k) := \prod_{i=1}^{k} r(v_i, \phi_i, r_{i+1}, \phi_{i+1}).$$

Note that unlike to the above definition of $\det_c$ and r, the potential depends only on the transvection groups $r_1^K, \dots, r_k^K$ and not on the particular choice of the v_i and the ϕ_i. Clearly, the two-way directed cycle $(r_1, \dots, r_k)$ is singular if and only if $\mathrm{Pot}(r_1, \dots, r_k) = (-1)^k$.

The above function $r(c, \gamma, d, \delta)$ has the property

$$r(d, \delta, c, \gamma) = \frac{1}{r(c, \gamma, d, \delta)} \tag{Eq. 5}$$

which can be used to calculate the potential of the symmetric difference of two two-way directed cycles glued by a subpath. More concretely, let $(r_1, r_2, \dots, r_k)$ be a two-way directed cycle and let us assume that for some $1 \leq i < j \leq k$ there is a two-way directed path $r_i, q_1, \dots, q_l, r_j$.



In this picture, the field elements $\alpha_1, \alpha_2, \dots, \alpha_k, \beta_1, \dots, \beta_{l+1}, \gamma_1, \dots, \gamma_{l+1}$ are the fractions defined in (Eq. 4). Thus, by the definition of the potential of directed cycles we have

$$\begin{aligned}
\mathrm{Pot}(r_1, \dots, r_k) &= \alpha_1 \cdots \alpha_k, \\
\mathrm{Pot}(r_1, \dots, r_i, q_1, \dots, q_l, r_j, \dots, r_k) &= \alpha_1 \cdots \alpha_{i-1} \beta_1 \cdots \beta_{l+1} \alpha_j \cdots \alpha_k, \\
\mathrm{Pot}(r_i, \dots, r_j, q_l, \dots, q_1) &= \alpha_i \cdots \alpha_{j-1} \gamma_{l+1} \cdots \gamma_1.
\end{aligned}$$

Using that $\gamma_i = \beta_i^{-1}$ for every i by (Eq. 5) it follows that

$$\begin{aligned}
\mathrm{Pot}(r_1, \dots, r_k) = \mathrm{Pot}(r_1, \dots, r_i, q_1, \dots, q_l, r_j, \dots, r_k) \\
\cdot \mathrm{Pot}(r_i, \dots, r_j, q_l, \dots, q_1)
\end{aligned} \tag{Eq. 6}$$

In particular, a cycle obtained by gluing two singular cycles along their joint subpathes is singular itself.

Let $Y \supset Y_1$ be any set of transvections and let us assume by a way of contradiction that every cycle in $\Gamma(Y)$ is singular. (Note that this implies that Y does not contain any one-way directed edge by Remark 4.4/(2) and by the strongly connected property of Y.) Let $s_1 = 1 + w_1 \otimes \psi_1,\ s_2 = 1 + w_2 \otimes \psi_2 \in Y$ be two neighbouring vertices in $\Gamma(Y)$ and let us assume that $s_3 = s_2 s_1 s_2^{-1}$ is a transvection not in Y. Then $s_3 = 1 + w_3 \otimes \psi_3$ where $w_3 = w_1 + \psi_2(w_1)w_2$ and $\psi_3 = \psi_1 - \psi_1(w_2)\psi_2$. One can check that (s_1, s_3) and (s_2, s_3) are (two-way directed) edges. Let $t = 1 + u \otimes \mu \in Y$ be any transvection in Y. A small calculation shows that

$$\begin{aligned}
\det_c(w_3, \psi_3, w_1, \psi_1, u, \mu) &= \psi_3(w_1)\psi_1(u)\mu(w_3) + \psi_1(w_3)\mu(w_1)\psi_3(u) \\
&= -\psi_1(w_2)\psi_2(w_1)\psi_1(u)(\mu(w_1) + \psi_2(w_1)\mu(w_2)) \\
&\quad + \psi_2(w_1)\psi_1(w_2)\mu(w_1)(\psi_1(u) - \psi_1(w_2)\psi_2(u)) \\
&= -\psi_1(w_2)\psi_2(w_1)\Big(\psi_1(u)\psi_2(w_1)\mu(w_2) + \mu(w_1)\psi_1(w_2)\psi_2(u)\Big) \\
&= -\psi_1(w_2)\psi_2(w_1) \cdot \det_c(w_1, \psi_1, w_2, \psi_2, u, \mu) = 0.
\end{aligned}$$

Thus, if (s_3, s_1, t) is a cycle, then it is singular. A similar calculation shows that if (s_3, s_2, t) is a cycle, then it is singular, as well.

Starting from our assumption that every cycle in $\Gamma(Y)$ is singular we would like to conclude that this also holds for every cycle in $\Gamma(Y \cup \{s_3\})$. Of course, this trivially follows for cycles whose vertices are already in Y, so we should prove this only for cycles in $\Gamma(Y \cup \{s_3\})$ which contain s_3. Let us assume that $r_1, \dots, r_k \in Y$ such that $(r_1, \dots, r_k, s_3)$ is a cycle in $\Gamma(Y \cup \{s_3\})$. Since (s_3, r_1) and (r_k, s_3) are edges, both r_1 and r_k are connected with at least one of s_1 and s_2. If, for example, both of them are connected with s_1, then (s_1, r_1, s_3) and (s_1, s_3, r_k) both are singular cycles, so $\mathrm{Pot}(s_3, r_k, s_1) = \mathrm{Pot}(s_3, s_1, r_1) = -1$. By using (Eq. 6) twice we get

$$\mathrm{Pot}(r_1, \dots, r_k, s_1) = \mathrm{Pot}(r_1, \dots, r_k, s_3) \cdot \mathrm{Pot}(s_3, r_k, s_1) \cdot \mathrm{Pot}(s_3, s_1, r_1).$$

Furthermore, $(r_1, \dots, r_k, s_1)$ is a cycle in $\Gamma(Y)$, so it is singular by our assumption. Therefore, $\mathrm{Pot}(r_1, \dots, r_k, s_3) = (-1)^{k+1}$, which means that $(r_1, \dots, r_k, s_3)$ is also singular. Similarly, if, say, r_1 is connected with s_1 and r_k is connected with s_2 then by using (Eq. 6) three times we get

$$\begin{aligned}
\mathrm{Pot}(s_1, r_1, \dots, r_k, s_2) &= \mathrm{Pot}(s_1, r_1, s_3) \cdot \mathrm{Pot}(s_1, s_3, s_2) \\
&\quad \cdot \mathrm{Pot}(s_2, s_3, r_k)\, \mathrm{Pot}(r_1, \dots, r_k, s_3),
\end{aligned}$$

which implies that $\mathrm{Pot}(r_1, \dots, r_k, s_3) = (-1)^{k+1}$, that is, $(r_1, \dots, r_k, s_3)$ is singular again.

Now, let us assume that all cycles in $\Gamma(Y_1)$ are singular. The above argument shows that if we repeatedly add new transvections to Y_1 by conjugating previous ones with each other, then we can never get a non-singular cycle. But Y_1 generates $SL(V)$ and all transvections are conjugate in $SL(V)$, which implies that sooner or later we get all the transvections of $SL(V)$ that way. Since $n \geq 3$, the graph $\Gamma(\mathcal{T})$ contains a non-singular cycle, which is a contradiction.

Thus, we proved that $\Gamma(Y_1)$ contains a non-singular cycle. Let $(r_1, \dots, r_k)$ be a non-singular cycle such that k is as small as possible. We claim that $(r_1, \dots, r_k)$ is chordless. Otherwise, there is a (two-way directed) edge (r_i, r_j) in $\Gamma(Y_1)$ for some $1 \leq i < j - 1$. Using (Eq. 6), we get that at least one of the shorter cycles $(r_1, \dots, r_i, r_j, \dots, r_k)$ and $(r_i, r_{i+1}, \dots, r_j)$ must be non-singular, a contradiction. □

Using Lemmas 4.6 and 4.5 we get

Corollary 4.7. *There is a K-closed set of transvections $Y_2 \supset Y_1$ with $\ell_{Y_1}(Y_2) \leq O(n)$ such that $\Gamma(Y_2)$ contains a one-way directed edge.*

Lemma 4.8. *There is a K-closed set of transvections $Y_3 \supset Y_2$ of length $\ell_{Y_2}(Y_3) \leq O(n)$ such that for every $s, t \in \mathcal{T}$ there is a directed path from s into t in $Y_3 \cup \{s, t\}$ of length at most 2.*

Proof. Let $s, t \in \mathcal{T}$ be two transvections such that $(s, t) \notin E(\mathcal{T})$. Since $Y_2 \cup \{s, t\}$ is strongly connected, there is a directed path from s into t in $Y_2 \cup \{s, t\}$. Let $s = r_0, r_1, \dots, r_k, t = r_{k+1}$ be a directed path of shortest length in $\Gamma(Y_2 \cup \{s, t\})$ with $r_i = 1 + v_i \otimes \phi_i$ for every i, so there is no edge from r_i into r_j if $j > i + 1$. As in the proof of Theorem 3.2, we get that $\{v_1, \dots, v_k\} \subset V$ is linearly independent set, so $k \leq n$. Let $r_{s,t} := r_k \dots r_2 r_1 r_2^{-1} \dots r_k^{-1}$. Using Lemma 2.4/(b) and the dual of (c) repeatedly, we get that $s, r_{s,t}, t$ is a path. Clearly, $\ell_{Y_2}(r_{s,t}) \leq 2k - 1 \leq 2n$ also holds. Thus, the set

$$Y_3 := Y_2 \cup \{r_{s,t}^K \mid s, t \in \mathcal{T},\ (s, t) \text{ is not an edge}\}$$

has the required properties. □

Lemma 4.9. *There is a K-closed set of transvections $Y_4 \supset Y_3$ with $\ell_{Y_3}(Y_4) \leq O(1)$ such that for every transvection $r \in \mathcal{T}$, there are one-way edges (r, e_r) and (s_r, r) for some $e_r, s_r \in Y_4$.*

Proof. Let $r \in \mathcal{T}$ be any transvection and let (s, t) be a one-way directed edge in $\Gamma(Y_2)$ (whose existence is guaranteed by Corollary 4.7). Let $r = r_0, \dots, r_k = s$ be a path of shortest length in $\Gamma(\{r\} \cup Y_3)$. By the construction of Y_3, we have $0 \leq k \leq 2$. If $k = 0$, then (r, t) is one-way directed. Now, let $k = 1$. Using parts of Lemma 2.4, we get that

if (t,r) is not an edge, then $(r,[s,t])$ is a one-way directed edge, while if (t,r) is an edge, then $(r,t^{\lambda}st^{-\lambda})$ is a one-way directed edge for some $\lambda \in K$. Finally, if $k=2$, then we apply the case $k=1$ twice: First, we can apply it to r_1,s,t we get a one-way directed edge (r_1,e_{r_1}), then we can apply it to (r,r_1,e_{r_1}) to get a one-way directed edge (r,e_r).

The existence of a suitable transvection s_r can be proved in an analogous way. □

Lemma 4.10. *Let $r_1=1+v_1\otimes\phi_1,\dots,r_k=1+v_k\otimes\phi_k$ be a directed path of transvections where $k\leq 5$. Let us assume that at least one of (r_1,r_2) and (r_2,r_3) is one-way directed for $k=3$, while both of (r_1,r_2) and (r_{k-1},r_k) are one-way directed for $4\leq k\leq 5$. Furthermore, let $Z=\{r_1^K,\dots,r_k^K\}$. Then there is a $\psi\in V^*$ such that $s:=1+(v_1+v_k)\otimes\psi$ is a transvection with $\ell_Z(s^K)\leq c$ for some constant c. Similarly, there is a transvection $t=1+w\otimes(\phi_1+\phi_k)$ such that $\ell_Z(t^K)\leq c$.*

Proof. First, if (r_1,r_k) is an edge, then for any $\lambda\in K$ we have

$$r_k^{\lambda}r_1r_k^{-\lambda}=1+(v_1+\lambda\phi_k(v_1)v_k)\otimes(\phi_1-\lambda\phi_1(v_k)\phi_k).$$

Choosing $\lambda=1/\phi_k(v_1)$ we get that

$$s:=r_k^{1/\phi_k(v_1)}r_1r_k^{-1/\phi_k(v_1)}=1+(v_1+v_k)\otimes\Big(\phi_1-\frac{\phi_1(v_k)}{\phi_k(v_1)}\phi_k\Big)$$

satisfy the claim with $\psi=\phi_1-\frac{\phi_1(v_k)}{\phi_k(v_1)}\phi_k$. In a similar way, if (r_k,r_1) is an edge, then we can choose $s=r_1^{1/\phi_1(v_k)}r_kr_1^{-1/\phi_1(v_k)}$.

So for the remainder, we assume that there is no edge between r_1 and r_k in either direction.

If $k=3$ and, say, (r_1,r_2) is a one-way directed edge, then let $s_1:=[r_2^{1/\phi_2(v_1)},r_1]=1+v_2\otimes\phi_1$. Now, (s_1,r_3) is a one-way directed edge, so $s_2:=[r_3^{1/\phi_3(v_2)},s_1]=1+v_3\otimes\phi_1$. Therefore, $s:=r_1\cdot s_2=1+(v_1+v_3)\otimes\phi_1$. The case when (r_2,r_3) is one-way directed can be handled in a similar, but simpler, way.

Now, let $k=4$ and let both (r_1,r_2) and (r_3,r_4) be one-way directed. Choosing $s_1=[r_4^{1/\phi_4(v_3)},r_3]=1+v_4\otimes\phi_3$, the case $k\leq 3$ can be applied to the path r_1,r_2,s_1.

Finally, let $k=5$ and let us assume that both (r_1,r_2) and (r_4,r_5) are one-way directed. If at least one of $\{(r_1,r_4),(r_2,r_4),(r_2,r_5)\}$ is an edge or if one of $\{(r_1,r_3),(r_3,r_5)\}$ is a one-way directed edge, then the case $k\leq 4$ for some smaller path between r_1 and r_5 in $\Gamma(r_1,\dots,r_5)$, so let us assume that this is not the case.

If (r_4,r_1) is one-way directed, then choosing $s_1=[r_1^{1/\phi_1(v_4)},r_4]=1+v_1\otimes\phi_4$ and $s_2=[r_5^{1/\phi_5(v_4)},r_4]=1+v_5\otimes\phi_4$ we get $s:=s_1\cdot s_2=1+(v_1+v_5)\otimes\phi_4$. So for

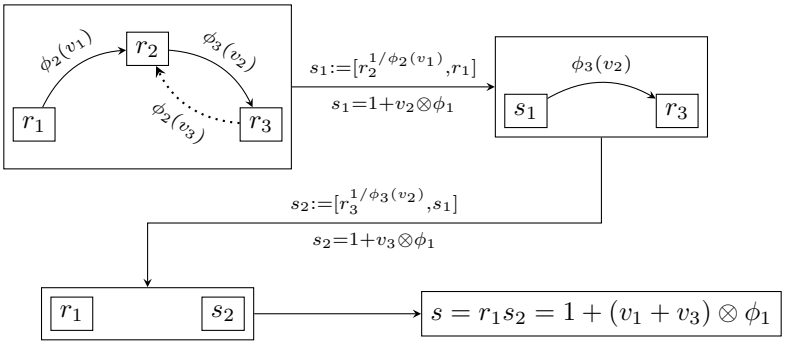
the remainder, we assume that there is no edge between r_1 and r_4 in either direction.

Now, if (r_3, r_1) is not an edge then (r_1, r_3), $(r_2, r_4) \notin E(\mathcal{T})$ by our assumption, so choosing $s_1 = r_3 r_2 r_3^{-1}$, the case $k = 4$ can be applied to the path r_1, s_1, r_4, r_5. The case when $(r_5, r_3) \notin E(\mathcal{T})$ is similar.

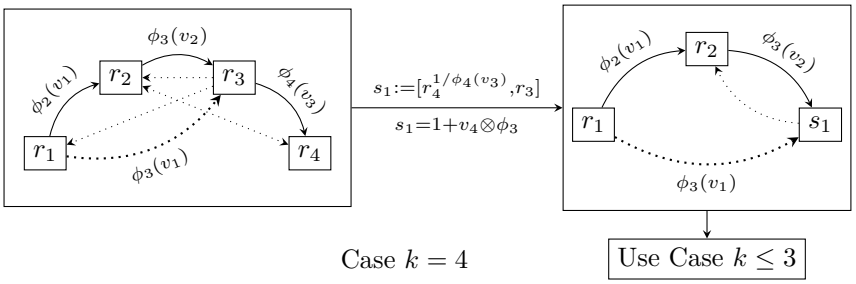
Finally, let us assume that both (r_5, r_3) and (r_3, r_1) are edges. Then there is a $\lambda \in K$ such that $s_1 := r_4^{\lambda} r_3 r_4^{-\lambda}$ satisfies that (r_5, s_1) is one-way directed (if (r_5, r_3) is one-way directed, then $\lambda = 0$ gives $s_1 = r_3$). Now, the case $k = 3$ can be applied to the path r_5, s_1, r_1.

The existence of a $t = 1 + w \otimes (\phi_1 + \phi_k)$ with $\ell_Z(t^K) \leq c$ can be proven in an analogous way.

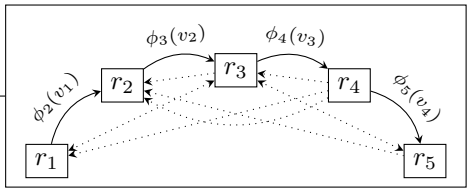
For the reader's convenience we illustrate the possible configurations which we had to handle during the proof. Dotted edge between two transvections means that we do not assume anything about whether the edge exists in $\Gamma(\mathcal{T})$ or not.)



Case $k = 3$

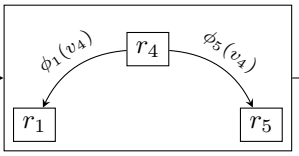


Case $k = 4$



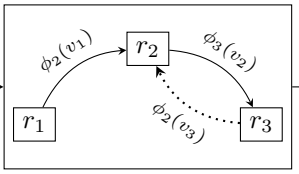
$(r_1, r_4) \in E(\mathcal{T})$ OR $(r_2, r_4) \in E(\mathcal{T})$ OR $(r_2, r_5) \in E(\mathcal{T})$
(r_1, r_3) is one-way OR (r_3, r_5) is one-way
→ Use Case $k \leq 4$

$(r_4,r_1) \in E(\mathcal{T})$
$(r_1,r_4) \notin E(\mathcal{T})$

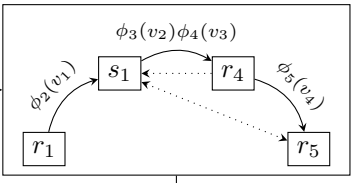


$s_1 := [r_1^{1/\phi_1(v_4)}, r_4] = 1 + v_1 \otimes \phi_4$
$s_2 := [r_5^{1/\phi_5(v_4)}, r_4] = 1 + v_5 \otimes \phi_4$
$s := s_1 s_2 = 1 + (v_1 + v_5) \otimes \phi_4$

$(r_3,r_1) \notin E(\mathcal{T})$
$(r_1,r_3) \notin E(\mathcal{T})$

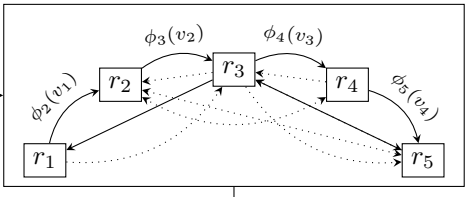


$s_1 := r_3 r_2 r_3^{-1}$

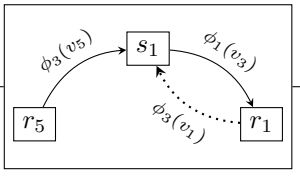


→ Use Case $k \leq 4$

$(r_3,r_1),(r_5,r_3) \in E(\mathcal{T})$
$(r_1,r_4),(r_4,r_1) \notin E(\mathcal{T})$



$\exists\, \lambda$: $(r_5, r_4^{\lambda} r_3 r_4^{-\lambda})$ is one-way
$s_1 = r_4^{\lambda} r_3 r_4^{-\lambda}$



→ Use Case $k = 3$

Case $k = 5$

□

Remark 4.11. One can check that the above argument provides a value $c = 25$.

Lemma 4.12. *There is a K-closed set of transvections $Y_5 \supset Y_4$ with $\ell_{Y_4}(Y_5) \leq O(n^5)$ such that for every $0 \neq v \in V$ there is a $0 \neq \psi \in V^*$ such that $1 + v \otimes \psi \in Y_5$. Similarly, for every $0 \neq \psi \in V^*$ there is a $0 \neq v \in V$ with $1 + v \otimes \psi \in Y_5$.*

Proof. We only prove the first claim, the second one can be proved in a similar way. Let $0 \neq v \in V$ be fixed. By Theorem 4.2, there are transvections $s_1 = 1 + a_1 \otimes \phi_1, \dots, s_k = 1 + a_k \otimes \phi_k \in Y_1$ such that $k \leq n$ and $v = \sum_{i=1}^{k} a_i$. For any fixed positive integer m, let $l(m)$ denote the smallest integer satisfying $m \leq 2^{l(m)}$, so $l(m) = \lceil \log_2(m) \rceil$. We prove the existence of a $\phi \in V^*$ such that $\ell_{Y_4}(1 + v \otimes \phi) \leq 25^{l(k)}$ by using induction on k.

The claim is clearly true for $k = 1$. For an arbitrary $k \leq n$ let $v_1 = \sum_{i=1}^{\lceil k/2 \rceil} a_i$ and $v_2 = \sum_{i=\lceil k/2 \rceil + 1}^{k} a_i$. Using that $\lceil k/2 \rceil \leq 2^{l(k)-1}$ and using induction on k, we get that there are $r_1 = 1 + v_1 \otimes \psi_1$ and $r_2 = 1 + v_2 \otimes \psi_2$ for some $\psi_1, \psi_2 \in V^*$ such that $\ell_{Y_4}(r_1, r_2) \leq 25^{l(k)-1}$. Using Lemma 4.9 and Lemma 4.8 we get a path from r_2 to r_1 in $\Gamma(r_1, r_2, Y_4)$ satisfying the conditions in Lemma 4.10. Thus, using Lemma 4.10 to this path, we get a transvection $r = 1 + v \otimes \psi$ with $\ell_{\{r_1, r_2, Y_4\}}(r) \leq c = 25$. Thus, $\ell_{Y_4}(r) \leq \ell_{\{r_1, r_2, Y_4\}}(r) \cdot \ell_{Y_4}(r_1, r_2) \leq 25 \cdot 25^{l(k)-1} = 25^{l(k)}$, as claimed.

Now, for any $k \leq n$ we have $25^{l(k)} = 25^{\lceil \log_2(k) \rceil} \leq 25 \cdot 2^{\log_2 25 \cdot \log_2(n)} \leq 25n^5$.

Let Y_5 be the union of all transvection groups in $Y_4^{(25n^5)}$. Then the conclusion of the lemma holds for Y_5. □

Lemma 4.13. *Let $\mathcal{T}$ be the set of all transvections. Then $\ell_{Y_5}(\mathcal{T}) = O(1)$.*

Proof. Let $0 \neq v \in V,\ 0 \neq \phi \in V^*$ satisfying $\phi(v) = 0$. We need to prove that $1 + v \otimes \phi \in Y_5^c$ for some constant c. By Lemma 4.12, there are transvections $s_1, s_2 \in Y_5$ such that $s_1 = 1 + v_1 \otimes \phi,\ s_2 = 1 + v \otimes \phi_2$ for some $v_1 \in V,\ \phi_2 \in V^*$. If $\langle v_1 \rangle = \langle v \rangle$ or $\langle \phi_2 \rangle = \langle \phi \rangle$, then the assertion follows since Y_5 is K-closed.

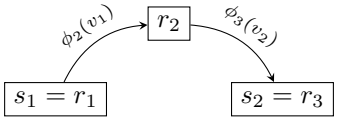
By our assumption, $\phi(v) = 0$, that is, (s_2, s_1) is not an edge. Therefore, if (s_1, s_2) is an edge, then it is a one-way edge, so $[s_2^{1/\phi_2(v_1)}, s_1] = 1 + v \otimes \phi \in Y_5^4$. Thus, for the remainder we assume that $\langle v_1 \rangle \neq \langle v \rangle,\ \langle \phi_2 \rangle \neq \langle \phi \rangle$ and there is no edge between s_1 and s_2 in either direction.

We claim that there is a path $s_1 = r_1, r_2, \dots, r_k = s_2$ with $k \leq 5$ such that $\ell_{Y_5}(r_1, \dots, r_k) \leq 5$, and none of $(r_2, r_1),\ (r_k, r_{k-1}),\ (r_k, r_2),\ (r_{k-1}, r_1)$ are edges. First, let $r_1 = s_1,\ r_k = s_2$ (the value of k will be specified later). Since $\langle v_1 \rangle \neq \langle v \rangle$, there is a $\psi \in V^*$ such that $\psi(v_1) \neq 0,\ \psi(v) = 0$. Now, by Lemma 4.12, $1 + u \otimes \psi \in Y_5$ for some $u \in V$. Furthermore, by Lemma 4.9, there is an $1 + u' \otimes \psi' \in Y_4$ such that $(1 + u \otimes \psi, 1 + u' \otimes \psi')$ is a one-way edge. Then $[1 + (\psi'(u))^{-1} u' \otimes \psi', 1 + u \otimes \psi] = 1 + u' \otimes \psi \in Y_5^4$. Let $U = \langle u, u' \rangle$, then $\dim(U) = 2$ and

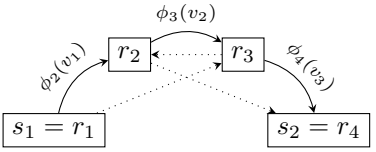
$$\{1 + x \otimes \psi \,|\, x \in U\} = (1 + u \otimes \psi)^K \cdot (1 + u' \otimes \psi)^K \subset Y_5^5.$$

The restriction of ϕ to U must have non-trivial kernel, so there is an $x \in U \setminus \{0\}$ with $\phi(x) = 0$. Now, let $t = 1 + x \otimes \psi$. Then $\ell_{Y_5}(t) \leq 5$, (r_1, t) is a one-way edge and (r_k, t) is not an edge. In an analogous way, we can prove the existence of a transvection t' such that $\ell_{Y_5}(t') \leq 5$, (t', r_k) is a one-way edge and (t', r_1) is not an edge. Now, one of the following holds:

(a) If $t^K = (t')^K$, then let $k = 3$ and let $r_2 = t$.
(b) If (t, t') is an edge, then let $k = 4$ and $r_2 = t,\ r_3 = t'$.
(c) The previous two cases do not hold. Then let $k = 5$, $r_2 = t,\ r_4 = t'$ and $r_3 \in Y_5$ a transvection such that r_2, r_3, r_4 is a path. (The existence of such an r_3 follows from Lemma 4.8.)



$k = 3$



$k = 4$

For each $1 \leq i \leq k$, let $r_i = 1 + v_i \otimes \phi_i$ (where $\phi_1 = \phi$ and $v_k = v$). Now,

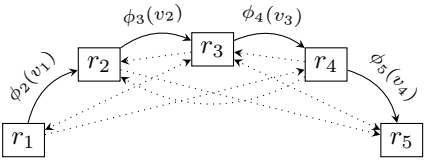
$$[r_3^{1/\phi_3(v_2)}, [r_2^{1/\phi_2(v_1)}, r_1]] = 1 + v_3 \otimes \phi_1 = 1 + v \otimes \phi \text{ for } k = 3$$

and

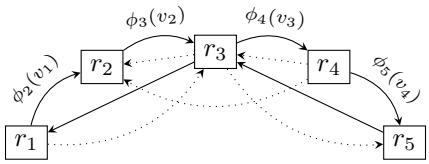
$$[r_4^{1/\phi_4(v_3)}, [r_3^{1/\phi_3(v_2)}, [r_2^{1/\phi_2(v_1)}, r_1]]] = 1 + v_4 \otimes \phi_1 = 1 + v \otimes \phi \text{ for } k = 4.$$

(Note that in these cases $\phi_1(v_i) = 0$ for $2 \leq i \leq k$, so we can use Lemma 2.1(c) repeatedly.)

Finally, let $k = 5$. Then we have the following graph.



(Note that, by assumption, there are no edges from r_2 to r_4, from r_4 to r_1 and from r_5 to r_2 in this graph.) If (r_1, r_4) or (r_2, r_5) is a (one-way) edge, then the case $k = 3$ can be used. Otherwise, if (r_3, r_1) or (r_5, r_3) is not an edge, then we can use $[r_2, r_1]$ or $[r_5, r_4]$ to reduce the problem of generating $1 + v_5 \otimes \phi_1$ to the case $k = 4$. Then an iterated commutator similar as above works. So for the remainder we assume that both (r_5, r_3) and (r_3, r_1) are edges.



Now, we can use Lemma 2.4(d), (b) and their dual to confirm the existence of $\lambda, \mu \in K$ such that $t_1 = r_4^{\mu} r_2^{\lambda} r_3 r_2^{-\lambda} r_4^{-\mu}$ satisfies that r_5, t_1, r_1 is a one-way directed path. Using again Lemma 4.8, there is a $t_2 \in Y_5$ such that r_1, t_2, r_5 is a path. Changing t_2 to $r_1 t_2 r_1^{-1}$ if necessary, we can assume that (t_1, t_2) is an edge. By Lemma 2.4(d), there is a $\nu \in K$ such that $t_3 = t_1^{\nu} t_2 t_1^{-\nu}$ satisfies that r_1, t_3, r_5 is a path and (t_3, r_5) is a one-way edge. Then $[[r_5^K, t_3], r_1] = (1 + v_5 \otimes \phi_1)^K = (1 + v \otimes \phi)^K$ holds. □

Proof of Theorem 1.5. Using Lemmas 4.2, 4.8, 4.9, 4.12, 4.13, 4.1 and Corollary 4.7 we get that

$$\ell_X(SL(V)) \leq \ell_X(Y_1)\ell_{Y_1}(Y_2)\dots\ell_{Y_5}(\mathcal{T})\ell_{\mathcal{T}}(SL(V))$$
$$\leq O(n^2)\cdot O(n)\cdot O(n)\cdot O(1)\cdot O(n^5)\cdot O(1)\cdot O(n^2) = O(n^{11}). \quad \square$$

5. A generalisation of a theorem of Humphries

As a side-effect of the proof of Theorem 1.5, we generalise and extend a theorem of Humphries. In [11], the author gave a sufficient and necessary condition when a set of transvections $S \subset SL(n,p)$ of size n generates $SL(n,p)$. (Note that n is the minimal possible size of a generating set S of transvections in $SL(n,p)$.)

For this section, let V be an $n > 2$-dimensional vector space over an arbitrary field K. Let $S, S' \subset SL(V)$ be two sets of transvections (of the same size). According to Humphrey, we say that there is a *t-equivalence* $S \to S'$ if there is a chain $S = S_0, S_1, \dots, S_t = S'$ of sets of transvections such that for each $1 \leq i \leq t$ there are $x, y \in S_{i-1}$ such that S_i is obtained from S_{i-1} by replacing x to its conjugate yxy^{-1}. Clearly, if there is a t-equivalence $S \to S'$, then S and S' generate the same subgroup of $SL(V)$.

Let $S = \{t_\alpha = 1 + v_\alpha \otimes \phi_\alpha \,|\, \alpha \in I\} \subset SL(n,K)$ be any set of transvections. We consider the following properties:

(P 1.) $\langle v_\alpha \,|\, \alpha \in I\rangle = V$ and $\langle \phi_\alpha \,|\, \alpha \in I\rangle = V^*$.
(P 2.) $\Gamma(S)$ is strongly connected.
(P 3.) There is a non-singular cycle in $\Gamma(S)$.
(P 3′.) There is a t-equivalence $S \to S'$ such that $\Gamma(S')$ contains a one-way directed edge.

Humphries result [11, Theorem 1.1] says that a set of transvections $S \subset SL(n,p)$ of size n generates $SL(n,p)$ if and only if S satisfies (P 1.), (P 2.) and (P 3′.). One expects that the same assertion should be true without the condition $|S| = n$, but Humphries' proof uses this condition in an essential way.

For a set of transvections $S = \{t_\alpha \,|\, \alpha \in I\} \subset SL(V)$, we use the notation S^K for the K-closure of S, i.e. $S^K = \cup_{\alpha\in I}^m t_\alpha^K$. Using the arguments of Section 4, we prove the following.

Theorem 5.1. *Let $S \subset SL(V)$ be a set of transvections. Then S^K generates $SL(V)$ if and only if $\{$(P 1.), (P 2.), (P 3.)$\}$ or $\{$(P 1.), (P 2.), (P 3′.)$\}$ holds for S.*

Note that if $|K|$ is a prime, then $t^K = \langle t\rangle$ for any transvection t, so S^K can be replaced with S in this Theorem. On the other hand, if L is a proper subfield of L, and $S \subset SL(n,L) \leq SL(n,K)$ then clearly $\langle S\rangle \neq SL(n,K)$ regardless of what conditions S satisfies.

First we prove the following

Theorem 5.2. *Let $S \subset SL(V)$ be a set of transvections. Then S generates an irreducible subgroup of $SL(V)$ if and only if $\{$(P 1.), (P 2.)$\}$ holds for S.*

Proof. Let $S = \{t_\alpha = 1 + v_\alpha \otimes \phi_\alpha \,|\, \alpha \in I\}$ and let $H = \langle S^K\rangle \leq SL(V)$. First, let us assume that $H \leq SL(V)$ is irreducible. Then the H-invariant subspaces $\langle v_\alpha \,|\, \alpha \in I\rangle > 0$ and $\cap_{\alpha\in I} \ker(\phi_\alpha) < V^*$ must be trivial, which proves (P 1.). Furthermore, (P 2.) also holds by the last paragraph of the proof of Lemma 4.2.

Now, let us assume that $\{$(P 1.), (P 2.)$\}$ holds for S. Let us assume that $0 \neq U \leq V$ is an H-invariant subspace and let $0 \neq u \in U$. By (P 1.), there is an $\alpha \in I$ such that $u \notin \ker(\phi_\alpha)$. Then $v_\alpha = t_\alpha(u) - u \in U$. Let $\beta \in I$ be any element of the index set. Property (P 2.) implies the existence of a path $t_\alpha = t_{\gamma_0}, t_{\gamma_1}, \ldots, t_{\gamma_k} = t_\beta$ in $\Gamma(S)$. Using induction on k, it follows that $v_\beta = v_{\gamma_k} \in \langle t_{\gamma_k}(v_{\gamma_{k-1}}) - v_{\gamma_{k-1}}\rangle \leq U$. Therefore, $U \geq \langle v_\alpha \,|\, \alpha \in I\rangle = V$ by property (P 1.). □

Proof of Theorem 5.1. Let $S \leq SL(V)$ be a set of transvections. First let us assume that S^K generates $SL(V)$. Then $\langle S\rangle$ acts irreducibly on V, so (P 1.) and (P 2.) follows by Theorem 5.2. It is easy to check that the proof of Lemma 4.6 can be applied to S^K. Thus, $\Gamma(S^K)$ contains a chordless non-singular cycle, which implies property (P 3.). Finally, (P 3′.) is a consequence of (P 3.) by Lemma 4.5.

For the converse direction, let use assume that S has properties $\{$(P 1.), (P 2.), (P 3.)$\}$. Then S also has property $\{$(P 1.), (P 2.), (P 3′.)$\}$ by Lemma 4.5. Finally, one can check that after Corollary 4.7, our argument only uses that Y_2 has properties $\{$(P 1.), (P 2.)$\}$ and $\Gamma(Y_2)$ contains a one-way directed edge, but it never refers to the identity $\langle Y_2\rangle = SL(V)$. So this argument can be applied to S^K instead of Y_2 to prove that $\langle S^K\rangle = SL(V)$. □

As a consequence of Theorems 5.2, 5.1 and the main result of [13] we get

Corollary 5.3. *Let us assume that $K \neq \mathbb{F}_2$ and $S \subset SL(V)$ is a set of transvections. Then S^K generates $Sp(V)$ if and only if S has properties (P 1.) and (P 2.) but it does not have (P 3.).*

References

[1] L. Babai, On the diameter of Eulerian orientations of graphs, in: Proc. 17th Ann. Symp. on Discrete Algorithms (SODA'06), ACM–SIAM, 2006, pp. 822–831.
[2] L. Babai, Á. Seress, On the diameter of permutation groups, Eur. J. Comb. 13 (1992) 231–243.
[3] L. Babai, R. Beals, Á. Seress, On the diameter of the symmetric group: polynomial bounds, in: Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms, ACM, New York, 2004, pp. 1108–1112.
[4] J. Bamberg, N. Gill, T.P. Hayes, H.A. Helfgott, Á. Seress, P. Spiga, Bounds on the diameter of Cayley graphs of the symmetric group, J. Algebraic Comb. 40 (2014) 1–22.
[5] E. Breuillard, B. Green, T. Tao, Approximate subgroups of linear groups, Geom. Funct. Anal. 21 (2011) 774–819.
[6] D. Gorenstein, Finite Groups, second edition, Chelsea Publishing Co., New York, 1980.
[7] Z. Halasi, A. Maróti, L. Pyber, Y. Qiao, An improved diameter bound for finite simple groups of Lie type, Bull. Lond. Math. Soc. 51 (2019) 645–657.
[8] H.A. Helfgott, Growth and generation in $SL_2(\mathbb{Z}/p\mathbb{Z})$, Ann. Math. (2) 167 (2008) 601–623.
[9] H.A. Helfgott, Growth in linear algebraic groups and permutation groups: towards a unified perspective, in: C.M. Campbell, C.W. Parker, M.R. Quick, E.F. Robertson, C.M. Roney-Dougal (Eds.), Groups St Andrews 2017 in Birmingham, Cambridge University Press, Cambridge, 2019, pp. 300–345.
[10] H.A. Helfgott, Á. Seress, On the diameter of permutation groups, Ann. Math. (2) 179 (2014) 611–658.
[11] S.P. Humphries, Generation of special linear groups by transvections, J. Algebra 99 (1986) 480–495.
[12] E. Kowalsky, Explicit growth and expansion for SL_2, Int. Math. Res. Not. 2013 (24) (2013) 5645–5708.
[13] J. McLaughlin, Some groups generated by transvections, Arch. Math. 18 (1967) 364–368.
[14] J. McLaughlin, Some subgroups of $SL_n(\mathbb{F}_2)$, Ill. J. Math. 13 (1969) 108–115.
[15] L. Pyber, E. Szabó, Growth in finite simple groups of Lie type, J. Am. Math. Soc. 29 (2016) 95–146.