ISRAEL JOURNAL OF MATHEMATICS **TBD** (2024), 1–52
DOI: 10.1007/s11856-024-2605-1

ON THE DIAMETER OF CAYLEY GRAPHS OF CLASSICAL GROUPS WITH GENERATING SETS CONTAINING A TRANSVECTION

BY

Martino Garonzi*

Departamento de matemática, Universidade de Brasília
Campus Universitário Darcy Ribeiro, Brasilia - DF, 70910–900, Brazil
e-mail: mgaronzi@gmail.com
URL: https://orcid.org/0000-0003-0041-3131

AND

Zoltán Halasi** and Gábor Somlai**,†

Eötvös Loránd University, Pázmány Péter sétány 1/c, H-1117, Budapest, Hungary
and
Alfréd Rényi Institute of Mathematics
Reáltanoda utca 13-15, H-1053, Budapest, Hungary
e-mail: halasi.zoltan@renyi.hu, gabor.somlai@ttk.elte.hu
URL: https://orcid.org/0000-0002-1305-5380
URL: https://orcid.org/0000-0001-5761-7579

* The first author acknowledges the support of Fundação de Apoio à Pesquisa do Distrito Federal (FAPDF) - demanda espontânea 03/2016, and of Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) - Grant numbers 302134/2018-2, 422202/2018-5.

** The work of the second and third authors 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). Their work was supported by the National Research, Development and Innovation Office (NKFIH) Grant No. K138596.

† The third author was also partially supported by the János Bolyai Research Fellowship and by the New National Excellence Program under the grant number UNKP-20-5-ELTE-231.

Received date March 8, 2022 and in revised form November 6, 2022

ABSTRACT

A well-known conjecture of Babai states that if G is any finite simple group and X is a generating set for G, then the diameter of the Cayley graph $\mathrm{Cay}(G, X)$ is bounded by $\log |G|^c$ for some universal constant c. In this paper, we prove such a bound for $\mathrm{Cay}(G, X)$ for $G = \mathrm{PSL}(n, q)$, $\mathrm{PSp}(n, q)$ or $\mathrm{PSU}(n, q)$ where q is odd, under the assumptions that X contains a transvection and $q \neq 9$ or 81.

1. Introduction

Given a finite group G and a set X of generators of G, the associated (undirected) Cayley graph $\mathrm{Cay}(G, X)$ is defined to have vertex set G and edge set $\{\{g, gx\} : g \in G,\ x \in X\}$. The diameter of $\mathrm{Cay}(G, X)$ equals the maximum over $g \in G$ of the length of a shortest expression of g as a product of generators in X and their inverses. The maximum of $\mathrm{diam}(\mathrm{Cay}(G, X))$, as X runs over all possible generating sets of G, is denoted by $\mathrm{diam}(G)$.

In 1988 Babai [2] proposed the following conjecture.

CONJECTURE 1.1: *If G is a non-abelian finite simple group, then*

$$\mathrm{diam}(G) \leq (\log |G|)^c$$

for some absolute constant c.

Despite many efforts, Babai's conjecture is still open for a general G. In view of the Classification Theorem, the above conjecture should be proved for three classes of simple groups: For alternating groups, for simple groups of Lie type of bounded rank and for classical groups of large rank. (Note that if $|G|$ is bounded, then the above inequality trivially holds for a large enough c, so one does not need to care about sporadic groups.)

In the paper [2], Babai and Seress established the following bound:

$$\mathrm{diam}(A_n) < \exp(\sqrt{n \log n}(1 + o(1))).$$

Moreover, a similar bound holds for arbitrary permutation groups of degree n (see [3]). Up to now, the best available bound for alternating groups was proved by Helfgott and Seress [16], which says that

$$\mathrm{diam}(\mathrm{Cay}(A_n, S)) = \exp(O((\log n)^4 \log \log n))$$

if S is any generating set of A_n.

The first infinite series of finite simple groups for which the conjecture was proved, by Helfgott [14], is $\mathrm{PSL}(2,p)$, where p is a prime. Later, with much more effort, he succeeded to extend his results to $\mathrm{PSL}(3,p)$ (see [15]).

In fact, for simple groups of Lie type of bounded rank, Babai's conjecture is now completely solved. Note that this covers the case of exceptional simple groups of Lie type. It is an easy consequence of the following "product theorem":

THEOREM 1.2 (Pyber–Szabó [23], Breuillard–Green–Tao [6]): *For any positive integer r, there is an $\varepsilon = \varepsilon(r) > 0$ such that if G is any finite simple group of Lie type of rank r and X is a generating set of G, then either $|X^3| > |X|^{1+\varepsilon}$ or $X^3 = G$.*

As a consequence of this deep result one gets a strong form of Babai's conjecture.

COROLLARY 1.3: *If G is a finite simple group of Lie type of bounded rank r, then for any generating set X of G*

$$\operatorname{diam}(\mathrm{Cay}(G,X)) = O\Big(\frac{\log|G|}{\log|X|}\Big)^c,$$

where c depends only on r.

Proof. Let $\varepsilon = \varepsilon(r)$ as in Theorem 1.2 and let k be the smallest integer satisfying $|X|^{(1+\varepsilon)^k} > |G|$. Assuming that $X^{(3^k)} \neq G$, by a repeated use of Theorem 1.2 we get that

$$|X^{(3^k)}| > |X|^{(1+\varepsilon)^k} > |G|,$$

a contradiction. So, $\operatorname{diam}(\mathrm{Cay}(G,X)) \leq 3^k$. On the other hand, by choosing $c = \log_{1+\varepsilon} 3$,

$$|X|^{(1+\varepsilon)^k} > |G| \iff (1+\varepsilon)^k > \frac{\log|G|}{\log|X|} \iff 3^k > \Big(\frac{\log|G|}{\log|X|}\Big)^c,$$

that is, k is the smallest integer satisfying $3^k > \left(\frac{\log|G|}{\log|X|}\right)^c$. Hence

$$\operatorname{diam}(\mathrm{Cay}(G,X)) \leq 3^k \leq 3\Big(\frac{\log|G|}{\log|X|}\Big)^c.$$

This concludes the proof. ∎

In contrast to the above results, for classical groups G of large rank, the best known bounds for $\operatorname{diam}(G)$ are very far from the bound predicted by Conjecture 1.1. The best known upper bounds for $\operatorname{diam}(G)$ obtained in [5], [13]

when G is a classical group of unbounded rank are exponential in q. More precisely, it was proved in [13] that

$$\operatorname{diam}(G) \leq q^{O(n \log n)^2}.$$

On the other hand, for large enough q compared to n, this was strenghtened by Bajpai, Dona and Helfgott [4] who proved that if G is a classical Chevalley group of rank n defined over the field $\mathbb{F}_q$, then

$$\operatorname{diam}(G) \leq (\log |G|)^{1947 n^4 \log(2n)}.$$

If we restrict our attention to generating sets $X \subset G$ with special conditions, then we might prove better bounds for $\operatorname{diam}(G, X)$. In particular, the second author proved a suitable bound for $\operatorname{diam}(\operatorname{Cay}(G, X))$ when $G = \mathrm{SL}(n, p)$, p is a prime and X is a generating set for G containing a transvection (see [12]).

Remark 1.4: In many cases, a simple group G is given as the image of a quasisimple group $\widetilde{G}$ under a surjective homomorphism $\tau : \widetilde{G} \to G$. Now, if X is any generating set for G, then $\widetilde{X} = \tau^{-1}(X)$ is a generating set for $\widetilde{G}$ satisfying $\operatorname{diam}(\operatorname{Cay}(\widetilde{G}, \widetilde{X})) = \operatorname{diam}(\operatorname{Cay}(G, X))$. Since we also have $|\widetilde{G}| \leq |G|^{O(1)}$, a positive answer to Babai's conjecture for $\widetilde{G}$ implies the positive answer to Babai's conjecture for G.

The main result of this paper is the following.

THEOREM 1.5: *Let V be an n-dimensional vector space over the finite field $\mathbb{F}_q$ where q is odd and G is one of* $\mathrm{SL}(V)$, $\mathrm{Sp}(V)$ *or* $\mathrm{SU}(V)$. *Let X be a generating set for G containing a transvection. Then* $\operatorname{diam}(\operatorname{Cay}(G, X)) \leq (n \log q)^c$ *for some constant c provided that*

- $q \neq 9$ *if* $G = \mathrm{Sp}(V)$;
- $q \neq 81$ *if* $G = \mathrm{SU}(V)$;
- $q \neq 9$ *and* $q \neq 81$ *if* $G = \mathrm{SL}(V)$.

So Babai's bound holds for these cases.

A common modification of conjectures of these types is the case of random generators, which was also highlighted by Lubotzky [21]. A special case of one of the main results of a beautiful paper by Eberhard and Jezernik (see [10, Theorem 1.1]) says that if G is a classical group over $\mathbb{F}_q$ of rank n where q is bounded and n is large enough, then by choosing $X = \{x, y, z\}$ randomly from G, there is a word $w \in F_3$ of length $n^{O(1)}$ with high probability (with probability $1 - e^{-cn}$

for some absolute constant c) such that $w(x, y, z)$ is an element of G of minimal degree. Since the elements of minimal degree of $\mathrm{SL}(V), \mathrm{SU}(V)$ and $\mathrm{Sp}(V)$ are exactly the transvections of these groups, a combined use of this result with the main result of this paper implies the following.

COROLLARY 1.6: *Let G be one of $\mathrm{SL}(n, q)$, $\mathrm{Sp}(n, q)$ or $\mathrm{SU}(n, q)$ where q satisfies the assumptions of Theorem 1.5. Let us also assume that q is bounded. Let $X = \{x, y, z\}$, where x, y, z are chosen randomly from G (i.e., independently and with uniform distribution). Then*

$$P(\mathrm{diam}(\mathrm{Cay}(G, X)) \leq (\log |G|)^C) \geq 1 - e^{-cn}$$

for some constants c, C. That is, Babai's conjecture holds for three random generators with high probability when n is large enough.

A part of our proof can be used to show the following, which holds even for infinite fields.

THEOREM 1.7: *Let us assume that V is an n-dimensional non-degenerate symplectic or unitary vector space over the field K, and $G = \mathrm{Sp}(V)$ or $G = \mathrm{SU}(V)$. Let us assume that X is a generating set for G containing a transvection subgroup over $K_0 \leq K$, where $K_0 = K$ in the symplectic case and $|K : K_0| = 2$ in the unitary case. Then* $\mathrm{diam}(\mathrm{Cay}(G, X)) \leq n^{O(1)}$ *provided that $|K_0| > 2$.*

2. Preliminaries

Throughout this paper, we use the notation $\ell_X(Y)$ for the length of Y over X. This is defined as follows. For any $X, Y \subset G$ with $Y \subset \langle X \rangle$ let $\ell_X(Y)$ be the smallest number k such that every element of Y can be written as a product of at most k elements from $X \cup X^{-1}$, that is,

$$\ell_X(Y) = \min\{k \in \mathbb{N} \mid Y \subset (X \cup X^{-1} \cup 1)^k\}.$$

Clearly, ℓ has the property $\ell_X(Z) \leq \ell_X(Y) \cdot \ell_Y(Z)$ for any $X, Y, Z \subset G$ with $Y \subset \langle X \rangle$, $Z \subset \langle Y \rangle$, which makes us possible to "cut" the proof of Theorem 1.5 into steps providing larger and larger generating sets having stronger and stronger properties. Using the next lemma, as a first step we can assume that X contains only transvections. We formulate it as a more general statement.

LEMMA 2.1: *In order to prove Babai's conjecture for quasisimple groups, it is sufficient to assume that the generating set X consists of conjugate elements.*

Proof. Let X be any generating set of the quasisimple group G. Then there exists a non-central element $t \in X$. For $m > 0$ an integer, let

$$Y_m := \{x_m \cdots x_1 t x_1^{-1} \cdots x_m^{-1} : x_1, \ldots, x_m \in X \cup \{1\}\}.$$

Consider the ascending chain $\{\langle Y_i \rangle\}_{i=1}^{\infty}$ of subgroups of G. Assume that

$$\langle Y_r \rangle = \langle Y_{r+1} \rangle = H$$

for some r and some subgroup H of G. It follows that H is closed under conjugation by elements of X. Since X is a generating set of G, the subgroup H must be normal in G. Since G is quasisimple, H is either central or equal to G. The subgroup H cannot be central, since it contains conjugates of t. We conclude that $H = G$. It also follows that the length of the chain $\{\langle Y_i \rangle\}_{i=1}^{\infty}$ is at most $r \leq \log_2 |G|$. Thus $Y := Y_m$ is a generating set for G for any integer m at least $\log_2 |G|$. Furthermore, we have

$$\ell_X(G) \leq \ell_X(Y) \cdot \ell_Y(G) \leq (2 \log_2 |G| + 1) \ell_Y(G).$$

So, if Babai's bound holds for $\mathrm{diam}(\mathrm{Cay}(G, Y)) = \ell_Y(G)$, then it also holds for $\mathrm{diam}(\mathrm{Cay}(G, X))$ (with a slightly larger c). ∎

2.1. LINEAR, SYMPLECTIC AND UNITARY GROUPS. Let F be any field and n a positive integer. Let V be a vector space of dimension n over F. Usually F will denote the finite field $\mathbb{F}_q$ of order q.

The group of all linear transformations on V is denoted by $\mathrm{GL}(V)$ or $\mathrm{GL}(n, F)$ or $\mathrm{GL}(n, q)$ in case $F = \mathbb{F}_q$. Let $f : V \times V \to F$ be a map. An element g in $\mathrm{GL}(V)$ is said to preserve f if $f(gu, gv) = f(u, v)$ for all u, $v \in V$. The set of all elements of $\mathrm{GL}(V)$ preserving f is called the isometry group of f. A vector $v \in V$ is called singular if $f(v, v) = 0$.

We will assume that the map f is any of two types. It will be a symplectic form, that is, a non-degenerate bilinear alternating form, or it will be a unitary form, that is, a non-degenerate conjugate-symmetric sesquilinear form. Note that f is called an alternating form if $f(v, v) = 0$ for all $v \in V$. From this it follows that f is skew-symmetric, that is, $f(u, v) = -f(v, u)$ for all u, $v \in V$. In case f is a symplectic form, n must be even and the isometry group of f is the symplectic group $\mathrm{Sp}(V)$ or $\mathrm{Sp}(n, F)$ or $\mathrm{Sp}(n, q)$ in case $F = \mathbb{F}_q$. Now

let f be a unitary form. Assume q is the square of an integer q_0. Let σ be the automorphism of F defined by the identity $\sigma(\lambda) = \lambda^{q_0}$ for every $\lambda \in F$. The form f is called a conjugate-symmetric sesquilinear form if

$$f(u, \lambda v + w) = \lambda f(u, v) + f(u, w) \quad \text{and} \quad f(w, v) = f(v, w)^{\sigma}.$$

The isometry group of a unitary form is the general unitary group $\mathrm{GU}(V)$ or $\mathrm{GU}(n, q)$. The group $\mathrm{GU}(n, q)$ has a subgroup, called the special unitary group $\mathrm{SU}(V)$ or $\mathrm{SU}(n, q_0)$ of index $q_0 + 1$ consisting of all elements with determinant 1.

Throughout the paper G will denote $\mathrm{Sp}(V)$, $\mathrm{SU}(V)$, or the special linear group $\mathrm{SL}(V)$. We will assume that $n \geq 4$. In particular, G is a quasisimple group.

2.2. Transvections. The notion of a transvection was probably first introduced by Artin. In this subsection we collect some basic facts about transvections which can be found in [1].

In [1, p. 160] an element t of $\mathrm{GL}(V)$ is called a transvection if it keeps every vector of some hyperplane W fixed and moves any vector $v \in V$ by some vector of W, that is, $t(v) - v \in W$ and $t(v) - v \neq 0$ if $v \notin W$.

Artin determined the form of a transvection. Let $\phi \neq 0$ be an element of the dual space V^* of V. The set of all vectors $v \in V$ such that $\phi(v) = 0$ is a hyperplane W. If $\psi \in V^*$ also describes W, then $\psi = c \cdot \phi$ for some $c \in F^{\times}$. Let t be a transvection and let W be the associated hyperplane. Let $\phi \in V^*$ be associated to W. Then t has the form $t(x) = x + \phi(x)w$ for $x \in V$ and for some fixed w in W, that is, $\phi(w) = 0$. Conversely, any map of this form has fixed point space the hyperplane associated to ϕ and every vector in V is moved by a multiple of w. If $t \neq 1$, then the 1-dimensional subspace of W spanned by w is called the direction of t.

It can be easily seen that every transvection is inside $\mathrm{SL}(V)$, and they are all conjugate to each other in $\mathrm{GL}(V)$. Moreover, if $n \geq 3$, then they are all conjugate even in $\mathrm{SL}(V)$.

2.3. Transvections in $\mathrm{SL}(V)$, $\mathrm{Sp}(V)$ and $\mathrm{SU}(V)$. Recall that throughout this paper G is any of the groups $\mathrm{SL}(V)$, $\mathrm{Sp}(V)$, $\mathrm{SU}(V)$ where F is the finite field $\mathbb{F}_q$ and V is a finite-dimensional vector space over F. In each case, let $\mathcal{T}$ denote the set of all transvections in G.

In the following we borrow several notions and concepts from [12]. Fix $0 \neq u \in V$ and $0 \neq \phi \in V^* = \mathrm{Hom}(V, F)$ such that $\phi(u) = 0$. The element $u \otimes \phi \in V \otimes V^*$ may be viewed as an endomorphism of V and $t = 1 + u \otimes \phi$ is a transvection satisfying $t(x) = x + \phi(x)u$ for all $x \in V$. In fact, the set of all transvections in $\mathrm{SL}(V)$ is

$$\mathcal{T}(\mathrm{SL}(V)) = \{1 + u \otimes \phi \mid 0 \neq u \in V, 0 \neq \phi \in V^*, \phi(u) = 0\}.$$

Note that $1 + u \otimes \phi = 1 + v \otimes \psi$ for some other choice of $0 \neq v \in V$ and $0 \neq \psi \in V^*$ such that $\psi(v) = 0$, if and only if, $v = \lambda u$ and $\phi = \lambda\psi$ for some nonzero $\lambda \in F$.

From now on, if we write $1 + v \otimes \psi$ for a transvection we assume $0 \neq v \in V$ and $0 \neq \psi \in V^*$ with $\psi(v) = 0$.

Let f be a symplectic or unitary form on V. For $u \in V$ let $\varphi_u \in V^* = \mathrm{Hom}(V, F)$ be the map defined as $\varphi_u(x) = f(u, x)$ for all $x \in V$. Let $t = 1 + u \otimes \phi$ be a transvection and let x, $y \in V$ be arbitrary subject to the conditions $x \in \ker(\phi)$ and $y \notin \ker(\phi)$. Assume that t preserves f. Then $f(x, y) = f(tx, ty) = f(x, y + \phi(y)u)$. From this it follows that $\phi(y)f(x, u) = 0$, that is, $x \in \ker(\varphi_u)$. Thus

$$\ker(\phi) = \ker(\varphi_u)$$

since both spaces have dimension $n - 1$. This is equivalent to saying that $\phi = \lambda\varphi_u$ for some $0 \neq \lambda \in F$. Notice also that $f(u, u) = \varphi_u(u) = \lambda^{-1}\phi(u) = 0$. We conclude that the set of transvections in $\mathrm{SL}(V)$ preserving the form f is contained in the set $\{1 + \lambda u \otimes \varphi_u \mid \lambda \in F^\times, 0 \neq u \in V, f(u, u) = 0\}$.

In case f is a symplectic form, the elements of this latter set are called symplectic transvections and are precisely the transvections contained in $\mathrm{Sp}(V)$ (see [25, Exercise 3.20]), so we have

$$\mathcal{T}(\mathrm{Sp}(V)) = \{1 + \lambda u \otimes \varphi_u \mid \lambda \in F^\times, 0 \neq u \in V, f(u, u) = 0\}.$$

Finally, in case f is a unitary form, the transvection $1 + \lambda u \otimes \varphi_u$ with $\lambda \in F$, $0 \neq u \in V$ and $f(u, u) = 0$ preserves f if and only if $\mathrm{Tr}(\lambda) := \lambda + \lambda^{q_0} = 0$ (see [25, Exercise 3.22]), so

$$\mathcal{T}(\mathrm{Sp}(U)) = \{1 + \lambda u \otimes \varphi_u \mid 0 \neq u \in V,\ f(u, u) = 0,\ \lambda \in F^\times,\ \mathrm{Tr}(\lambda) = 0\}.$$

2.4. Conjugating Transvections with Each Other. By Lemma 2.1, we can assume that X contains only transvections. Starting from X, our goal is to create the conjugate class of all the transvections. During the proof, our main tool to achieve this goal will be to take conjugates $t_2 t_1 t_2^{-1}$ for some already

generated transvections t_1 and t_2. The following lemma will be used many times during the proof.

LEMMA 2.2: *Let $t_1 := 1 + a_1 \otimes \phi_1$ and $t_2 := 1 + a_2 \otimes \phi_2$ be two transvections. Then*

$$t_2 t_1 t_2^{-1} = 1 + (a_1 + \phi_2(a_1)a_2) \otimes (\phi_1 - \phi_1(a_2)\phi_2).$$

Proof. For $x \in V$, we have

$$\begin{aligned}
&t_2 t_1 t_2^{-1}(x) \\
&\quad = (1 + a_2 \otimes \phi_2)(1 + a_1 \otimes \phi_1)(1 - a_2 \otimes \phi_2)(x) \\
&\quad = (1 + a_2 \otimes \phi_2)(1 + a_1 \otimes \phi_1)(x - \phi_2(x)a_2) \\
&\quad = (1 + a_2 \otimes \phi_2)(x - \phi_2(x)a_2 + \phi_1(x)a_1 - \phi_2(x)\phi_1(a_2)a_1) \\
&\quad = x + \phi_1(x)a_1 - \phi_2(x)\phi_1(a_2)a_1 + \phi_2(a_1)\phi_1(x)a_2 - \phi_2(a_1)\phi_2(x)\phi_1(a_2)a_2 \\
&\quad = x + (\phi_1(x) - \phi_1(a_2)\phi_2(x)) \cdot (a_1 + \phi_2(a_1)a_2) \\
&\quad = (1 + (a_1 + \phi_2(a_1)a_2) \otimes (\phi_1 - \phi_1(a_2)\phi_2))(x).
\end{aligned}$$

This proves the lemma. ∎

2.5. TRANSVECTION GROUPS. Let $t = 1 + u \otimes \phi \in \mathcal{T}$. For any $\lambda \in \mathbb{F}_q^\times$, let the transvection $1 + \lambda u \otimes \phi$ be denoted by t^λ. We have noted that $t^\lambda t^\mu = t^{\lambda+\mu}$ for any λ and μ in $\mathbb{F}_q^\times$. In particular, $t^{-1} = 1 + (-u) \otimes \phi = 1 - u \otimes \phi$.

For an arbitrary subset $\Lambda \subset \mathbb{F}_q$, let $t^\Lambda := \{t^\lambda \mid \lambda \in \Lambda\}$. This set is a group if and only if Λ is a subgroup of the additive group of $\mathbb{F}_q$.

In case G is a unitary group the notation q_0 was introduced to be the prime power which is the square root of q. For a unified treatment, we set $q_0 = q$ in the cases when G is a special linear or a symplectic group. Using this notation, for any $t \in \mathcal{T}$ we have that $\mathbb{F}_{q_0}$ is the largest subset Λ of $\mathbb{F}_q$ such that $t^\Lambda \subset \mathcal{T}$. We call $t^{\mathbb{F}_{q_0}}$ the full transvection subgroup of G containing t. Clearly, $t^{\mathbb{F}_{q_0}}$ contains the cyclic group $\langle t \rangle$ and the containment is proper if and only if q_0 is not a prime. More generally, for any subfield K of $\mathbb{F}_{q_0}$, we call t^K the transvection subgroup over K containing t. Let $Y \subset \mathcal{T}$ be any subset of transvections in G. For any subfield K of $\mathbb{F}_{q_0}$, the K-closure of Y is defined as the set $Y^K := \{t^\lambda \mid t \in Y, \lambda \in K^\times\}$. We say that Y is K-closed if $Y^K = Y$.

Let G be a symplectic or a unitary group. In this case for every singular vector $0 \neq v \in V$, there is a unique associated full transvection subgroup, which we denote by T_v. In the symplectic case, let $\lambda_0 = 1$ and let $\mathcal{F} = \mathbb{F}_{q_0} = \mathbb{F}_q$, while

in the unitary case, let λ_0 be a fixed element of $\mathbb{F}_q$ with $\mathrm{Tr}(\lambda_0) = 0$, and let $\mathcal{F} = \{\lambda \in \mathbb{F}_q \mid \mathrm{Tr}(\lambda) = 0\}$. In both cases $1 + \lambda_0 \cdot v \otimes \varphi_v \in \mathcal{T}$. Now $\mathcal{F} = \mathbb{F}_{q_0}\lambda_0$ is a one-dimensional $\mathbb{F}_{q_0}$-subspace of $\mathbb{F}_q$ and

$$T_v = (1 + v \otimes \varphi_v)^{\mathcal{F}} = (1 + \lambda_0 \cdot v \otimes \varphi_v)^{\mathbb{F}_{q_0}}.$$

Note that $T_v = T_{\lambda v}$ for every singular vector $0 \neq v \in V$ and for every $\lambda \in F_q^\times$.

2.6. Transvection Graph. As before, let G be any of the groups $\mathrm{SL}(V)$, $\mathrm{Sp}(V)$, $\mathrm{SU}(V)$ and let $\mathcal{T}$ be the set of all transvections in G. Again, let F be the finite field $\mathbb{F}_q$. Let Y be any subset of $\mathcal{T}$. We introduce the transvection graph $\Gamma(Y)$ (as in [7]) and its labelled version $\widetilde{\Gamma}(Y)$.

The directed graph $\Gamma(Y)$ has vertex set Y and two vertices $s = 1 + u \otimes \phi$ and $t = 1 + v \otimes \psi$ are connected by a directed edge $[s, t]$ running from s to t if and only if $\psi(u) \neq 0$. The set of edges in $\Gamma(Y)$ will be denoted by $E(Y)$. We say that the edge $[s, t] \in E(Y)$ is one-way directed if $[t, s] \notin E(Y)$, otherwise (s, t) is called a two-way directed edge. This difference of notation will become clear when we introduce cycles.

In the symplectic and unitary case, for any two transvections $s = 1+\lambda\cdot u\otimes\varphi_u$ and $t = 1+\mu\cdot v\otimes\varphi_v$ in Y we have $[s, t] \in E(Y)$ if and only if $\varphi_v(u) = f(v, u) \neq 0$ if and only if $[t, s] \in E(Y)$. Hence every edge in $\Gamma(Y)$ is two-way directed and $\Gamma(Y)$ can be seen as an undirected graph. In that case $\Gamma(Y)$ is just the same as the non-commuting graph (which was probably first mentioned in [22]) of G restricted to the vertex set Y. In general, $st \neq ts$ if and only if at least one of $[s, t] \in E(Y)$ and $[t, s] \in E(Y)$ holds. Therefore, in the linear case we can think of $\Gamma(Y)$ as a refinement of the non-commuting graph on Y.

We next define the labelled transvection graph $\widetilde{\Gamma}(Y)$. First, for every $s \in Y$, we fix $u_s \in V$ and $\phi_s \in V^*$ satisfying $s = 1 + u_s \otimes \phi_s$. Then $\widetilde{\Gamma}(Y)$ is a complete directed graph with label $l(s, t) = \phi_t(u_s) \in \mathbb{F}_q$ for every $(s, t) \in Y \times Y$. (Note that $l(s, t)$ does not only depend on s and t but also on how u_s, ϕ_s, u_t, ϕ_t were chosen.) Note that $\Gamma(Y)$ can be derived from $\widetilde{\Gamma}(Y)$ by deleting all labels and all edges with label 0.

In the special case when $Y = \mathcal{T}$ we get the full transvection graph $\Gamma(\mathcal{T})$ and its labelled version $\widetilde{\Gamma}(\mathcal{T})$. For $Y \subseteq \mathcal{T}$, the graphs $\Gamma(Y)$ and $\widetilde{\Gamma}(Y)$ can be seen as the subgraphs of $\Gamma(\mathcal{T})$ and $\widetilde{\Gamma}(\mathcal{T})$ induced by Y, respectively.

We now define the weight of a cycle in $\widetilde{\Gamma}(Y)$. This concept will be the main tool in Section 4.2.

For any integer $k \geq 2$ and for any transvections $s_1, s_2, \ldots, s_k \in Y$, let

$$w(s_1, s_2, \ldots, s_k) := l(s_1, s_2)l(s_2, s_3)\ldots l(s_{k-1}, s_k)l(s_k, s_1)$$

be the weight of the k-tuple $(s_1, \ldots, s_k)$. If $w(s_1, \ldots, s_k) \neq 0$ then we say that $(s_1, \ldots, s_k)$ is a cycle (or closed path) in $\Gamma(Y)$.

Note that unlike the labels $l(s_i, s_j)$, the weight $w(s_1, \ldots, s_k)$ depends only on the transvections $s_1, \ldots, s_k$. In order to distinguish paths from cycles, we use the notation $[s_1, s_2, \ldots, s_k]$ for a directed path of length k and the notation $(s_1, s_2, \ldots, s_k)$ for a directed k-cycle.

2.7. The Dual of a Transvection Graph. We introduce the dual of a transvection graph $\Gamma(X)$. For this goal, let $f : V \times V \to \mathbb{F}_q$ be any fixed non-degenerate symmetric bilinear function on V. Then for any $x \in \mathrm{End}(V)$ let $x^* \in \mathrm{End}(V)$ be the adjoint of x with respect to f, that is,

$$f(x(u), v) = f(u, x^*(v))$$

for any $u, v \in V$. Furthermore, for any $u \in V$ let again $\varphi_u \in V^*$ defined as $\varphi_u(v) = f(u, v)$ for all $v \in V$, so

$$\mathcal{T} = \{1 + u \otimes \varphi_v \,|\, u, v \in V \setminus \{0\},\, f(u, v) = 0\}.$$

Then we have the following.

Lemma 2.3:

(1) *The map $x \to x^*$ is $\mathbb{F}_q$-linear, which is product-reversing, that is, $(x_1x_2)^* = x_2^*x_1^*$ for every $x_1, x_2 \in \mathrm{End}(V)$. As a consequence*

$$\ell_X(Y) = \ell_{X^*}(Y^*)$$

for every $X, Y \subset \mathrm{SL}(V)$.

(2) *If $t = 1 + u \otimes \varphi_v$ is a transvection for some $u, v \in V$, then t^* is also a transvection. Moreover, $t^* = 1 + v \otimes \varphi_u$.*

(3) *For any set X of transvections, let $X^* := \{t^* \,|\, t \in X\}$. Then the adjoint map defines a weight-preserving anti-isomorphism $\Gamma(X) \to \Gamma(X^*)$, which means that*

$$(t_1, t_2) \in E(X) \iff (t_2^*, t_1^*) \in E(X^*) \text{ and } w(t_1, t_2, \ldots, t_k) = w(t_k^*, t_{k-1}^*, \ldots, t_1^*)$$

for any $t_1, t_2, \ldots, t_k \in X$.

Proof. It is well known that the adjoint map is a product-reversing $\mathbb{F}_q$-linear transformation of $\mathrm{End}(V)$. Therefore, for any elements $x_1, x_2, \ldots, x_k \in X$ and $y \in Y$, we have

$$y = x_1 x_2 \cdots x_k \iff y^* = x_k^* x_{k-1}^* \cdots x_1,$$

so $\ell_X(Y) = \ell_{X^*}(Y^*)$.

Let $t = 1 + u \otimes \varphi_v$ be a transvection with $f(u, v) = 0$. Then for every $x, y \in V$ we have

$$\begin{aligned} &f(t(x), y) \\ &\quad = f(x + \varphi_v(x)u, y) = f(x, y) + \varphi_v(x) f(u, y) = f(x, y) + f(v, x) f(u, y) \\ &\quad = f(x, y) + \varphi_u(y) f(x, v) = f(x, (1 + v \otimes \varphi_u) y), \end{aligned}$$

which means that $t^* = 1 + v \otimes \varphi_u$. Since $f(v, u) = f(u, v) = 0$, it is clear that t^* is a transvection.

Let $t_1, t_2, \ldots, t_k \in X$ and write each t_i as $t_i = 1 + u_i \otimes \varphi_{v_i}$ where $f(u_i, v_i) = 0$. Now,

$$\begin{aligned} (t_1, t_2) \in E(X) &\iff \varphi_{v_2}(u_1) \neq 0 \iff f(v_2, u_1) \neq 0 \\ &\iff \varphi_{u_1}(v_2) \neq 0 \iff (t_2^*, t_1^*) \in E(X^*). \end{aligned}$$

Furthermore,

$$w(t_1, \ldots, t_k) = \prod_{i=1}^{k} \varphi_{v_{i+1}}(u_i) = \prod_{i=1}^{k} \varphi_{u_i}(v_{i+1}) = w(t_k^*, \ldots, t_1^*).$$

This concludes the proof. ∎

3. Determining groups generated by transvections in terms of the transvection graph

Throughout this section let $Y \subset \mathcal{T} = \mathcal{T}(SL(V))$ be any fixed subset of transvections and $H = \langle Y \rangle \leq \mathrm{SL}(V)$. The goal of this section is to give answers to the following general question: How do the properties of the (weighted) transvection graph $\Gamma(Y)$ reflect the properties of the generated subgroup $H = \langle Y \rangle \leq \mathrm{SL}(V)$?

In the following we give several conditions of this type.

3.1. Determining the Irreducibility of H. Recall that a directed graph Γ is called strongly connected if for every two distinct vertices a, b in Γ there exists a directed path from a to b. It is easy to see that strongly connectedness is equivalent to the following condition: for every non-empty proper subset Z of vertices of Γ there exists an edge going from a vertex inside Z to a vertex outside Z.

We define the V-part and the V^*-part of $Y \subset \mathcal{T}$ as follows

$$ {}_V Y := \{v \in V \mid \exists \phi \in V^* \text{ s.t. } 1 + v \otimes \phi \in Y\},$$
$$ Y_{V^*} := \{\phi \in V^* \mid \exists v \in V \text{ s.t. } 1 + v \otimes \phi \in Y\}.$$

Theorem 3.1: *H acts irreducibly on V if and only if the following three conditions hold.*

(1) *${}_V Y$ spans V;*
(2) *Y_{V^*} spans V^*;*
(3) *$\Gamma(Y)$ is strongly connected.*

Proof. We may assume that the dimension of V is at least 2 and that $|Y| \geq 2$.

Let $0 \neq u \in V$ and let $U := \langle h(u) : h \in H \rangle$. Observe that since U is the smallest H-invariant subspace containing u, the condition that H acts irreducibly on V is equivalent to $U = V$ for every $0 \neq u \in V$.

Let $t_1 = 1 + a_1 \otimes \phi_1$ and $t_2 = 1 + a_2 \otimes \phi_2$ be two distinct transvections in Y. Assume that there is a directed edge in $\Gamma(Y)$ from t_1 to t_2, that is, $\phi_2(a_1) \neq 0$. Observe that if $a_1 \in U$, then $a_2 = \phi_2(a_1)^{-1} \cdot (t_2(a_1) - a_1) \in U$, since U is H-invariant.

Assume that the three conditions of the statement hold.

Since Y_{V^*} spans V^*, there exists $t = 1 + a \otimes \phi \in Y$ such that $\phi(u) \neq 0$. Observe that $a = \phi(u)^{-1} \cdot (t(u) - u) \in U$.

Since $\Gamma(Y)$ is strongly connected, the fifth and the third paragraphs imply that ${}_V Y \subseteq U$. The first condition provides $V = U$.

Assume now that H acts irreducibly on V.

Let $t = 1 + w \otimes \phi \in Y$ be arbitrary and let Z be the set of all transvections in Y that may be reached from t by a directed path in $\Gamma(Y)$. Observe that $\langle {}_V Z \rangle$ is H-invariant because if $w_0 \in {}_V Z$ and $y = 1 + v \otimes \psi \in Y$ then $y(w_0) = w_0 + \psi(w_0)v \in \langle {}_V Z \rangle$ by the definition of $\Gamma(Y)$ and of Z. Moreover $w \in {}_V Z$, so $\langle {}_V Z \rangle = V$. Part (1) follows. We claim that $Z = Y$. Assume for a contradiction that $Y \setminus Z \neq \emptyset$. Let $\psi \in (Y \setminus Z)_{V^*}$. Then ψ vanishes on ${}_V Z$

and so also on V, which is impossible. Since t was chosen arbitrarily, $\Gamma(Y)$ must be strongly connected, giving (3).

We now prove (2). Let $L := \langle Y_{V^*} \rangle \leq V^*$. Then Y_{V^*} contains a basis $\{\phi_1, \ldots, \phi_m\}$ for L, where $m \leq n = \dim(V)$. Note that

$$I := \bigcap_{j=1}^{m} \ker(\phi_j) = \{0\},$$

for if $0 \neq v \in I$ then v is fixed by H contradicting the fact that H acts irreducibly on V (and the dimension of V is at least 2). We have $m \leq n$ subspaces of V of codimension 1 whose intersection is trivial. This implies that $m = n$, in other words $L = V^*$. ∎

3.2. Determining the Defining Field for H. The next problem we deal with is to determine the smallest subfield $L \leq \mathbb{F}_q$ such that H is realisable over L, that is, such that H is conjugate to a subgroup of $\mathrm{SL}(n, L) \leq \mathrm{SL}(V)$. It turns out that L can be determined from the weights of cycles of $\Gamma(Y)$.

Proposition 3.2 ([8, Proposition 2]): *The field generated by the weights of all the cycles in $\Gamma(Y)$ is equal to the field generated by the traces of the matrices in H.*

Remark 3.3: If H is irreducible on V, then the field generated by the traces of the matrices in $H \leq \mathrm{SL}(V)$ is exactly the smallest subfield L such that H is realisable over L (see [18, Corollary 9.23]). Thus, the field generated by the weights of all the cycles in $\Gamma(Y)$ is equal to the smallest subfield L of $\mathbb{F}_q$ such that H conjugates into $\mathrm{SL}(n, L) \leq \mathrm{SL}(n, q) \simeq \mathrm{SL}(V)$ (see [8, Corollary 1]).

Lemma 3.4: *Let G be one of $\mathrm{SL}(n, q), \mathrm{Sp}(n, q)$ or $\mathrm{SU}(n, q)$. Moreover, assume that $n \geq 3$ if G is the unitary group. Then the traces of the elements of G generate $\mathbb{F}_q$.*

Proof. Let V be the underlying n-dimensional space over $\mathbb{F}_q$, so G can be identified with $\mathrm{SL}(V), \mathrm{Sp}(V)$ or $\mathrm{SU}(V)$.

First let us assume that $G = \mathrm{SL}(V)$ or $\mathrm{Sp}(V)$. Let us choose a decomposition $V = U \oplus U'$ with $\dim(U) = 2$. In case of $G = \mathrm{Sp}(V)$ let us also assume that U is a non-degenerate subspace of V and $U' = U^{\perp}$. Let $\{x, y\}$ be a basis of U, symplectic in case $G = \mathrm{Sp}(V)$, and let $g \in \mathrm{GL}(V)$ satisfying

$$g(x) = y + (\lambda - n + 2)x,\ g(y) = -x,\ g(u) = u \text{ for all } u \in U'.$$

Then g has trace equal to λ, and it belongs to G.

Now, let $G = \mathrm{SU}(V)$. Let U be a 3-dimensional non-degenerate subspace of V, so $V = U \oplus U^\perp$ and U possesses a basis x, y, z with

$$f(x,x) = f(y,y) = f(x,z) = f(y,z) = 0, \quad f(x,y) = f(z,z) = 1.$$

Let b be a generator of the cyclic group $\mathbb{F}_q^\times$ and let $g \in \mathrm{SU}(V)$ be defined as

$$g(x) = b^{-q_0} y, \quad g(y) = bx, \quad g(z) = -b^{q_0-1} z, \quad g(u) = u \text{ for all } u \in U^\perp.$$

Then the trace of g is $-b^{q_0-1} + n - 3$. Let p^r be the cardinality of the subfield generated by the traces of elements. Note that $b^{q_0-1} \in \mathbb{F}_{p^r}$ and the order of b^{q_0-1} is q_0+1. Then $p^r - 1$ is divisible by q_0+1, which implies that $p^r = q$. ∎

3.3. Theorems of Dickson and Wagner. The first important result regarding subgroups generated by transvections is due to Dickson, who gave a full description of subgroups of $\mathrm{SL}(2,q)$ generated by two non-commuting transvections.

Theorem 3.5 (Dickson, see [11, Chapter 2, Theorem 8.4]): *Assume r is the power of an odd prime. Let δ be a generator of $\mathbb{F}_r$ and set*

$$L = \left\langle \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ \delta & 1 \end{pmatrix} \right\rangle.$$

Then we have either

(1) $L = \mathrm{SL}(2, \mathbb{F}_r)$ *or*
(2) $r = 9$, $|Z(L)| = 2$, $L/Z(L) \cong A_5$, *and L contains a subgroup isomorphic to* $\mathrm{SL}(2,3)$.

Dickson also gave a full description of subgroups of $\mathrm{SL}(2,q)$. The following may be found in [17, pp. 213–214] and [9, p. 285].

Theorem 3.6 (Dickson's Theorem): *Let p be a prime and f a positive integer. The subgroups of* $\mathrm{PSL}(2, p^f)$ *are the following.*

(i) *Elementary abelian p-group.*
(ii) *Cyclic group whose order z divides $(p^f \pm 1)/k$, where $k = (p^f - 1, 2)$.*
(iii) *D_{2z}, where z is as in (ii).*
(iv) *A_4, where $p > 2$ or $p = 2$ and $f \equiv 0 \pmod 2$.*
(v) *S_4, where $p^{2f} - 1 \equiv 0 \pmod{16}$.*
(vi) *A_5, where $p = 5$ or $p^{2f} - 1 \equiv 0 \pmod 5$.*

(vii) *The semidirect product of an elementary abelian group of order* p^m *and a cyclic group of order* t*, where* $t \mid p^m - 1$ *and* $t \mid p^f - 1$.
(viii) $\mathrm{PSL}(2, p^m)$*, where* $m \mid f$ *and* $\mathrm{PGL}(2, p^m)$*, where* $2m \mid f$.

Let again V be an n-dimensional vector space over $\mathbb{F}_q$. In most of our cases n will be unbounded. We will be interested in subspaces U of V of bounded dimension. Let U be a subspace of V of dimension k. Let W be another subspace of V such that $V = U \oplus W$. In particular, the dimension of W is $n - k$. Let 1_W denote the identity map on W. In the special case when $W = V$ we denote 1_W by 1.

Let H be a subgroup of $\mathrm{SL}(U)$. Let us denote the subgroup

$$\{g \in \mathrm{SL}(V) \mid g(U) = U,\, g(W) = W,\, g_U \in H,\, g_W = 1_W\} \leq \mathrm{SL}(V)$$

by $H \oplus 1_W$. Usually we use this construction for $H = \mathrm{SL}(U)$, or $\mathrm{Sp}(U)$ or $\mathrm{SU}(U)$.

In a similar way, for a matrix group $M \leq \mathrm{SL}(k, q)$, let $M \oplus 1_{n-k}$ denote the subgroup

$$\left\{ m \oplus 1_{n-k} := \begin{pmatrix} m & 0 \\ 0 & 1_{n-k} \end{pmatrix} \,\middle|\, m \in M \right\},$$

where 1_{n-k} denotes the identity matrix of size $(n-k) \times (n-k)$. Let H be a subgroup of $\mathrm{SL}(V)$ and let M be a subgroup of $\mathrm{SL}(k, q)$. By writing

$$H \simeq M \oplus 1_{n-k}$$

it is meant that there is a direct sum decomposition $V = U \oplus W$ with $\dim(U) = k$ and a subgroup $H_0 \simeq H$ of $\mathrm{SL}(U)$ such that $H = H_0 \oplus 1_W$ and the matrix form of H_0 is equal to M in a suitable basis of U. In the symplectic and unitary cases when we write $H \simeq M \oplus 1_{n-k}$ we tacitly assume that the underlying decomposition $V = U \oplus W$ is orthogonal with respect to f, that is, f_U is non-degenerate and $W = U^\perp$. In these cases H_0 will be a subgroup of $\mathrm{Sp}(U)$ or $\mathrm{SU}(U)$.

Now, we give a partial generalisation of Theorem 3.5.

THEOREM 3.7: *Let* $(s, t) \in E(\mathcal{T})$ *be a two-way directed edge and let* $K, \Lambda \subseteq \mathbb{F}_q^\times$ *containing* 1*. Set* $\delta := w(s, t)$ *and* $M := \mathbb{F}_p(K, \Lambda, \delta)$*. Then*

$$\langle s^K, t^\Lambda \rangle \simeq \mathrm{SL}(2, M) \oplus 1_{n-2},$$

unless $p = 2$ *or* $M = \mathbb{F}_9$.

Proof. Write the transvections as $s = 1 + v \otimes \phi$ and $t = 1 + w \otimes \psi$. Since (s, t) is a two-way edge, $\phi(w) \neq 0$ and $\psi(v) \neq 0$. Therefore

$$V = \langle v, w \rangle \oplus (\ker(\phi) \cap \ker(\psi)).$$

Choose a basis of V whose first two vectors are $\phi(w)v, w$ and the other vectors form a basis of the $n-2$-dimensional space $\ker(\phi) \cap \ker(\psi)$. Then we get

$$H = \langle s^K, t^\Lambda \rangle \simeq H_0 \oplus 1_{n-2}$$

for some $H_0 \leq \mathrm{SL}(2, q)$. Therefore we may assume that $n = 2$, $V = \langle v, w \rangle$ and $H \leq \mathrm{SL}(2, q)$. Furthermore, in the basis $\{\phi(w)v, w\}$ the transvections s^κ, t^λ, for $\kappa \in K$, $\lambda \in \Lambda$, have matrix form

$$\begin{pmatrix} 1 & \kappa \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ \delta\lambda & 1 \end{pmatrix}.$$

Now, the image of H in $\mathrm{PSL}(2, q)$ (denoted by $\bar{H}$) is one of the groups appearing in Theorem 3.6. $\bar{H}$ has two non-commuting p-subgroups, namely the images of $\langle s^{\mathrm{K}} \rangle$ and $\langle t^{\Lambda} \rangle$. Thus, $\bar{H}$ cannot be of type (i), (ii) or (vii). For the remainder of the proof, let us assume that $p \neq 2$. Then $\bar{H}$ cannot be of type (iii). Now, let us assume that $p = 3$ and $\bar{H}$ is any of type (iv) or (v) or (vi), then we must have $\mathrm{K}\Lambda \subset \mathbb{F}_3$. If $\delta \notin \mathbb{F}_9$, then by Theorem 3.5 we have

$$|H| = |\langle s, t \rangle| = |\mathrm{SL}(2, \mathbb{F}_3(\delta))| \geq |\mathrm{SL}(2, 27)| > 2|A_5| \geq |H|,$$

a contradiction. Thus $\delta \in \mathbb{F}_9$, which means that either $M = \mathbb{F}_9$ or $M = \mathbb{F}_3$. In the latter case we get that $H = \mathrm{SL}(2, 3)$ by Theorem 3.5. Now let us assume that $\bar{H}$ is of type (vi) and $p = 5$. Again, we must have $\mathrm{K}\Lambda \subset \mathbb{F}_5$. Similarly as before, if $\delta \notin \mathbb{F}_5$, then

$$|H| = |\langle s, t \rangle| = |\mathrm{SL}(2, \mathbb{F}_5(\delta))| > 2|A_5| \geq |H|,$$

a contradiction.

Thus, we get that $\bar{H}$ is of type (viii), so $H = \mathrm{SL}(2, L)$ for some subfield L of $\mathbb{F}_q$. Since M equals the subfield generated by all the weights of $\Gamma(\{s^{\mathrm{K}}, t^{\Lambda}\})$, we get that $L = M$ by Section 3.2. ∎

The analogous question, which irreducible subgroups of $\mathrm{SL}(n, q)$ are generated by transvections when $n \geq 3$, was solved by Wagner.

THEOREM 3.8 ([24, Theorem 1.1]): *Let V be a vector space over $\mathbb{F}_q$ of dimension $n \geq 3$ and let H be a subgroup of $\mathrm{SL}(V)$. Let us assume that H has the following properties.*

(1) *H is generated by transvections,*
(2) *H acts irreducibly on V,*
(3) *H contains a transvection subgroup of order larger than 2.*

Then H is isomorphic to one of $\mathrm{SL}(n, L)$, $\mathrm{SU}(n, L)$ or $\mathrm{Sp}(n, L)$ for some subfield L of $\mathbb{F}_q$.

Remark 3.9:

(1) Since the order of any transvection is $p > 2$ by our assumption, property (3) will always follow automatically.
(2) Originally, in [24] this theorem was formed for subgroups of $\mathrm{PSL}(V)$ generated by elations. (An elation is just the image of a transvection under the natural map $\mathrm{SL}(V) \to \mathrm{PSL}(V)$.) But it is easy to see that Wagner's theorem is translated into the above theorem when we consider subgroups of $\mathrm{SL}(V)$.

3.4. DETERMINING THE TYPE OF H. As we have seen in Remark 3.3, the weights of the cycles determine the smallest subfield L of $\mathbb{F}_q$ such that an irreducible subgroup of $\mathrm{SL}(n, q)$ generated by a set of transvections is realisable over L. Now, we introduce two other parameters of cycles of $\Gamma(\mathcal{T})$. With their help, one can determine whether a set of transvections generates $\mathrm{SL}(V), \mathrm{SU}(V)$ or $\mathrm{Sp}(V)$.

Definition 3.10: For any $r_1, r_2, \ldots, r_k \in \mathcal{T}$ we define

$$d_s(r_1, r_2, \ldots, r_k) := w(r_1, r_2, \ldots, r_k) + (-1)^{k+1} w(r_k, r_{k-1}, \ldots, r_1),$$
$$d_u(r_1, r_2, \ldots, r_k) := w(r_1, r_2, \ldots, r_k) + (-1)^{k+1} w(r_k, r_{k-1}, \ldots, r_1)^{\sqrt{q}}.$$

(d_u is only defined if q is a perfect square.) We say that a cycle $(r_1, \ldots, r_k) \subset \mathcal{T}$ is symplectic (or singular) if $d_s(r_1, \ldots, r_k) = 0$. Similarly, we say that $(r_1, \ldots, r_k) \subset \mathcal{T}$ is unitary if $d_u(r_1, r_2, \ldots, r_k) = 0$.

Remark 3.11:

(1) In [12], the notation $\det_c$ was used instead of d_s, but we changed it to reflect its connection with the symplectic group.
(2) By the definition of d_s and d_u, if (s, t) is a 2-cycle, that is, a two-way directed edge, then (s, t) is always symplectic, while it is unitary if and only if $w(s, t) \in \mathbb{F}_{\sqrt{q}}$.

(3) By an abuse of notation, we sometimes allow ourselves to say that an arbitrary tuple $(r_1, \ldots, r_k)$ is symplectic/unitary even if it is not a cycle (in neither direction).
(4) Clearly, a one-way directed cycle (i.e., a cycle, which is not a cycle in the reverse direction) is never symplectic or unitary, while a k-tuple which is not a cycle (in both of the two possible directions) is both symplectic and unitary.

THEOREM 3.12: *Let us assume that $n \geq 3$ and let $Z \subset \mathcal{T} \subset \mathrm{SL}(n,q)$ be a set of transvections such that Z generates an irreducible subgroup of $\mathrm{SL}(n,q)$ and the weights of all cycles of $\Gamma(Z)$ generate $\mathbb{F}_q$. Then $\langle Z \rangle$ is one of $\mathrm{SL}(n,q)$, $\mathrm{SU}(n,q)$ or $\mathrm{Sp}(n,q)$ and we have*

$$\begin{aligned}
\langle Z \rangle = \mathrm{Sp}(n,q) &\iff \text{every cycle of } \Gamma(Z) \text{ is symplectic},\\
\langle Z \rangle = \mathrm{SU}(n,q) &\iff \text{every cycle of } \Gamma(Z) \text{ is unitary},\\
\langle Z \rangle = \mathrm{SL}(n,q) &\iff \Gamma(Z) \text{ contains both a non-symplectic and}\\
&\qquad\qquad \text{a non-unitary cycle}.
\end{aligned}$$

Proof. By Theorem 3.8 and by Remark 3.3, we have that $\langle Z \rangle$ must be isomorphic to one of $\mathrm{SL}(n,q)$, $\mathrm{Sp}(n,q)$ or $\mathrm{SU}(n,q)$.

Let us assume that $\langle Z \rangle = \mathrm{SU}(n,q)$ and let $(r_1, r_2, \ldots, r_k)$ be an arbitrary cycle in $\Gamma(Z)$. Using the description of unitary transvections, each r_i can be written in the form $r_i = 1 + \lambda_i u_i \otimes \varphi_{u_i}$ where $u_i \in V$ is singular and

$$\mathrm{Tr}(\lambda_i) = \lambda_i + \lambda_i^{\sqrt{q}} = 0.$$

Then we have

$$\begin{aligned}
w(r_1, r_2, \ldots, r_k) &= \varphi_{u_2}(\lambda_1 u_1)\varphi_{u_3}(\lambda_2 u_2)\cdots\varphi_{u_1}(\lambda_k u_k)\\
&= \prod_{i=1}^{k} \lambda_i \cdot f(u_2,u_1)f(u_3,u_2)\cdots f(u_1,u_k)\\
&= (-1)^k \prod_{i=1}^{k} \lambda_i^{\sqrt{q}} \cdot f(u_1,u_2)^{\sqrt{q}} f(u_2,u_3)^{\sqrt{q}} \cdots f(u_k,u_1)^{\sqrt{q}}\\
&= (-1)^k (f(u_1,\lambda_2 u_2)f(u_2,\lambda_3 u_3)\cdots f(u_k,\lambda_1 u_1))^{\sqrt{q}}\\
&= (-1)^k w(r_k, r_{k-1}, \ldots, r_1)^{\sqrt{q}},
\end{aligned}$$

so $d_u(r_1, r_2, \ldots, r_k) = 0$. A similar calculation shows that if $\langle Z \rangle = \mathrm{Sp}(n,q)$, then every cycle in $\Gamma(Z)$ is symplectic.

Let us assume that every cycle of $\Gamma(Z)$ is symplectic. Then

$$\langle Z \rangle \leq \langle Z^{\mathbb{F}_q} \rangle = \mathrm{Sp}(n,q)$$

by [12, Corollary 5.3].

Thus, it remains to prove that if Z generates $\mathrm{SL}(n,q)$, then $\Gamma(Z)$ must contain a non-unitary cycle. We prove this by using an argument which is very similar to the one given in the proof of [12, Lemma 4.6].

For the remainder of the proof let q be a square and $\langle Z \rangle = \mathrm{SL}(n,q)$. Let us consider a set of transvections $Z' \supset Z$ and let us assume that every cycle in $\Gamma(Z')$ is unitary. If there were a one-way directed edge $(s,t) \in E(Z')$, then, since $\Gamma(Z')$ is strongly connected, there would be a directed cycle $(s,t,r_1,\ldots,r_m)$ in $\Gamma(Z')$. Clearly, such a cycle must be non-unitary. Therefore, every edge of $\Gamma(Z')$ is two-way directed. Furthermore, we know that $w(s,t) \in \mathbb{F}_{\sqrt{q}}$ for every $(s,t) \in E(Z')$ by Remark 3.11/(2). Now, for every (two-way) directed cycle $(r_1,r_2,\ldots,r_k)$ in $\Gamma(Z')$ one can define

$$P_u(r_1,\ldots,r_k) := \frac{w(r_1,r_2,\ldots,r_k)}{w(r_k,r_{k-1},\ldots,r_1)^{\sqrt{q}}}.$$

Surely, a directed cycle $(r_1,\ldots,r_k)$ is unitary if and only if $P_u(r_1,\ldots,r_k)=(-1)^k$. One can show that under the assumption that every 2-cycle is unitary, the cycle parameter P_u is well-behaved under "gluing" two-way directed cycles. More concretely, if a directed cycle $(r_1,\ldots,r_k)$ is obtained from two directed cycles $(r_1,\ldots,r_i,q_1,\ldots,q_l,r_j,\ldots,r_k)$ and $(r_i,\ldots,r_j,q_l,\ldots,q_1)$ $(i<j)$ glued along their joint subpath $r_i,q_1,q_2,\ldots,q_l,r_j$, then

$$\begin{aligned} &P_u(r_1,\ldots,r_k) \\ &\qquad = P_u(r_1,\ldots,r_i,q_1,\ldots,q_l,r_j,\ldots,r_k)\cdot P_u(r_i,\ldots,r_j,q_l,\ldots,q_1). \end{aligned} \tag{1}$$

(Observe that the joint subpath must be oppositely directed in the two cycles.) Indeed,

$$\begin{aligned} &P_u(r_1,\ldots,r_i,q_1,\ldots,q_l,r_j,\ldots,r_k)\cdot P_u(r_i,\ldots,r_j,q_l,\ldots,q_1) \\ &= \frac{w(r_1,\ldots,r_i,q_1,\ldots,q_l,r_j,\ldots,r_k)\cdot w(r_i,\ldots,r_j,q_l,\ldots,q_1)}{w(r_k,\ldots,r_j,q_l,\ldots,q_1,r_i,\ldots,r_1)^{\sqrt{q}}\cdot w(q_1,\ldots,q_l,r_j,\ldots,r_i)^{\sqrt{q}}} \\ &= \frac{w(r_1,\ldots,r_k)\cdot\prod_{s=1}^{l-1} w(q_s,q_{s+1})\cdot w(r_i,q_1)\cdot w(q_l,r_j)}{w(r_k,\ldots,r_1)^{\sqrt{q}}\cdot\prod_{s=1}^{l-1} w(q_s,q_{s+1})^{\sqrt{q}}\cdot(w(q_1,r_i)\cdot w(r_j,q_l))^{\sqrt{q}}} \\ &= P_u(r_1,\ldots,r_k). \end{aligned}$$

Let $(s_1, s_2) \in E(Z')$ be any (two-way) directed edge and let $t := s_2 s_1 s_2^{-1}$. We claim that every cycle of $\Gamma(Z' \cup \{t\})$ is unitary. Clearly, this should be checked for cycles containing t. We prove this for 2-cycles and for some 3-cycles first.

Let

$$s_1 = 1 + u_1 \otimes \phi_1, \quad s_2 = 1 + u_2 \otimes \phi_2,$$

so

$$t = 1 + (u_1 + \phi_2(u_1)u_2) \otimes (\phi_1 - \phi_1(u_2)\phi_2).$$

For any $r = 1 + v \otimes \psi \in Z'$ we have

$$\begin{aligned} w(r,t) &= (\phi_1 - \phi_1(u_2)\phi_2)(v) \cdot \psi(u_1 + \phi_2(u_1)u_2) \\ &= \phi_1(v)\psi(u_1) + \phi_1(v)\phi_2(u_1)\psi(u_2) \\ &\quad - \phi_1(u_2)\phi_2(v)\psi(u_1) - \phi_1(u_2)\phi_2(v)\phi_2(u_1)\psi(u_2) \\ &= w(s_1, r) + w(r, s_1, s_2) - w(s_2, s_1, r) - w(s_1, s_2)w(r, s_2). \end{aligned}$$

Since $(s_1, r),\ (s_1, s_2),\ (r, s_2)$ are unitary, their weights are in $\mathbb{F}_{\sqrt{q}}$. Furthermore, $w(r, s_1, s_2) = -w(s_2, s_1, r)^{\sqrt{q}}$ implies that

$$w(r, s_1, s_2) - w(s_2, s_1, r) \in \mathbb{F}_{\sqrt{q}},$$

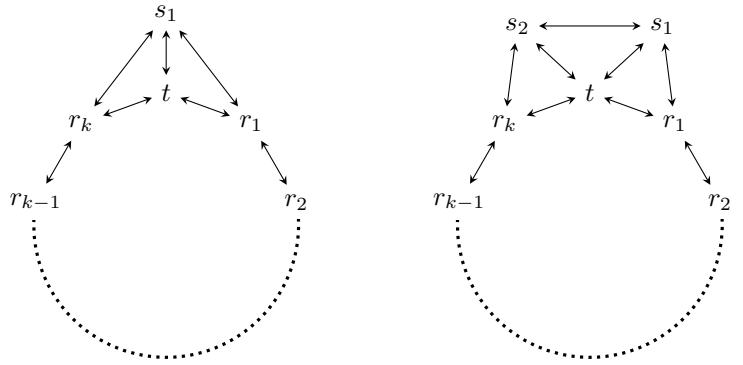
as well. So, $w(r, t) \in \mathbb{F}_{\sqrt{q}}$, that is, (r, t) is unitary.

Furthermore,

$$\begin{aligned} d_u(r, t, s_2) &= (\phi_1 - \phi_1(u_2)\phi_2)(v) \cdot \phi_2(u_1 + \phi_2(u_1)u_2) \cdot \psi(u_2) \\ &\quad + ((\phi_1 - \phi_1(u_2)\phi_2)(u_2) \cdot \psi(u_1 + \phi_2(u_1)u_2) \cdot \phi_2(v))^{\sqrt{q}} \\ &= \phi_1(v)\phi_2(u_1)\psi(u_2) - \phi_1(u_2)\phi_2(v)\phi_2(u_1)\psi(u_2) \\ &\quad + (\phi_1(u_2)\psi(u_1)\phi_2(v))^{\sqrt{q}} + (\phi_1(u_2)\phi_2(u_1)\psi(u_2)\phi_2(v))^{\sqrt{q}} \\ &= w(r, s_1, s_2) - w(s_1, s_2)w(s_2, r) + w(s_2, s_1, r)^{\sqrt{q}} \\ &\quad + w(s_1, s_2)^{\sqrt{q}} w(s_2, r)^{\sqrt{q}} \\ &= d_u(r, s_1, s_2) - w(s_1, s_2)w(s_2, r) + w(s_1, s_2)w(s_2, r) = 0. \end{aligned}$$

A similar calculation shows that $d_u(r, t, s_1) = 0$.

Finally, let $(r_1, r_2, \ldots, r_k, t)$ be any cycle in $\Gamma(Z' \cup \{t\})$. Then both r_1 and r_k are connected with at least one of s_1 and s_2. Indeed, writing $r_i = 1 + v_i \otimes \psi_i$ for every i, the label of (r_k, t) is $\phi_1(v_k) - \phi_1(u_2)\phi_2(v_k)$ and it is nonzero, so at least one of $\phi_1(v_k)$ and $\phi_2(v_k)$ must be nonzero. Depending on the role of s_1 and s_2, there are two possibilities:



Using the gluing property,

$$P_u(r_1, r_2, \ldots, r_k, t) = \frac{P_u(r_1, r_2, \ldots, r_k, s_1)}{P_u(s_1, r_1, t) \cdot P_u(s_1, t, r_k)} = \frac{(-1)^{k+1}}{(-1)^3 \cdot (-1)^3} = (-1)^{k+1}$$

in the first case and

$$\begin{aligned} P_u(r_1, r_2, \ldots, r_k, t) &= \frac{P_u(r_1, r_2, \ldots, r_k, s_2, s_1)}{P_u(s_1, r_1, t) \cdot P_u(s_1, t, s_2) \cdot P_u(s_2, t, r_k)} \\ &= \frac{(-1)^{k+2}}{(-1)^3 \cdot (-1)^3 \cdot (-1)^3} = (-1)^{k+1} \end{aligned}$$

in the second case. Hence $(r_1, r_2, \ldots, r_k, t)$ is unitary.

Starting from Z we can construct new transvections by a repeated conjugation of previously constructed transvections with each other. By our previous argument, if every cycle of $\Gamma(Z)$ were unitary, then we could never get a non-unitary cycle with such repeated conjugations. However, since $\langle Z \rangle = \mathrm{SL}(n, q)$ and all the transvections are conjugate in $\mathrm{SL}(n, q)$, we finally get all the transvections. Clearly, the full transvection graph $\Gamma(\mathcal{T})$ contains also non-unitary cycles, which proves that $\Gamma(Z)$ must contain a non-unitary cycle. ∎

4. The proof of the main theorem

By Lemma 2.1 and the previous section, in order to prove Theorem 1.5, we can assume that the following properties hold for X.

(P1) G is any of $\mathrm{SL}(V)$, $\mathrm{Sp}(V)$, $\mathrm{SU}(V)$ and $X \subset \mathcal{T}(G)$.

(P2) The sets ${}_V X$ and X_{V^*} span V and V^* respectively and $\Gamma(X)$ is strongly connected.

(P3) The weights of all the cycles of $\Gamma(X)$ generate $\mathbb{F}_q$.

(P4) If $G = \mathrm{Sp}(V)$, then every cycle in $\Gamma(\mathcal{T})$ is symplectic.

(P5) If $G = SL(V)$, then $\Gamma(X)$ contains a non-unitary cycle.

Note that all of these properties remain true when we add new transvections to X.

4.1. Decreasing the Diameter of $\Gamma(X)$. The goal of this section is to obtain a set of transvections such that the diameter of the associated transvection graph is small.

Lemma 4.1: *There is a set of transvections $X' \supset X$, with $\ell_X(X') \leq O(n)$ and with the following property. For any $s, t \in \mathcal{T}$ with $[s,t] \notin E(\mathcal{T})$ there is an $r \in X'$ such that $[s,r], [r,t] \in E(\mathcal{T})$. In other words,* $\mathrm{diam}(\Gamma(X'')) \leq 2$ *for every $X' \subseteq X'' \subseteq \mathcal{T}$.*

Proof. Let $s, t \in \mathcal{T}$ be two transvections of the form $s = 1 + v \otimes \phi$, $t = 1 + w \otimes \psi$ with $[s,t] \notin E(\mathcal{T})$. If there is no edge going from s to any element of X then $\nu(v) = 0$ for every $\nu \in X_{V^*}$, which contradicts the fact that $\langle X_{V^*} \rangle = V^*$ since $v \neq 0$. Similarly, if there is no edge going from any element of X to t then $\psi(u) = 0$ for every $u \in {}_V X$, which contradicts the fact that $\langle {}_V X \rangle = V$ since $\psi \neq 0$. Since $\Gamma(X)$ is strongly connected, there exists a path $[s, r_1, \ldots, r_k, t]$ in $\Gamma(X)$, where $r_1, \ldots, r_k \in X$. Choose such a path of minimal length.

We claim that $k \leq n+1$. To prove this, write $r_i = 1 + v_i \otimes \phi_i$ for every $i = 1, \ldots, k$. Observe that, by minimality of k, $\phi_{i+1}(v_j) = 0$ unless $j = i$ or (possibly) $j > i+1$, and $\phi_{i+1}(v_i) \neq 0$ for every $i = 1, \ldots, k-1$. Let $\alpha_1, \ldots, \alpha_{k-1} \in F$ be such that $\sum_{j=1}^{k-1} \alpha_j v_j = 0$. For any $i \in \{1, \ldots, k-1\}$ we have

$$0 = \phi_{i+1}\left(\sum_{j=1}^{k-1} \alpha_j v_j \right) = \sum_{j=i}^{k-1} \alpha_j \cdot \phi_{i+1}(v_j). \text{(Eq. i)}$$

Starting from equation $k-1$ and going backwards, using that $\phi_{i+1}(v_i) \neq 0$ for all $i = k-1, k-2, \ldots, 1$, it is clear that $\alpha_j = 0$ for all $j = 1, \ldots, k-1$. This implies that $\{v_1, \ldots, v_{k-1}\}$ is linearly independent, therefore $k \leq n+1$.

We claim that s, r, t is a path in $\Gamma(\mathcal{T})$, where

$$r = r_k r_{k-1} \cdots r_2 r_1 r_2^{-1} \cdots r_{k-1}^{-1} r_k^{-1}.$$

Observe that this claim will conclude the proof because $k \leq n+1$, so adding the element r to X for every $s, t \in \mathcal{T}$ we get a set of transvections X' with $\ell_X(X') \leq 2n+1$ satisfying the required property.

Define $t_i := r_i r_{i-1} \cdots r_2 r_1 r_2^{-1} \cdots r_{i-1}^{-1} r_i^{-1}$ for all i with $1 \leq i \leq k$, in particular $t_1 = r_1$ and $t_k = r$. Write $t_i = 1 + w_i \otimes \psi_i$. We claim that $\psi_i(v) \neq 0$ for every $i \geq 1$. If $i = 1$ then this is clear. Assume now $i \geq 2$. By Lemma 2.2,

$$t_i = r_i t_{i-1} r_i^{-1} = 1 + (w_{i-1} + \phi_i(w_{i-1})v_i) \otimes (\psi_{i-1} - \psi_{i-1}(v_i)\phi_i).$$

We deduce that there exist nonzero scalars λ_i such that $w_1 = \lambda_1 v_1$, $\psi_1 = \lambda_1^{-1}\phi_1$ and, if $i \geq 2$, then

$$w_i = \lambda_i(w_{i-1} + \phi_i(w_{i-1})v_i) \in \langle v_1, \ldots, v_i \rangle,$$
$$\psi_i = \lambda_i^{-1}(\psi_{i-1} - \psi_{i-1}(v_i)\phi_i),$$

therefore $\psi_i(v) = \lambda_i^{-1}\psi_{i-1}(v) \neq 0$ by using induction on i.

We also need $\psi(w_k) \neq 0$. Since $w_{k-1} \in \langle v_1, \ldots, v_{k-1} \rangle$, we have

$$\psi(w_k) = \psi(\lambda_k(w_{k-1} + \phi_k(w_{k-1})v_k)) = \lambda_k \phi_k(w_{k-1})\psi(v_k).$$

We will prove by induction that $\phi_{i+1}(w_i) \neq 0$ for every $i = 1, \ldots, k-1$. Note that $w_1 = \lambda_1 v_1$ hence $\phi_2(w_1) = \lambda_1 \phi_2(v_1) \neq 0$. Now assume $i \geq 2$. Then, since $w_{i-1} \in \langle v_1, \ldots, v_{i-1} \rangle$, assuming $\phi_i(w_{i-1}) \neq 0$, we have

$$\phi_{i+1}(w_i) = \phi_{i+1}(\lambda_i(w_{i-1} + \phi_i(w_{i-1})v_i)) = \lambda_i \phi_i(w_{i-1})\phi_{i+1}(v_i) \neq 0.$$

This concludes the proof. ∎

Since X' has small length over X, we may replace X by the set X' and keep calling it X. By the previous lemma, from now on we can assume the property

(P6) The (directed) diameter of $\Gamma(X')$ is at most 2 for any $X' \supseteq X$.

In what follows, we say that a directed path $[r_1, r_2, \ldots, r_k]$ in $\Gamma(\mathcal{T})$ is two-way directed, if it is also a directed path in the reverse direction, i.e. if (r_i, r_{i+1}) is two-way directed for every $1 \leq i < k$. In a similar way, we can also define the concept of two-way directed cycles, as well. A transvection graph $\Gamma(Y)$ is said to be two-way connected, if for all vertices $r, s \in Y$, there is a two-way directed path in $\Gamma(Y)$ connecting r and s. If $\Gamma(Y)$ is two-way connected, then the two-way diameter of $\Gamma(Y)$ is defined as the smallest k such that every two vertices in Y are connected in $\Gamma(Y)$ by a two-way directed path of length at most k.

Note that if G is either $\mathrm{SU}(V)$ or $\mathrm{Sp}(V)$, then the diameter of $\Gamma(X)$ coincides with its two-way diameter because in these cases every edge is a two-way directed edge. So the next claim is only interesting for $G = \mathrm{SL}(V)$.

LEMMA 4.2: *There exists a set of transvections X', containing X, such that $\ell_X(X') \leq O(1)$ with the following property. For every $s,t \in \mathcal{T}$ there are $r_1,\dots,r_k \in X'$ with $k \leq 5$ such that $s,r_1,\dots,r_k,t$ is a two-way directed path. In other words, the two-way directed diameter of $\Gamma(X'')$ is at most 6 for every $X' \subseteq X'' \subseteq \mathcal{T}$.*

Proof. First, we increase X to have the property that there is a two-way directed edge from any $s \in \mathcal{T}$ into X. Let us assume that $s \in \mathcal{T}$ is an arbitrary vertex which is not connected to any element of X by a two-way directed edge. Since $\langle X_{V^*}\rangle = V^*$, there is an $r \in X$ with $[s,r] \in E(\mathcal{T})$. By property (P6), the diameter of $X \cup \{s\}$ is at most 2, so there is a $t \in X$ such that $[r,t,s]$ is a directed path. Let us write

$$s = 1 + u \otimes \phi,\ r = 1 + v \otimes \psi \quad \text{and} \quad t = 1 + w \otimes \chi.$$

Let

$$t_s = trt^{-1} = 1 + (v + \chi(v)w) \otimes (\psi - \psi(w)\chi).$$

By our assumption $\phi(v) = \chi(u) = 0$, while $w(s,r,t) = \psi(u)\chi(v)\phi(w) \neq 0$. Thus, we have

$$\phi(v + \chi(v)w) = \phi(w)\chi(v) \neq 0 \quad \text{and} \quad (\psi - \psi(w)\chi)(u) = \psi(u) \neq 0,$$

so (s,t_s) is a two-way directed edge. Adding t_s to X for every s which is not connected to X by a two-way directed edge, the resulting X_1 will have the required property.

In view of the previous condition, we would like to increase X_1 to an X' such that the two-way distance of any $r_1,r_2 \in X_1$ in $\Gamma(X')$ is at most 4. (Note that this property only implies that the two-way diameter of X' is at most 6.) By property (P6), for any $r_1,r_2 \in X_1$ there is an $r \in X_1$ such that $[r_1,r,r_2]$ is a directed path. Therefore, it is enough to prove that we can increase X_1 to $X' \subset \mathcal{T}$ (whose length over X is small enough) such that for any one-way directed edge $[r_1,r_2] \in X_1$ there is a $t \in X'$ such that $[r_1,t,r_2]$ is a two-way directed path. After this the proof will become complete.

Let $r_1 = 1 + u_1 \otimes \phi_1$ and $r_2 = 1 + u_2 \otimes \phi_2$ be elements of X_1 such that $[r_1,r_2]$ is a one-way directed edge, i.e., $\phi_2(u_1) \neq 0,\ \phi_1(u_2) = 0$. By property (P6), there is a $t = 1 + v \otimes \psi \in X$ such that $[r_2,t,r_1]$ is a directed path of length 2. If this is a two-way path, then there is nothing to be done. Otherwise, we distinguish two cases:

CASE 1: *Exactly one of* $[r_2, t]$ *and* $[t, r_1]$ *is one-way directed.*

Let us assume that, say, $[r_2, t]$ is a one-way directed edge but (t, r_1) is a 2-cycle. Let $t_1 := r_1 t r_1^{-1} = 1 + (v + \phi_1(v)u_1) \otimes (\psi - \psi(u_1)\phi_1)$. Now,

$$(\psi - \psi(u_1)\phi_1)(u_1) = \psi(u_1) \neq 0, \qquad \phi_1(v + \phi_1(v)u_1) = \phi_1(v) \neq 0,$$
$$(\psi - \psi(u_1)\phi_1)(u_2) = \psi(u_2) \neq 0, \qquad \phi_2(v + \phi_1(v)u_1) = \phi_1(v)\phi_2(u_1) \neq 0,$$

so $[r_2, t_1, r_1]$ is two-way directed.

CASE 2: *Both of the edges in the path* $[r_2, t, r_1]$ *are one-way directed.*

Let again $t_1 = r_1 t r_1^{-1}$. The above calculation now shows that $[r_2, t_1, r_1]$ is a directed path such that (r_2, t_1) is a 2-cycle, while $[t_1, r_1]$ is a one-way directed edge. Now, the same argument as in Case 1 can be applied but to the element $t_2 = r_2 t_1 r_2^{-1}$ to get a two-way directed path $[r_2, t_2, r_1]$. ∎

Again, we may replace X by X' to ensure the following hereditary property.

(P7) The two-way diameter of $\Gamma(X')$ is at most 6 for any $X' \supseteq X$.

4.2. GENERATING THE $\mathbb{F}_{q_0}$-CLOSURE OF X. Throughout this section let $X \subseteq G$ be a set of transvections possessing all properties from (P1) to (P7). The goal of this section is to generate the $\mathbb{F}_{q_0}$-closure of X in short length over X.

Our main tool here is the concept of weight of cycles. In what follows, when we talk about a directed cycle $(r_1, r_2, \ldots, r_k)$ in $\Gamma(\mathcal{T})$, we generally mean that the indices of its vertices are elements of $\mathbb{Z}_k$, so $r_{k+1} = r_1,\ r_{k+2} = r_2$, etc.

For every integer $k \geq 2$, let $L_k = L_k(X)$ be the subfield of $\mathbb{F}_q$ generated by the weights of the cycles in $\Gamma(X)$ of length at most k. The sequence $\{L_i\}_{i=2}^{\infty}$ is an increasing sequence of subfields in $\mathbb{F}_q$.

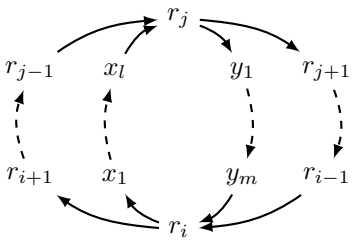
LEMMA 4.3: *Assume that the integer* $k \geq 3$ *is such that* $L_{k-1} < L_k$. *Then:*

(1) $k \in \{3, 4, 5\}$.
(2) *If* $k \in \{4, 5\}$ *and* $(r_1, \ldots, r_k)$ *is a cycle whose weight is not in* L_{k-1}, *then there is an index* i *with* $1 \leq i \leq k$ *such that* $[r_i, r_{i+2}] \notin E(X)$.

Proof. Let $(r_1, \ldots, r_k)$ be a k-cycle in $\Gamma(X)$ whose weight is not in L_{k-1}. Choosing indices i and j with $1 \leq i < j \leq k$, we claim that at least one of the following holds:

- $[r_i, r_{i+1}, \ldots, r_j]$ is a path of minimum length in $\Gamma(X)$ from r_i to r_j;
- $[r_j, r_{j+1}, \ldots, r_k, r_1, \ldots, r_i]$ is a path of minimum length in $\Gamma(X)$ from r_j to r_i.

Assume that the claim is false. Then $j \geq i+2$ and $(i,j) \neq (1,k)$. Let $x_1, \ldots, x_l,\ y_1, \ldots, y_m \in X$ such that $[r_i, x_1, \ldots, x_l, r_j]$ is a shorter path than $[r_i, r_{i+1}, \ldots, r_j]$, that is, $l < j-i-1$, and $[r_j, y_1, \ldots, y_m, r_i]$ is a shorter path than the path $[r_j, r_{j+1}, \ldots, r_k, r_1, \ldots, r_i]$, that is, $m < k-j+i-1$. So we have the following picture:



Let us consider the four directed cycles of the above picture. By assumption, the three cycles

$$C_1 := (r_i, x_1, \ldots, x_l, r_j, r_{j+1}, \ldots, r_{i-1}),$$
$$C_2 := (r_i, r_{i+1}, \ldots, r_{j-1}, r_j, y_1, \ldots, y_m),$$
$$C_3 := (r_i, x_1, \ldots, x_l, r_j, y_1, \ldots, y_m)$$

all have lengths smaller than k, so their weights are inside L_{k-1}. However,

$$w(r_1, \ldots, r_k) = \frac{w(C_1) \cdot w(C_2)}{w(C_3)},$$

so $w(r_1, \ldots, r_k) \in L_{k-1}$, a contradiction. This completes the proof of the claim.

Now, let us assume that $k \geq 6$ and let us choose $i = 1,\ j = 4$. By the previous paragraph, either $[r_1, r_2, r_3, r_4]$ is a path of shortest length between r_1 and r_4 or $[r_4, r_5, r_6, \ldots, r_1]$ is a path of shortest length between r_4 and r_1, in $\Gamma(X)$. However, both of these paths have length at least 3, while the diameter of $\Gamma(X)$ is at most 2 by (P6), a contradiction. So $k \leq 5$, as claimed.

Now, let us assume that $k \in \{4, 5\}$. If, say, $[r_1, r_3] \in E(X)$, then $[r_1, r_2, r_3]$ is not the shortest path from r_1 to r_3, so, using the first paragraph of the proof again, we get that $[r_3, \ldots, r_k, r_1]$ must be a shortest path from r_3 to r_1 in $\Gamma(X)$. Since $\operatorname{diam}(\Gamma(X)) \leq 2$ by (P6), we get that $k = 4$ and $[r_3, r_1] \notin E(X)$. Hence the last claim follows with $i = 3$. ∎

In what follows for any subset $Y \subset \mathcal{T}$ we use the notation $Y^{(k)}$ for the set of all transvections which can be written as a product of at most k many elements of $Y \cup Y^{-1}$, that is, $Y^{(k)} := \{t \in \mathcal{T} \mid \ell_Y(t) \leq k\}$.

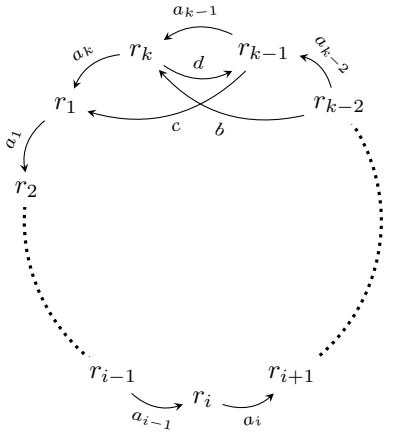
Using the previous lemma along with property (P3), we get that $L_5(X) = \mathbb{F}_q$. The following lemma allows us to assume that $L_3(X) = \mathbb{F}_q$ and $|\mathbb{F}_q : L_2(X)| \leq 2$.

LEMMA 4.4: *There exists a set of transvections X' with $\ell_X(X') = O(1)$ such that $L_3(X') = \mathbb{F}_q$ and $|\mathbb{F}_q : L_2(X')| \leq 2$.*

Proof. Let $k \geq 3$ be an integer. Let

$$r_1 = 1 + u_1 \otimes \phi_1,\ r_2 = 1 + u_2 \otimes \phi_2, \ldots,\ r_k = 1 + u_k \otimes \phi_k \in \Gamma(X)$$

be arbitrary transvections. Besides that let $a_i := \phi_{i+1}(u_i)$ for every index i with $1 \leq i \leq k$ and let $b := \phi_k(u_{k-2})$, $c := \phi_1(u_{k-1})$ and $d := \phi_{k-1}(u_k)$. Note that these are certain labels in the transvection graph. We have the following labelled graph:



(Note that only those edges appear on this picture, which have roles in the forthcoming arguments.) By Lemma 2.2,

$$\begin{aligned} r_k r_{k-1} r_k^{-1} &= 1 + (u_{k-1} + \phi_k(u_{k-1})u_k) \otimes (\phi_{k-1} - \phi_{k-1}(u_k)\phi_k) \\ &= 1 + (u_{k-1} + a_{k-1}u_k) \otimes (\phi_{k-1} - d\phi_k). \end{aligned}$$

Now, we calculate the weight of $(r_1, \ldots, r_{k-2}, r_k r_{k-1} r_k^{-1})$.

$$\begin{aligned} w(r_1, \ldots, r_{k-2}, r_k r_{k-1} r_k^{-1}) &= \prod_{i=1}^{k-3} a_i \cdot ((\phi_{k-1} - d\phi_k)(u_{k-2}))(\phi_1(u_{k-1} + a_{k-1}u_k)) \\ &= \prod_{i=1}^{k-3} a_i \cdot (a_{k-2} - db)(c + a_{k-1}a_k). \end{aligned}$$

Assume now that $(r_1, \ldots, r_k)$ is a cycle of minimal length with the property that its weight is not in $L_{k-1}(X)$. Observe that $w(r_1, \ldots, r_{k-1}) \in L_{k-1}(X)$ because, if it is nonzero, then it is the weight of a $(k-1)$-cycle. We would like to apply the above process to $(r_1, \ldots, r_k)$. By Lemma 4.3, we have $k \leq 5$. Let us assume that $k > 3$. Then by using Lemma 4.3 again, we can also assume that $b = \phi_k(u_{k-2}) = 0$. Then we have

$$\begin{aligned} w(r_1, \ldots, r_{k-2}, r_k r_{k-1} r_k^{-1}) &= \prod_{i=1}^{k-3} a_i \cdot a_{k-2}(c + a_{k-1}a_k) \\ &= w(r_1, \ldots, r_k) + w(r_1, \ldots, r_{k-1}). \end{aligned}$$

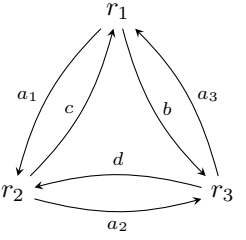
By our assumption, $w(r_1, \ldots, r_k) \notin L_{k-1}(X)$ but $w(r_1, \ldots, r_{k-1}) \in L_{k-1}(X)$, so

$$w(r_1, \ldots, r_{k-2}, r_k r_{k-1} r_k^{-1}) \notin L_{k-1}(X).$$

Since $r_k r_{k-1} r_k^{-1} \in X^{(3)}$ we get that the weights of the cycles in $\Gamma(X^{(3)})$ of length at most 4 generate $\mathbb{F}_q$. Using this argument once again, we get that $L_3(X^{(9)}) = \mathbb{F}_q$.

It remains to prove that we can extend $X^{(9)}$ to an X' with $\ell_X(X') = O(1)$ such that $|\mathbb{F}_q : L_2(X')| \leq 2$. If $L_2(X^{(27)}) = \mathbb{F}_q$, then we can choose $X' = X^{(27)}$, so let us assume that $L_2(X^{(27)}) \neq \mathbb{F}_q$ for the remainder.

Since $L_3(X^{(9)}) = \mathbb{F}_q$, we know that the weights of the 3 cycles in $\Gamma(X^{(9)})$ along with $L_2(X^{(9)})$ generate $\mathbb{F}_q$. Let (r_1, r_2, r_3) be any 3-cycle in $\Gamma(X^{(9)})$, whose weight is not inside $L_2(X^{(9)})$. Using the same notation as in the first part of the proof, we have the following diagram:



Now, $r_1, r_2, r_3, r_3 r_2 r_3^{-1} \in X^{(27)}$, so the weights

$$m_1 := w(r_1, r_2) = a_1 c, \qquad m_2 := w(r_2, r_3) = a_2 d$$
$$m_3 := w(r_3, r_1) = a_3 b, \qquad m_4 := w(r_1, r_3 r_2 r_3^{-1}) = (a_1 - db)(c + a_2 a_3)$$

are elements of $L_2(X^{(27)})$. Let $\delta = a_1a_2a_3 = w(r_1, r_2, r_3) \notin L_2(X^{(9)})$. Then we have

$$m_4 = (a_1 - db)(c + a_2a_3) = m_1 + \delta - \frac{m_1m_2m_3}{\delta} - m_2m_3,$$

so δ is a root of the polynomial

$$x^2 + (m_1 - m_4 - m_2m_3)x - m_1m_2m_3 \in L_2(X^{(27)})[x].$$

Thus, we get that the weight of any 3-cycle in $\Gamma(X^{(9)})$ is in a second degree extension of $L_2(X^{(27)})$. But such weights along with $L_2(X^{(27)})$ generate $\mathbb{F}_q$, and $\mathbb{F}_q$ contains only at most one second degree extension of $L_2(X^{(27)})$. This readily implies that $|\mathbb{F}_q : L_2(X^{(27)})| = 2$. ∎

Now, we show that if $G \neq \mathrm{SL}$, then $L_2(X) = \mathbb{F}_{q_0}$ can be achieved very easily. (Recall that $q_0 = q$ unless $G = \mathrm{SU}(n, q)$, when $q_0 = \sqrt{q}$.) In contrast, proving this in case of $G = \mathrm{SL}$ seems to be much more difficult (see Section 4.3).

LEMMA 4.5: *If $L_3(X) = \mathbb{F}_q$ and $|\mathbb{F}_q : L_2(X)| \leq 2$, then $L_2(X^{(3)}) = \mathbb{F}_q$ in the symplectic case and $L_2(X) = \mathbb{F}_{q_0}$ in the unitary case.*

Proof. By our assumption, we already know that $|L_2(X)| \geq \sqrt{q}$. Now, in the unitary case the claim follows, since $L_2(\mathcal{T}) \leq \mathbb{F}_{q_0}$ holds by a combined use of Theorem 3.12 and Remark 3.11/(2).

In the symplectic case, we take the last part of the previous proof. Let $r_i = 1 + \lambda_i u_i \otimes \varphi_{u_i} \in X$ for $i = 1, 2, 3$. In this case we have

$$\begin{aligned} dbc &= \varphi_{u_2}(\lambda_3u_3)\varphi_{u_3}(\lambda_1u_1)\varphi_{u_1}(\lambda_2u_2) \\ &= \lambda_1\lambda_2\lambda_3(-\varphi_{u_3}(u_2))(-\varphi_{u_1}(u_3))(-\varphi_{u_2}(u_1)) = -a_1a_2a_3 \end{aligned}$$

hence $(r_1, r_3r_2r_3^{-1})$ is a 2-cycle in $X^{(3)}$ whose weight is

$$\begin{aligned} w(r_1, r_3r_2r_3^{-1}) &= (a_1 - db)(c + a_2a_3) = a_1a_2a_3 - dbc + a_1c - dba_2a_3 \\ &= 2w(r_1, r_2, r_3) + w(r_1, r_2) - w(r_1r_3)w(r_2r_3). \end{aligned}$$

Since the characteristic of $\mathbb{F}_q$ is different from 2, we have that $w(r_1, r_2, r_3)$ is an element of the subfield generated by the 2-cycles of $\Gamma(X^{(3)})$ for any 3-cycle (r_1, r_2, r_3) of $\Gamma(X)$. Thus, the weights of the 2-cycles of $\Gamma(X^{(3)})$ generate $\mathbb{F}_q$. ∎

Note that the above proof already used that q is odd, when G is a symplectic group. Our next proof heavily relies on Theorem 3.7, so it uses our full assumption on q.

LEMMA 4.6: *Let $L := L_2(X) \leq \mathbb{F}_{q_0}$ and let us assume that $L \neq \mathbb{F}_9$. Then we have the following:*

(1) $\ell_X(X^L) \leq O((\log |L|)^c)$.
(2) *If X contains a transvection subgroup $s_0^{L_0}$ over L_0 for some subfield $L_0 \leq L$, then $\ell_X(X^L) \leq O(|L : L_0|^c)$.*

Proof. First, we show that if we can generate a transvection subgroup s^L for some $s \in X$, then we can generate X^L in length $\ell_{X\cup s^L}(X^L) = O(1)$. Indeed, let us assume that such an s^L is already generated and let (s, r) be a two-way directed edge in $\Gamma(X)$. Using Theorem 3.7, we know that $\langle s^L, r\rangle \simeq \mathrm{SL}(2, L) \oplus 1_{n-2}$. In particular, $r^L \leq \langle s^L, r\rangle$. By using Corollary 1.3, we get the following for some constant c_0.

$$\ell_{X\cup s^L}(r^L) \leq \ell_{\{s^L,r\}}(\langle s^L, r\rangle) \leq \ell_{\{s^L,r\}}(\mathrm{SL}(2, L)) = O\Big(\frac{\log |\mathrm{SL}(2, L)|}{\log |L|}\Big)^{c_0} = O(1).$$

Now, for an arbitrary $t \in X$ let $s, r_1, r_2, \ldots, r_k, t$ be a two-way directed path in $\Gamma(X)$ with $k \leq 5$. (such a path exists by property (P7).) Using the above argument repeatedly to the two-way edges $(s, r_1),\ (r_1, r_2),\ \ldots, (r_k, t)$, we get that $\ell_{X\cup s^L}(t^L) = O(1)$, as claimed.

Now, we turn to the problem of generating a transvection subgroup over L. If $|L| = 3$, the statement is trivial, so for the remainder we can assume that $|L| \neq 3, 9$. Then there is a 2-cycle (s_1, t_1) in $\Gamma(X)$ such that $\mathbb{F}_p(w(s_1, t_1)) \not\leq \mathbb{F}_9$. Let $K_1 := \mathbb{F}_p(w(s_1, t_1)) \leq L$. Using Theorem 3.7 along with Corollary 1.3, we get that

$$\begin{aligned}\ell_X(s_1^{K_1}) &\leq \ell_{\{s_1,t_1\}}(s_1^{K_1}) \leq \ell_{\{s_1,t_1\}}(\mathrm{SL}(2, K_1)) \\ &= O((\log_2 |\mathrm{SL}(2, K_1)|)^{c_0}) = O((\log_2 |K_1|)^{c_0})\end{aligned}$$

for some constant c_0. Now, if $K_1 = L$, then we are done.

Now assume that $K_1 < L$ and let (s_2, t_2) be a two-way directed edge in $\Gamma(X)$ with $w(s_2, t_2) \notin K_1$ such that the distance of s_1 and s_2 is the smallest possible, and let $K_2 := K_1(w(s_2, t_2))$. Using the same procedure as in the first paragraph to a shortest two-way path $s_1, r_1, \ldots, r_k, s_2$, we can generate $s_2^{K_1}$ in length $\ell_{X\cup s_1^{K_1}}(s_2^{K_1}) = O(1)$. (Note that $w(s_1, r_1),\ w(r_1, r_2),\ \ldots, w(r_k, s_2)$ are all elements of K_1 by our assumption.) Then

$$\langle s_2^{K_1}, t_2\rangle = \mathrm{SL}(2, K_2) \oplus 1_{n-2} \geq s_2^{K_2}$$

by Theorem 3.7 so, by using Corollary 1.3, we get that

$$\begin{aligned}\ell_{X\cup s_1^{K_1}}(s_2^{K_2}) &\le O(1)\cdot \ell_{\{s_2^{K_1},t_2\}}(\langle s_2^{K_2},t_2\rangle)\\ &\le O\Big(\frac{\log|\mathrm{SL}(2,K_2)|}{\log|K_1|}\Big)^{c_0} = O(|K_2:K_1|^{c_0}).\end{aligned}$$

In general, if $K_i < L_2(X)$ and the K_i-closure of some $s_i \in X$ is already generated, then we choose (s_{i+1},t_{i+1}) in $\Gamma(X)$ with $K_{i+1} := K_i(w(s_{i+1},t_{i+1})) > K_i$ and we apply the above procedure to get $s_{i+1}^{K_{i+1}}$ in length

$$\ell_{X\cup s_i^{K_i}}(s_{i+1}^{K_{i+1}}) = O(|K_{i+1}:K_i|^{c_0}).$$

Finally, we get a strictly increasing chain of subfields $K_1 < K_2 < \cdots < K_m = L$ with $m \le \log\log|L|$ and a 2-cycle (s_i,t_i) in $\Gamma(X)$ with $K_{i+1} = K_i(w(s_{i+1},t_{i+1}))$ for every $i < m$. Using the above procedure we can generate s_m^L in length

$$\begin{aligned}\ell_X(s_m^L) &= O((\log|K_1|)^{c_0})\prod_{i=1}^{m-1} O(|K_{i+1}:K_i|^{c_0})\\ &= O(\log|L|^{c_0})O(1)^{\log\log|L|} \le O((\log|L|)^c).\end{aligned}$$

So the proof of the first claim is complete.

Finally, the second claim follows by using essentially the same argument, but we do not use the second paragraph. Instead, we choose $s_1 := s_0$ and $K_1 := L_0$. ∎

Now, we are in the position to generate the $\mathbb{F}_{q_0}$-closure of X in short length over X in many cases.

COROLLARY 4.7: *Let $M = \mathbb{F}_{q_0}$ unless $G = \mathrm{SL}(n,q)$ with q perfect square, in which case let $M = \mathbb{F}_{\sqrt{q}}$. Then $\ell_X(X^M) = O((\log q)^c)$.*

Proof. By Lemmas 4.4 and 4.5, we can assume that $L_2(X) = \mathbb{F}_{q_0}$ unless $G = \mathrm{SL}(n,q)$ with q a perfect square when we can assume that $L_2(X) \ge \mathbb{F}_{\sqrt{q}}$ and $L_3(X) = \mathbb{F}_q$. An application of Lemma 4.6/(1) gives the result. ∎

Replacing X with X^M we can assume that

(P8) (a) If $G = \mathrm{Sp}(n,q)$ or $G = \mathrm{SU}(n,q)$ or q is not a perfect square, then X is $\mathbb{F}_{q_0}$-closed.

(b) If $G = \mathrm{SL}(n,q)$ and q is a perfect square, then X is M-closed with $M = \mathbb{F}_{\sqrt{q}}$ and $L_3(X) = \mathbb{F}_q$.

4.3. Gluing Triangles. The ultimate goal of this section is to finish the proof of Theorem 1.5 for the case when $G = \mathrm{SL}(V)$. In view of the previous sections, we can assume that X satisfies properties (P1)–(P7) and (P8)/(b). Therefore, X is M-closed for $M := \mathbb{F}_{\sqrt{q}}$ and X contains a triangle whose weight is not in M. By our assumptions on q, we have $|M| \geq 5$.

The main difficulty we face now is to construct a 2-cycle (s,t) in short length over X such that $w(s,t) \notin M$. In fact, we will be able to generate such a 2-cycle in length $\ell_X(s,t) = O(1)$. To prove this, we will use the concept of non-unitary cycles (see Definition 3.10).

We highlight that in the foregoing discussion, a large part of our calculations remain valid only until the weight of each 2-cycle of the examined parts of the transvection graph are inside M, that is, until each such 2-cycle is unitary (see Remark 3.11/(2)). Since our primary goal is to construct a non-unitary 2-cycle, it does not cause any problem, if we tacitly assume that during our proof all intermediate two-way directed edges are unitary 2-cycles.

By Theorem 3.12, $\Gamma(X)$ must contain a non-unitary cycle. Previously (see Lemmas 4.3 and 4.4) we have seen how to construct a triangle of bounded length over X whose weight is outside M. Using a similar argument, we can also construct a non-unitary triangle of bounded length over X.

Lemma 4.8: *We can assume that X contains a non-unitary triangle.*

Proof. First, if X contains a one-way directed edge $[r,s] \in E(X)$, then by property (P6) there is a $t \in X$ such that (r,s,t) is a one-way directed cycle, which is non-unitary by Remark 3.11/(4). Therefore, for the remainder of this proof we can assume that every edge in $\Gamma(X)$ is two-way directed.

Let $(r_1, r_2, \ldots, r_k)$ be a non-unitary cycle of minimal length in $\Gamma(X)$. (Such a cycle exists by property (P5).) By our assumption, every 2-cycle is unitary in $\Gamma(X)$, so we can use the gluing property of P_u (see Eq. (1)), to conclude that $(r_1, r_2, \ldots, r_k)$ cannot be obtained by gluing two shorter (and, therefore, unitary) cycles of $\Gamma(X)$.

Therefore, we get that $(r_1, r_2, \ldots, r_k)$ must be a chordless cycle of length $k \leq 5$, where "chordless" means that $[r_i, r_j]$ is an edge if and only if $i - j \equiv \pm 1 \pmod{k}$. If $k > 3$, then let $s := r_k r_{k-1} r_k^{-1}$. Now, $(r_1, r_2, \ldots, r_{k-2}, s)$ is a (two-way) directed cycle of length $k-1$. Furthermore, s is connected by both r_k and r_{k-1}

with a two-way edge. Using the gluing property we have

$$P_u(r_1,\dots,r_k) = P_u(r_1,r_2,\dots,r_{k-2},s)\cdot P_u(r_1,s,r_k) \\ \times P_u(r_k,s,r_{k-1})\cdot P_u(r_{k-1},s,r_{k-2}).$$

Thus, we get a non-unitary cycle of length at most $k-1 \leq 4$. If we still do not have a non-unitary triangle, we use the above argument again for a non-unitary 4-cycle. So, we conclude that $X^{(9)}$ surely contains a non-unitary triangle. ∎

For the remainder of this section, if s_1, s_2, s_3, and so on, are transvections, we use the usual notation $s_i = 1 + u_i \otimes \phi_i$. Recall that, by Lemma 2.2,

$$s_i^{s_j} = s_j^{-1} s_i s_j = 1 + (u_i - \phi_j(u_i)u_j) \otimes (\phi_i + \phi_i(u_j)\phi_j).$$

LEMMA 4.9: *For any $g \in G$, the conjugation by g defines a graph isomorphism on $\Gamma(\mathcal{T})$. Moreover,*

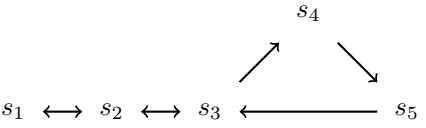
$$w(r_1,r_2,\dots,r_k) = w(r_1^g, r_2^g, \dots, r_k^g)$$

for any $r_1, r_2, \dots, r_k \in \mathcal{T}$.

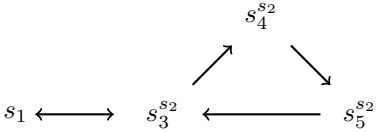
Proof. Both claims can be easily proved using the fact that if $r = 1 + u \otimes \phi$ is a transvection, then

$$r^g = g^{-1} r g = 1 + g^{-1}(u) \otimes (\phi \circ g). \qquad ∎$$

LEMMA 4.10: *Let s_1, s_2, s_3, s_4, s_5 be transvections such that the following happens.*



Assume that $[s_1, s_3]$ is not a double edge. If $[s_1, s_3]$ is a single edge, then the triangle (s_1, s_3, s_2) is non-unitary. If $[s_3, s_1]$ is a single edge then the triangle (s_1, s_2, s_3) is non-unitary. Assume now that $[s_1, s_3]$ and $[s_3, s_1]$ are not edges. Then the following happens:



Proof. If one of $[s_1, s_3]$ and $[s_3, s_1]$ is a single edge then the result follows from Remark 3.11/(4). We have $s_3^{s_2} = 1 + (u_3 - \phi_2(u_3)u_2) \otimes (\phi_3 + \phi_3(u_2)\phi_2)$, hence

$$(\phi_3 + \phi_3(u_2)\phi_2)(u_1) = \phi_3(u_2)\phi_2(u_1) \neq 0,$$
$$\phi_1(u_3 - \phi_2(u_3)u_2) = -\phi_2(u_3)\phi_1(u_2) \neq 0.$$

Therefore $(s_1, s_3^{s_2})$ is a double edge. By Lemma 4.9, we are done. ∎

LEMMA 4.11: *If s_1, s_2, s_3 are transvections then*

$$w(s_1, s_2, s_3) - w(s_1, s_3, s_2) = w(s_1, s_3) - w(s_1, s_2)w(s_2, s_3) - w(s_1^{s_2}, s_3).$$

Proof. Note that $s_1^{s_2} = 1 + (u_1 - \phi_2(u_1)u_2) \otimes (\phi_1 + \phi_1(u_2)\phi_2)$, so

$$\begin{aligned} w(s_1^{s_2}, s_3) &= (\phi_3(u_1) - \phi_2(u_1)\phi_3(u_2)) \cdot (\phi_1(u_3) + \phi_1(u_2)\phi_2(u_3)) \\ &= w(s_1, s_3) + w(s_1, s_3, s_2) - w(s_1, s_2, s_3) - w(s_1, s_2)w(s_2, s_3). \end{aligned}$$

The result follows. ∎

Remark 4.12: Lemma 4.11 implies that whenever we have a triangle (s_1, s_2, s_3) such that $w(s_1, s_2, s_3) \notin M$ but $w(s_1, s_3, s_2) \in M$, then there are $s, t \in \mathcal{T}$ with $w(s, t) \notin M$ and $\ell_X(s, t) \leq 3$, which is our main goal. So, for the rest of this section we can assume that whenever $w(s_1, s_2, s_3) \notin M$ for some already generated $s_1, s_2, s_3 \in \mathcal{T}$, then $w(s_1, s_3, s_2)$ is also not in M.

The following lemma will be used many times in this section without referring to it.

LEMMA 4.13: *Let $s_1, s_2, s_3, s_4 \in X$, $\lambda \in M$ and let*

$$A := w(s_1, s_2, s_3),$$
$$B := w(s_1, s_2, s_4, s_3) - w(s_1, s_2, s_3, s_4),$$
$$C := -w(s_3, s_4) \cdot w(s_1, s_2, s_4),$$
$$D := d_u(s_1, s_2, s_3),$$
$$E := w(s_1, s_2, s_4, s_3) - w(s_1, s_2, s_3, s_4) + w(s_2, s_1, s_4, s_3)^{\sqrt{q}} - w(s_2, s_1, s_3, s_4)^{\sqrt{q}},$$
$$F := -w(s_3, s_4) \cdot d_u(s_1, s_2, s_4).$$

Then

$$w(s_1, s_2, s_3^{s_4^\lambda}) = A + \lambda B + \lambda^2 C,$$
$$d_u(s_1, s_2, s_3^{s_4^\lambda}) = D + \lambda E + \lambda^2 F,$$

and we have the following:

(1) *If either $A \in M$, $C \notin M$ or $A \notin M$, $C \in M$ then for every $\lambda \in M^\times$ at least one of $w(s_1, s_2, s_3^{s_4^\lambda})$, $w(s_1, s_2, s_3^{s_4^{-\lambda}})$ does not belong to M.*

(2) *If at least one of D, E, F is nonzero, then there are at most two values of λ for which $d_u(s_1, s_2, s_3^{s_4^\lambda}) = 0$.*

Proof. Observe that $s_3^{s_4^\lambda} = 1 + (u_3 - \lambda\phi_4(u_3)u_4) \otimes (\phi_3 + \lambda\phi_3(u_4)\phi_4)$. Therefore

$$\begin{aligned} w(s_1,&s_2, s_3^{s_4^\lambda}) \\ &= \phi_2(u_1) \cdot (\phi_3(u_2) + \lambda\phi_3(u_4)\phi_4(u_2)) \cdot (\phi_1(u_3) - \lambda\phi_4(u_3)\phi_1(u_4)) \\ &= \phi_2(u_1)\phi_3(u_2)\phi_1(u_3) - \lambda\phi_2(u_1)\phi_3(u_2)\phi_4(u_3)\phi_1(u_4) \\ &\quad + \lambda\phi_2(u_1)\phi_3(u_4)\phi_4(u_2)\phi_1(u_3) - \lambda^2\phi_2(u_1)\phi_3(u_4)\phi_4(u_2)\phi_4(u_3)\phi_1(u_4) \\ &= A + \lambda B + \lambda^2 C. \end{aligned}$$

In particular,

$$w(s_1, s_2, s_3^{s_4^\lambda}) + w(s_1, s_2, s_3^{s_4^{-\lambda}}) = 2A + 2\lambda^2 C.$$

Using the permutation (1 2) to the indices in the formulas given for A, B, C, an easy calculation shows that

$$d_u(s_1, s_2, s_3^{s_4^\lambda}) = w(s_1, s_2, s_3^{s_4^\lambda}) + w(s_2, s_1, s_3^{s_4^\lambda})^{\sqrt{q}} = D + \lambda E + \lambda^2 F.$$

The remaining claims of the lemma are clear. ∎

LEMMA 4.14: *Assume that we are in the case $G = \mathrm{SL}(V)$ and that $|M| = \sqrt{q}$. Assume at least one of the following holds:*

(1) *There exist $s_1, s_2, s_3 \in X$ with $w(s_1, s_2, s_3) \notin M$ and $d_u(s_1, s_2, s_3) \neq 0$.*

(2) *There exist $s_1, s_2, s_3, s_4 \in X$ with $w(s_1, s_2, s_3) \notin M$ and $d_u(s_4, s_3, s_2) \neq 0$.*

Then there exist $t_1, t_2 \in \mathcal{T}$ such that $w(t_1, t_2) \notin M$ and $\ell_X(\{t_1, t_2\}) = O(1)$.

Proof. Assume case (1) holds. We have a triangle (s_1, s_2, s_3) with the property that $w(s_1, s_2, s_3) \notin M$ and $d_u(s_1, s_2, s_3) \neq 0$. Note that we can assume that $w(s_1, s_2, s_3) - w(s_1, s_3, s_2) \in M$ by Lemma 4.11, so

$$w(s_1, s_2, s_3) + w(s_1, s_3, s_2) \neq 0.$$

Indeed, if $w(s_1, s_2, s_3) + w(s_1, s_3, s_2) = 0$ then

$$2w(s_1, s_2, s_3) = w(s_1, s_2, s_3) - w(s_1, s_3, s_2) \in M$$

and the characteristic being odd, this would imply $w(s_1, s_2, s_3) \in M$, a contradiction.

Define $W := \langle u_1, u_2, u_3 \rangle_{\mathbb{F}_q}$. The fact that $w(s_1, s_2, s_3) + w(s_1, s_3, s_2) \neq 0$ implies that $\dim_{\mathbb{F}_q}(W) = 3$ (in other words, u_1, u_2 and u_3 are linearly independent over $\mathbb{F}_q$) and that we have a direct sum decomposition

$$V = \langle u_1, u_2, u_3 \rangle \oplus (\ker(\phi_1) \cap \ker(\phi_2) \cap \ker(\phi_3)).$$

This is because if a linear combination $\mu_1 u_1 + \mu_2 u_2 + \mu_3 u_3$ (where $\mu_1, \mu_2, \mu_3 \in \mathbb{F}_q$) belongs to $\ker(\phi_1) \cap \ker(\phi_2) \cap \ker(\phi_3)$ (in particular, this happens if this linear combination is zero) then applying ϕ_1, ϕ_2 and ϕ_3 to such a linear combination we obtain a homogeneous linear system in μ_1, μ_2, μ_3 whose determinant is precisely $w(s_1, s_2, s_3) + w(s_1, s_3, s_2)$. Moreover, a very similar argument shows that ϕ_1, ϕ_2 and ϕ_3 are linearly independent, and therefore span the dual space W^*.

Since (s_1, s_2, s_3) is a 3-cycle, the transvection graph $\Gamma(\{s_1, s_2, s_3\})$ is strongly connected, therefore the group

$$H = \langle s_1^M, s_2^M, s_3^M \rangle \leq \mathrm{GL}(V)$$

is identified (via the direct sum decomposition above) with an irreducible subgroup of $\mathrm{GL}(W)$ (by Theorem 3.1). By Theorem 3.8 and by Section 3.2, H is a special linear, unitary or symplectic group of dimension 3 over

$$M(w(s_1, s_2, s_3)) = \mathbb{F}_q.$$

But since the 3-cycle (s_1, s_2, s_3) is non-symplectic and non-unitary, the group H cannot be symplectic nor unitary, so $H \cong \mathrm{SL}(3, q)$ by Theorem 3.12.

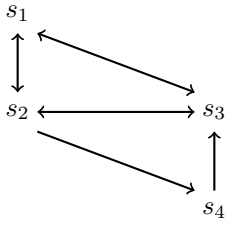
Let $t_1, t_2 \in \mathcal{T}(H) \subset \mathcal{T}$ with $w(t_1, t_2) \notin M$. By Corollary 1.3,

$$\ell_X(t_1, t_2) \leq \ell_{\{s_1^M, s_2^M, s_3^M\}}(H) = O\Big(\frac{\log |\mathrm{SL}(3, q)|}{\log(3|M|)}\Big)^c = O(1),$$

as claimed.

Assume case (2) holds. Now, we can assume that $w(s_1, s_3, s_2) \notin M$ by Remark 4.12. By applying a permutation to the indices if necessary, we can assume that (s_1, s_2, s_3) and (s_2, s_4, s_3) are two triangles, with $w(s_1, s_2, s_3) \notin M$ and $d_u(s_2, s_4, s_3) \neq 0$. We can also assume that $d_u(s_1, s_2, s_3) = 0$, furthermore $w(s_2, s_3, s_4) \in M$ and $w(s_2, s_4, s_3) \in M^\times$ by case (1). We call a triangle "good" if it is non-unitary and its weight is not in M. If we find a good triangle we are done by case (1), so what we will do is look for a good triangle. By Remark 3.11/(4), (s_1, s_2), (s_2, s_3) and (s_3, s_1) are double edges. So we have the

following picture.



In each of the following cases we use Lemma 4.13 for a well-chosen triangle.

CASE 1: $[s_4, s_1] \notin E(\mathcal{T})$ *or* $w(s_2, s_4, s_1) \notin M$.

We have

$$w(s_2, s_4, s_1^{s_3^\lambda}) = A + B\lambda + C\lambda^2$$

with

$$A = w(s_2, s_4, s_1), \quad B = w(s_2, s_4, s_3, s_1) - w(s_2, s_4, s_1, s_3) \quad \text{and}$$
$$C = -w(s_1, s_3) \cdot w(s_2, s_4, s_3) \in M^\times.$$

If $[s_4, s_1] \notin E(\mathcal{T})$, then $A = 0$ and

$$B = w(s_4, s_3, s_1, s_2) = \frac{w(s_1, s_2)w(s_3, s_1)w(s_2, s_4, s_3)}{w(s_1, s_3, s_2)} \notin M,$$

so $w(s_2, s_4, s_1^{s_3^\lambda}) \notin M$ for every $\lambda \in M^\times$.

On the other hand, if $A \notin M$, then since $C \in M$, $w(s_3, s_1) \cdot d_u(s_2, s_4, s_3) \neq 0$ and $|M| \geq 5$, there exists $\lambda \in M$ such that $(s_2, s_4, s_1^{s_3^\lambda})$ is a good triangle.

CASE 2: $[s_4, s_1] \in E(\mathcal{T})$ *and* ($[s_1, s_4] \in E(\mathcal{T})$ *or* $[s_3, s_4] \notin E(\mathcal{T})$).

We have $w(s_3, s_2, s_1^{s_4^\lambda}) = A + B\lambda + C\lambda^2$ with $A = w(s_3, s_2, s_1) \notin M$ and $C = -w(s_1, s_4) \cdot w(s_3, s_2, s_4) \in M$. In particular, at least one of $w(s_3, s_2, s_1^{s_4^\lambda})$ and $w(s_3, s_2, s_1^{s_4^{-\lambda}})$ is not in M for any $\lambda \in M$.

Now, we have $d_u(s_3, s_2, s_1^{s_4^\lambda}) = \lambda E - \lambda^2 F$ for some $E, F \in \mathbb{F}_q$. If $[s_1, s_4] \in E(\mathcal{T})$, then

$$F = w(s_1, s_4) \cdot d_u(s_2, s_4, s_3) \neq 0,$$

so there are at most two values $\lambda \in M$ such that $d_u(s_3, s_2, s_1^{s_4^\lambda}) = 0$. On the other hand, if $[s_1, s_4] \notin E(\mathcal{T})$ and $[s_3, s_4] \notin E(\mathcal{T})$, then $d_u(s_3, s_2, s_1^{s_4^\lambda}) = \lambda E$,

where

$$E = w(s_3, s_2, s_4, s_1) - w(s_3, s_2, s_1, s_4) + w(s_2, s_3, s_4, s_1)^{\sqrt{q}} - w(s_2, s_3, s_1, s_4)^{\sqrt{q}}$$
$$= w(s_3, s_2, s_4, s_1) \neq 0,$$

so $d_u(s_3, s_2, s_1^{s_4^\lambda}) \neq 0$ for every $\lambda \in M^\times$. Since $|M| \geq 5$, there is a $\lambda \in M$ such that $(s_3, s_2, s_1^{s_4^\lambda})$ is a good triangle.

Case 3: $[s_4, s_1]$ *is a one-way directed edge,* (s_4, s_3) *is a two-way directed edge,* $w(s_1, s_2, s_4) \in M^\times$.

Since

$$A = w(s_1, s_2, s_3) \notin M, \quad C = -w(s_3, s_4) \cdot w(s_1, s_2, s_4) \in M,$$
$$F = -w(s_3, s_4) \cdot d_u(s_1, s_2, s_4) \neq 0 \quad \text{and } |M| \geq 5,$$

we can find $\lambda \in M$ such that the triangle $(s_1, s_2, s_3^{s_4^\lambda})$ is good. ∎

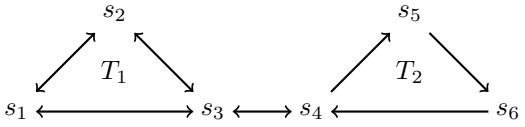
From now on, if a triangle is unitary and its weight is not in M then we call it of "type 1" and if a triangle is non-unitary and its weight is in M then we call it of "type 2". Note that the edges of a triangle of type 1 are two-way directed by Remark 3.11/(4).

Since $\Gamma(X)$ contains a non-unitary triangle by Lemma 4.8 and $L_3(X) = \mathbb{F}_q$ by (P8)(b), using Lemma 4.14/(1) we may assume that in $\Gamma(X)$ there are triangles of both types. What we want is to find one of the two situations described in Lemma 4.14 in the transvection graph spanned by a power of X whose exponent is bounded by a constant. So in the following discussion we may assume (by way of contradiction) that in every power of X whose exponent is bounded by a constant, all non-unitary triangles are of type 2 and all triangles whose weight is not in M are of type 1. Under this assumption, our goal is to generate four transvections satisfying Lemma 4.14/(2) whose length over X is bounded.

Assume we have two triangles $T_1 = (s_1, s_2, s_3)$ and $T_2 = (s_4, s_5, s_6)$, with T_1 of type 1 and T_2 of type 2. By Remark 3.11/(4), T_1 consists of double edges. First, we want to reduce the problem to the case in which T_1 and T_2 share a vertex.

Let us assume that T_1 and T_2 do not have a vertex in common. Let, say, s_3 and s_4 be chosen such that their two-way distance in $\Gamma(X)$ is the smallest possible one, and let $s_3, r_1, \ldots, r_k, s_4$ be a shortest two-way path in $\Gamma(X)$, so $k \leq 5$ by property (P7). If $k \geq 1$, then we may apply Lemma 4.9 and Lemma 4.10 to

change T_2 to a triangle of type 2, whose distance from T_1 is shorter. Using this process at most 5 times, we may assume that we are in the following situation.



If $d_u(s_5, s_6, s_3) \neq 0$, then we may replace T_2 with (s_5, s_6, s_3), while if $w(s_1, s_2, s_4) \notin M$ then we may replace T_1 with (s_1, s_2, s_4) and reduce to the case in which T_1 and T_2 share a vertex. Now, let us assume that $d_u(s_5, s_6, s_3) = 0$ and $w(s_1, s_2, s_4) \in M$. Then

$$d_u(s_5, s_6, s_3^{s_4^\lambda}) = \lambda E + \lambda^2 F$$

with $F = w(s_3, s_4) \cdot d_u(s_5, s_6, s_4) \neq 0$, so there is at most one $\lambda \in M^\times$ such that $d_u(s_5, s_6, s_3^{s_4^\lambda}) = 0$. Since $|M| \geq 5$, we can choose $\lambda \in M^\times$ such that both

$$d_u(s_5, s_6, s_3^{s_4^\lambda}) \neq 0 \quad \text{and} \quad d_u(s_5, s_6, s_3^{s_4^{-\lambda}}) \neq 0.$$

Since $w(s_1, s_2, s_3) \notin M$ but $w(s_1, s_2, s_4) \in M$, at least one of $w(s_1, s_2, s_3^{s_4^\lambda})$, $w(s_1, s_2, s_3^{s_4^{-\lambda}})$ is not in M. Up to replacing λ with $-\lambda$, we may assume that $w(s_1, s_2, s_3^{s_4^\lambda}) \notin M$. By our choice of λ, $d_u(s_3^{s_4^\lambda}, s_5, s_6) \neq 0$, so at least one of $(s_3^{s_4^\lambda}, s_5, s_6)$, $(s_3^{s_4^\lambda}, s_6, s_5)$ is a triangle. If, say, $(s_3^{s_4^\lambda}, s_5, s_6)$ is a triangle, then it is a triangle of type 2 by our assumption.

So we are in the situation where T_1 and T_2 share a vertex as it can be seen on Figure 1.

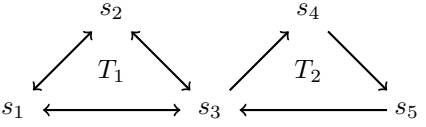
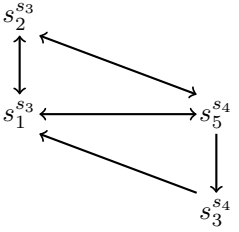


Figure 1. Triangles of type 1 and 2.
Type 1: unitary, weight not in M.
Type 2: non-unitary, weight in M.

We consider several cases, which are distinguished by the directed edges appearing on these two triangles.

CASE 1: $[s_2, s_5]$, $[s_3, s_5]$, $[s_2, s_4]$, $[s_5, s_1]$, $[s_4, s_1]$, $[s_4, s_3]$ *are not edges in Figure 1.*

We would like to prove that $w(s_1^{s_3}, s_2^{s_3}, s_5^{s_4}) \notin M$ and that $d_u(s_1^{s_3}, s_5^{s_4}, s_3^{s_4}) \neq 0$.



Note that

$$\begin{aligned}
s_1^{s_3} &= 1 + (u_1 - \phi_3(u_1)u_3) \otimes (\phi_1 + \phi_1(u_3)\phi_3),\\
s_2^{s_3} &= 1 + (u_2 - \phi_3(u_2)u_3) \otimes (\phi_2 + \phi_2(u_3)\phi_3),\\
s_3^{s_4} &= 1 + (u_3 - \phi_4(u_3)u_4) \otimes (\phi_3 + \phi_3(u_4)\phi_4),\\
s_5^{s_4} &= 1 + (u_5 - \phi_4(u_5)u_4) \otimes (\phi_5 + \phi_5(u_4)\phi_4).
\end{aligned}$$

Since

$$\phi_5(u_2) = \phi_5(u_3) = \phi_4(u_2) = \phi_1(u_5) = \phi_1(u_4) = \phi_3(u_4) = 0,$$

we have

$$\begin{aligned}
w(s_1^{s_3}, s_2^{s_3}, s_5^{s_4}) &= \phi_2(u_1)(-\phi_5(u_4)\phi_3(u_2)\phi_4(u_3)) \cdot (\phi_1(u_3)\phi_3(u_5))\\
&= -w(s_1, s_2, s_3) \cdot w(s_3, s_4, s_5) \notin M,
\end{aligned}$$

being $w(s_1, s_2, s_3) \notin M$ and $0 \neq w(s_3, s_4, s_5) \in M$. So we may assume that $(s_1^{s_3}, s_2^{s_3}, s_5^{s_4})$ is a triangle of type 1. This implies that it is a two-way directed cycle, so $[s_1^{s_3}, s_5^{s_4}] \in E(\mathcal{T})$. Moreover, by Lemma 4.9, $[s_5^{s_4}, s_3^{s_4}]$ is a single edge, being a conjugate of the single edge $[s_5, s_3]$, and $[s_3^{s_4}, s_1^{s_3}] \in E(\mathcal{T})$ since

$$(\phi_1 + \phi_1(u_3)\phi_3)(u_3 - \phi_4(u_3)u_4) = \phi_1(u_3) \neq 0$$

being

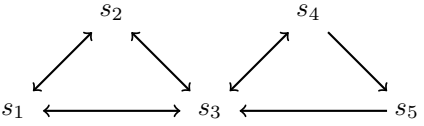
$$\phi_1(u_4) = \phi_3(u_4) = 0.$$

Thus, $(s_1^{s_3}, s_5^{s_4}, s_3^{s_4})$ is a one-way directed triangle, so

$$d_u(s_1^{s_3}, s_5^{s_4}, s_3^{s_4}) \neq 0$$

and we are done.

CASE 2: $[s_4, s_3] \in E(\mathcal{T})$.



Let $\lambda \in M$ (to be specified).

Assume first that $w(s_2, s_1, s_4) \neq 0$. We may assume that both triangles (s_1, s_2, s_3) and (s_1, s_3, s_2) are of type 1, by Remark 4.12. The triangle (s_2, s_1, s_4) is of type 2 (in which case we are done) unless (s_1, s_4) and (s_4, s_2) are double edges (by Remark 3.11/(4)), so we may assume this is the case.

Since $w(s_2, s_4) \cdot d_u(s_5, s_3, s_4) \neq 0$, there are at most two values of $\lambda \in M$ such that $d_u(s_5, s_3, s_2^{s_4^\lambda}) = 0$. If $w(s_3, s_1, s_4) \notin M$ then (s_3, s_1, s_4) is a triangle of type 1 and we are done, so now assume that $w(s_3, s_1, s_4) \in M$. Since $w(s_3, s_1, s_2) \notin M$, at least one of $w(s_3, s_1, s_2^{s_4^\lambda}) \notin M$ or $w(s_3, s_1, s_2^{s_4^{-\lambda}}) \notin M$ for any $\lambda \in M$. Since $|M| \geq 5$, there exists $\lambda \in M$ such that $d_u(s_5, s_3, s_2^{s_4^\lambda}) \neq 0$ and $w(s_3, s_1, s_2^{s_4^\lambda}) \notin M$. Thus, we are done in this case.

Assume now that $w(s_2, s_1, s_4) = 0$. Since $w(s_2, s_3) \cdot d_u(s_3, s_4, s_5) \neq 0$, there are at most two values of $\lambda \in M$ such that

$$d_u(s_2^{s_3^{-\lambda}}, s_4, s_5) = d_u(s_2, s_4^{s_3^\lambda}, s_5^{s_3^\lambda}) = 0.$$

Since $w(s_1, s_2, s_3) \notin M$, by Remark 4.12 we may assume that $w(s_2, s_1, s_3) \notin M$, so that $w(s_3, s_4) \cdot w(s_2, s_1, s_3) \notin M$ being $w(s_3, s_4) \in M^\times$. We would like to apply Lemma 4.13 in order to obtain that $w(s_2, s_1, s_4^{s_3^\lambda}) \notin M$, while at the same time $d_u(s_2^{s_3^{-\lambda}}, s_4, s_5) \neq 0$. It was proved in Lemma 4.13 that $w(s_2, s_1, s_4^{s_3^\lambda})$ can be expressed as $A + B\lambda + C\lambda^2$. In our case we have

$$A = w(s_2, s_1, s_4) = 0.$$

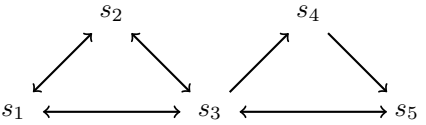
If $\alpha, \beta \in M^\times$, $\alpha \neq \beta$ satisfy $B\alpha + C\alpha^2 \in M$ and $B\beta + C\beta^2 \in M$, then Cramer's rule implies $B, C \in M$, which is not the case since

$$C = -w(s_3, s_4) \cdot w(s_2, s_1, s_3) \notin M.$$

Thus $w(s_2, s_1, s_4^{s_3^\lambda})$ is in M for at most one value of $\lambda \in M^\times$.

Since $|M| \geq 5$, we obtain that there is a $\lambda \in M$ such that $w(s_2, s_1, s_4^{s_3^\lambda}) \notin M$ and $d_u(s_2, s_4^{s_3^\lambda}, s_5^{s_3^\lambda}) \neq 0$. We are done.

CASE 3: $[s_3, s_5] \in E(\mathcal{T})$.



Let us apply the adjoint map to the above picture (see Section 2.7), and let us choose

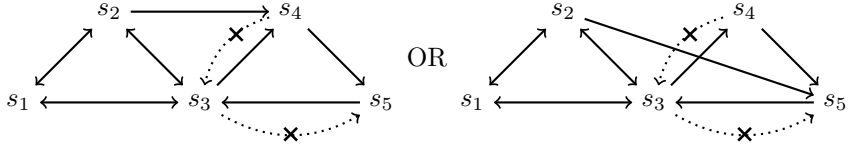
$$s'_1 = s_1^*, \quad s'_2 = s_2^*, \quad s'_3 = s_3^*, \quad s'_4 = s_5^*, \quad s'_5 = s_4^*.$$

Using Lemma 2.3, we get a graph $\Gamma(s'_1, s'_2, s'_3, s'_4, s'_5)$ as in Case 2. Thus, by Case 2, there is a two-cycle $(t_1, t_2) \in E(\mathcal{T})$ with $w(t_1, t_2) \notin M$ and

$$\ell_{s_1^*,\dots,s_5^*}(t_1, t_2) = O(1).$$

So, by Lemma 2.3, we get a two-cycle $(t_1^*, t_2^*) \in E(\mathcal{T})$ with $w(t_1^*, t_2^*) = w(t_1, t_2) \notin M$ and $\ell_{s_1,\dots,s_5}(t_1^*, t_2^*) = O(1)$.

CASE 4: $[s_4, s_3] \notin E(\mathcal{T})$, $[s_3, s_5] \notin E(\mathcal{T})$ *and at least one of* $[s_2, s_4] \in E(\mathcal{T})$ *and* $[s_2, s_5] \in E(\mathcal{T})$.



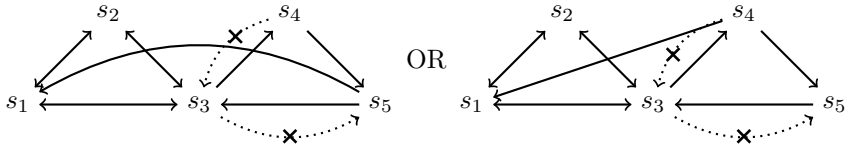
Let $\lambda \in M$ (to be specified).

Since $[s_4, s_3] \notin E(\mathcal{T})$, $w(s_3, s_4) = 0$, moreover $w(s_1, s_2, s_3) \notin M$, therefore, up to replacing λ with $-\lambda$, we may assume that $w(s_1, s_2, s_3^{s_4^\lambda}) \notin M$, so that $(s_1, s_2, s_3^{s_4^\lambda})$ is a triangle of type 1. Therefore, it is two-way directed, so $[s_3^{s_4^\lambda}, s_2] \in E(\mathcal{T})$. Since at least one of $[s_2, s_4]$ and $[s_2, s_5]$ is an edge and $|M| \geq 3$, we may choose $\lambda \in M$ such that $\phi_5(u_2) \neq \pm\lambda\phi_5(u_4)\phi_4(u_2)$ (the $\pm$ sign is needed because in this argument we have possibly replaced λ with $-\lambda$). Therefore

$$\phi_5(u_2) + \lambda\phi_5(u_4)\phi_4(u_2) \neq 0, \text{ that is, } [s_2, s_5^{s_4^\lambda}] \in E(\mathcal{T}).$$

Note that $[s_5^{s_4^\lambda}, s_3^{s_4^\lambda}]$ is a single edge being a conjugate of the single edge $[s_5, s_3]$. By Remark 3.11/(4), the triangle $(s_2, s_5^{s_4^\lambda}, s_3^{s_4^\lambda})$ is of type 2.

CASE 5: *$[s_4, s_3] \notin E(\mathcal{T})$, $[s_3, s_5] \notin E(\mathcal{T})$ and at least one of $[s_5, s_1] \in E(\mathcal{T})$ and $[s_4, s_1] \in E(\mathcal{T})$.*



As in Case 3, one can reduce this case to Case 4 by using the adjoint map.

Now, we can finish the proof of our main theorem when $G = \mathrm{SL}(V)$.

Proof of Theorem 1.5 for $G = \mathrm{SL}(V)$. In the previous sections, we already showed how we can generate an M-closed set of transvections X. The above discussion implies that, in the case $G = \mathrm{SL}(V)$, up to extending the M-closed X with exponent bounded by $O(1)$, we can reach $L_2(X) = \mathbb{F}_q$. Applying Lemma 4.6/(2), we may assume that X is $\mathbb{F}_q$-closed. In particular, such an X contains a full transvection group over $\mathbb{F}_q$ in length $(\log |G|)^{O(1)}$. Now, the theorem follows from [12, Theorem 1.5]. ∎

4.4. ADDITION OF PARAMETERS DEFINING TRANSVECTIONS. In view of the previous section, from now on we assume that V is a symplectic or unitary space with defining form f, $G = \mathrm{Sp}(V)$ or $G = \mathrm{SU}(V)$ and $X \subset G$ is an $\mathbb{F}_{q_0}$-closed subset of generating transvections with $\mathrm{diam}(\Gamma(X)) \leq 2$. For simpler notation, we can add 1 to X and assume that $1 \in X$.

Recall that for any nonzero singular vector $v \in V$, the transvection subgroup associated to v is $T_v = (1 + v \otimes \varphi_v)^{\mathcal{F}}$, where in the symplectic case $\mathcal{F} = \mathbb{F}_q$, while in the unitary case $\mathcal{F} = \mathbb{F}_{q_0} \cdot \lambda_0$ where $\lambda_0 \in \mathbb{F}_q^\times$ is any field element with $\mathrm{Tr}(\lambda_0) = 0$. In what follows, for any nonzero singular $v \in V$, we denote by t_v an arbitrary nonzero element of the associated transvection subgroup T_v.

By our assumptions, we have $X = \bigcup \{T_v \,|\, v \in {}_V X\}$. Our main goal is to prove that

$$\ell_X(T_{\sum_{i=1}^k a_i}) = O(n^c)$$

for any $T_{a_1}, \ldots, T_{a_k} \subset X$ where $k \leq n$ and $\sum_{i=1}^k a_i$ is singular. Since $\langle {}_V X \rangle = V$ this implies that $\ell_X(\mathcal{T}) = O(n^c)$. (Note that since $T_{a_i} = T_{\lambda_i a_i}$, we do not need to take arbitrary linear combinations.) First we prove this for $k = 2$.

LEMMA 4.15: *Let $T_a, T_b \subset X$ for some $a, b \in V$ such that $a + b$ is singular. Then $\ell_X(T_{a+b}) \leq c$ for some constant c.*

Proof. If $\langle a\rangle = \langle b\rangle$, then $T_{a+b} = T_a = T_b$, so there is nothing to prove. If (t_a, t_b) is a (two-way) directed edge, then let $W = \langle a, b\rangle$. Now, $V = W \oplus W^\perp$ where $W^\perp = \ker(\varphi_a) \cap \ker(\varphi_b)$, and the restriction to W defines an isomorphism $\langle T_a, T_b\rangle \to \langle (T_a)_W, (T_b)_W\rangle \simeq \mathrm{SL}(2, K)$ for some field $K \leq \mathbb{F}_q$. By Theorem 3.7, K must be equal to $\mathbb{F}_{q_0}$. Since $(T_{a+b})_W \leq \mathrm{SL}(2, \mathbb{F}_{q_0})$, we can apply the strong form of Babai's conjecture for $\mathrm{SL}(2, q_0)$ to get

$$\ell_X(T_{a+b}) \leq \ell_{T_a \cup T_b}(T_{a+b}) = O\Big(\frac{\log |\mathrm{SL}(2, q_0)|}{\log |T_a \cup T_b|}\Big) = O(1).$$

Now, let us assume that (t_a, t_b) is not an edge. Since $\Gamma(X)$ has diameter 2, there exists a path t_a, t_c, t_b in $\Gamma(X)$ for some $t_c \in X$. So $c \in_V X$ satisfies $f(c, a) \neq 0$, and $f(b, c) \neq 0$.

Since $f(b, a) = 0$, $f(a, a) = 0$, $f(c, a) \neq 0$ we deduce that $c \notin \langle a, b\rangle$. This implies that $\langle a, b, c\rangle$ is a space of dimension 3. Let

$$w := -f(c, b)a + f(c, a)b.$$

It is not hard to see that $\langle w\rangle$ is the radical of the space $\langle a, b, c\rangle$. Since X_{V^*} spans V^*, there exists $t_d \in X$ with $f(d, w) \neq 0$. Therefore, at least one of (t_a, t_d) and (t_b, t_d) is an edge in $\Gamma(X)$. Let $W := \langle a, b, c, d\rangle \leq V$. We claim that the radical of W is trivial. Indeed, $d \notin \langle w\rangle^\perp$ while $\langle a, b, c\rangle$ is contained in $\langle w\rangle^\perp$ so $\langle w\rangle^\perp \cap W = \langle a, b, c\rangle$. Therefore, the radical of W is contained in the radical of $\langle a, b, c\rangle$, which is $\langle w\rangle$. We conclude that the radical of W is trivial since w is not orthogonal to $d \in W$.

Let $X' = (T_a)_W \cup (T_b)_W \cup (T_c)_W \cup (T_d)_W$. We prove that it generates a subgroup of $SL(W)$, isomorphic to either $\mathrm{Sp}(4, L)$ or $\mathrm{SU}(4, L)$ for a subfield $L \leq \mathbb{F}_q$ with $|\mathbb{F}_q : L| \leq 2$.

In order to do so we check that the conditions of Theorem 3.8 hold. Clearly, conditions (1) and (3) hold. The irreducibility of X' will follow if we verify the conditions of Theorem 3.1. Plainly, a, b, c, d span W and the non-degeneracy of W implies condition (2). The transvection graph induced by X' is connected since a and b are not orthogonal to c and d is not orthogonal to at least one of a and b. Thus the corresponding transvections in the transvection subgroups are connected and hence the induced subgraph of the transvection graph is strongly connected since every edge is two-way directed in this case. Thus, by Theorem 3.8, $\langle X'\rangle \simeq \mathrm{Sp}(4, L)$ or $\mathrm{SU}(4, L)$ for some subfield $L \leq \mathbb{F}_q$. Finally, $X' \subset \mathrm{SL}(W)$ is a union of full transvection groups over $\mathbb{F}_{q_0}$, so the weights of the cycles in $\Gamma(X')$ generate a subfield $L \geq \mathbb{F}_{q_0}$, so $|\mathbb{F}_q : L| \leq 2$ holds.

The bounded-rank case of the Babai conjecture (in its strong form) can be applied to deduce that

$$\ell_X(T_{a+b}) \le \ell_{X'}((T_{a+b})_W) \le O\Big(\frac{\log|\mathrm{SL}(4,q)|}{\log|X'|}\Big) = O(1). \qquad \blacksquare$$

Remark 4.16: In the symplectic case, we managed to find elementary arguments to prove this lemma with specific constant $c = 21$ as follows.

CASE 1 $(\mathbf{a},\mathbf{b}) \in \mathbf{E}(\mathcal{T})$: Let $T_a = (1+a\otimes\varphi_a)^{\mathbb{F}_q}$, $T_b = (1+b\otimes\varphi_b)^{\mathbb{F}_q}$. Choosing $\lambda := \varphi_b(a)^{-1}$, we get

$$T_{a+b} = (1+\lambda b\otimes\varphi_b)(1+a\otimes\varphi_a)^{\mathbb{F}_q}(1-\lambda b\otimes\varphi_b),$$

so $\ell_X(T_{a+b}) \le 3$.

CASE 2 $(\mathbf{a},\mathbf{b}) \notin \mathbf{E}(\mathcal{T})$: Let $T_c, T_d \subset X$ as in the proof of Lemma 4.15.

First, let us assume that $f(c, a+b) \neq 0$. Since (a,c) and (b,c) are edges by construction, we have $\ell_X(T_{a+c}) \le 3$ and $\ell_X(T_{b-c}) \le 3$ by Case 1. Then we have

$$f(a+c, b-c) = f(c, a+b) \neq 0,$$

so using Case 1 again we get that $\ell_{T_{a+c}\cup T_{b-c}}(T_{a+b}) \le 3$. Hence $\ell_X(T_{a+b}) \le 9$.

Now, let us assume that $f(c, a+b) = 0$. By construction, $f(d, a+b) \neq 0$, which implies that at least one of $f(d,a)$ and $f(d,b)$ is not zero. If both of them are not zero, then we can use the same argument as in the previous case (but using d instead of c). So let us assume, say, $f(d,a) \neq 0$ and $f(d,b) = 0$.

Let $d' = d$ if $f(c,d) \neq 0$, while $d' = a+d$ if $f(c,d) = 0$. By our assumptions and by Case 1,

$$\ell_X(T_{d'}) \le 3, \quad (t_c, t_{d'}), (t_a, t_{d'}), (t_{a+b}, t_{d'}) \in E(\mathcal{T}) \text{ but } (t_b, t_{d'}) \notin E(\mathcal{T}).$$

Using the construction of Case 1 again, we can deduce that $\ell_X(T_{\tau c+d'}) \le 5$ for every $\tau \in \mathbb{F}_q^\times$. Since $q > 2$, we may choose a $\tau \neq 0$ satisfying $f(\tau c + d', a) \neq 0$. Now, $t_{\tau c+d'}$ is a neighbour of each of t_a, t_b, t_{a+b}, so the argument of the first paragraph of Case 2 (but using $\tau c + d'$ instead of c) can be used to construct t_{a+b}. Using this construction in a careful way, one can show that $\ell_X(T_{a+b}) \le 21$.

Now, we are able to generate all transvections in the symplectic case.

THEOREM 4.17: *Let us assume that $G = \mathrm{Sp}(V)$ and let c be the constant in Lemma 4.15. Then we have $\ell_X(\mathcal{T}) \le cn^{\log_2 c}$. In view of Remark 4.16, the bound $\ell_X(\mathcal{T}) \le 21n^{4.4}$ holds.*

Proof. Since $\mathcal{T} = \bigcup_{0\neq v\in V} T_v$ it is enough to generate T_v for any $0 \neq v \in V$. Our argument is essentially the same as the proof of [12, Lemma 4.12]. For the convenience of the reader, we present a detailed proof.

Let $v \neq 0$ be a vector in V. Since ${}_VX$ spans V, there are $a_1, \dots, a_k \in {}_VX$ such that $k \leq n$ and $v = \sum_{i=1}^{k} a_i$. Let $l(k) = \lceil \log_2 k \rceil$ so $l(k)$ is the smallest integer satisfying $k \leq 2^{l(k)}$. We prove that $\ell_X(T_v) \leq c^{l(k)} \leq cn^{\log_2 c}$ by using induction on k.

The claim trivially holds for $k = 1$. For an arbitrary $k \leq n$, let

$$v_1 = \sum_{i=1}^{\lceil k/2\rceil} a_i \quad \text{and} \quad v_2 = \sum_{i=\lceil k/2\rceil+1}^{k} a_i.$$

Let $X' = X \cup T_{v_1} \cup T_{v_2}$. Since

$$l(\lceil k/2\rceil) \leq l(k) - 1,$$

by induction we have $\ell_X(X') \leq c^{l(k)-1}$. On the other hand, $\ell_{X'}(T_v) \leq c$ by Lemma 4.15. So, $\ell_X(T_v) \leq \ell_X(X') \cdot \ell_{X'}(T_v) \leq c^{l(k)}$, as claimed. ∎

For the remainder, let V be a unitary space over $\mathbb{F}_q$ and $G = \mathrm{SU}(V)$. Now, $q_0 = \sqrt{q}$, so $\mathbb{F}_q$ is a 2-dimensional $\mathbb{F}_{q_0}$-space.

LEMMA 4.18: *Let $T_{v_1}, T_{v_2}, T_{v_3} \subset X$ for some $v_1, v_2, v_3 \in V$ and let us assume that $v_1 + v_2 + v_3$ is singular. Then $\ell_X(T_{v_1+v_2+v_3}) \leq c'$ for some constant c'.*

Proof. Let $\alpha_{ij} = f(v_i, v_j)$ for every $i \neq j$.

First, let us assume that $\mathrm{Tr}(\alpha_{ij}) = 0$ for some $i \neq j$. If, for example, $\mathrm{Tr}(\alpha_{12}) = 0$, then $f(v_1 + v_2, v_1 + v_2) = \alpha_{12} + \alpha_{21} = \mathrm{Tr}(\alpha_{12}) = 0$, so $v_1 + v_2$ is singular. Applying Lemma 4.15 twice, we get that

$$\ell_X(T_{v_1+v_2+v_3}) \leq \ell_X(T_{v_1+v_2}) \cdot \ell_{X\cup T_{v_1+v_2}}(T_{v_1+v_2+v_3}) \leq c^2.$$

Thus, for the remainder we assume that $\mathrm{Tr}(\alpha_{ij}) \neq 0$ for every $i \neq j$. In particular, each $\alpha_{ij} \neq 0$ $(i \neq j)$, that is, $t_{v_1}, t_{v_2}, t_{v_3}$ is a triangle in $\Gamma(X)$.

Now, let us assume that $\alpha_{31}/\alpha_{21} \notin \mathbb{F}_{q_0}$. We claim that there is a $\lambda \in \mathbb{F}_q$ such that both $\lambda v_1 + v_2$ and $(1 - \lambda)v_1 + v_3$ are singular. For any $\lambda \in \mathbb{F}_q$ we have

$$\lambda v_1 + v_2 \text{ is singular} \iff \mathrm{Tr}(\lambda\alpha_{21}) = f(\lambda v_1 + v_2, \lambda v_1 + v_2) = 0.$$

Similarly,

$$(1 - \lambda)v_1 + v_3 \text{ is singular} \iff \mathrm{Tr}((1 - \lambda)\alpha_{31}) = 0 \iff \mathrm{Tr}(\lambda\alpha_{31}) = \mathrm{Tr}(\alpha_{31}).$$

For any $\gamma \in \mathbb{F}_q^\times$, the function $x \to \mathrm{Tr}(x\gamma)$ is an $\mathbb{F}_{q_0}$-linear map from $\mathbb{F}_q$ onto $\mathbb{F}_{q_0}$, so $\{x \in \mathbb{F}_q \mid \mathrm{Tr}(x\gamma) = \delta\} \subset \mathbb{F}_q$ is an affine line of the $\mathbb{F}_{q_0}$-space $\mathbb{F}_q$. Thus, we need a $\lambda \in \mathbb{F}_q$ in the intersection of the affine lines

$$\{x \in \mathbb{F}_q \mid \mathrm{Tr}(x\alpha_{21}) = 0\} \quad \text{and} \quad \{x \in \mathbb{F}_q \mid \mathrm{Tr}(x\alpha_{31}) = \mathrm{Tr}(\alpha_{31}).\}$$

Our assumption $\alpha_{31}/\alpha_{21} \notin \mathbb{F}_{q_0}$ exactly means that these lines are not parallel, so there is a unique such λ. Using Lemma 4.15 again, we get that

$$\ell_X(T_{v_1+v_2+v_3}) \le \ell_X(T_{\lambda v_1+v_2}) \cdot \ell_{X \cup T_{\lambda v_1+v_2}}(T_{v_1+v_2+v_3}) \le c^2.$$

Applying any permutation of the indices we can also assume that α_{12}/α_{32} and α_{23}/α_{13} are in $\mathbb{F}_{q_0}$ for the remainder.

Finally, let $k := \alpha_{31}/\alpha_{21} = \alpha_{13}/\alpha_{12} \in \mathbb{F}_{q_0}^\times$, $l := \alpha_{23}/\alpha_{13} = \alpha_{32}/\alpha_{31} \in \mathbb{F}_{q_0}^\times$, and $m := \alpha_{12}/\alpha_{32} = \alpha_{21}/\alpha_{23} \in \mathbb{F}_{q_0}^\times$. Then we have

$$\begin{aligned}
\alpha_{21} &= m\alpha_{23} = lm\alpha_{13} = klm\alpha_{12},\\
\alpha_{13} &= \frac{\alpha_{12}}{\alpha_{21}}\alpha_{31} = \frac{\alpha_{31}}{klm},\\
\alpha_{32} &= \frac{\alpha_{31}}{\alpha_{13}}\alpha_{23} = klm\alpha_{23}.
\end{aligned}$$

Furthermore, $0 \ne \mathrm{Tr}(\alpha_{12}) = \alpha_{12}(1 + klm)$, so $klm \ne -1$. Thus, we get

$$\begin{aligned}
\begin{vmatrix} 0 & \varphi_{v_1}(v_2) & \varphi_{v_1}(v_3) \\ \varphi_{v_2}(v_1) & 0 & \varphi_{v_2}(v_3) \\ \varphi_{v_3}(v_1) & \varphi_{v_3}(v_2) & 0 \end{vmatrix} &= \alpha_{12}\alpha_{23}\alpha_{31} + \alpha_{21}\alpha_{32}\alpha_{13}\\
&= \alpha_{12}\alpha_{23}\alpha_{31}(1 + klm) \ne 0.
\end{aligned}$$

It follows that $W := \langle v_1, v_2, v_3\rangle$ is a non-degenerate 3-dimensional subspace of V and $V = W \oplus W^\perp$. Thus, $T_{v_i} = (T_{v_i})_W \oplus 1_{W^\perp}$ for each i, and

$$T_{v_1+v_2+v_3} = (T_{v_1+v_2+v_3})_W \oplus 1_{W^\perp}.$$

Defining $X' = (T_{v_1})_W \cup (T_{v_2})_W \cup (T_{v_3})_W \subset \mathrm{SU}(W)$ we have $\langle X'\rangle = \mathrm{SU}(W)$ by Theorem 3.1 and Theorem 3.8, so we can apply the strong form of Babai's conjecture to deduce that

$$\ell_X(T_{v_1+v_2+v_3}) \le \ell_{X'}((T_{v_1+v_2+v_3})_W) \le O\Big(\frac{\log|\mathrm{SU}(3,q)|}{\log|X'|}\Big) = O(1).$$

The proof is complete. ∎

LEMMA 4.19: *Let $v = \sum_{i=1}^s v_i$ for some singular vectors $v_1, \dots, v_s \in V$. Then $v - \lambda v_i$ is singular for some $1 \le i \le s$ and for some $\lambda \in \mathbb{F}_{q_0}$.*

Proof. If v is singular, then we can take $\lambda = 0$. Otherwise,

$$0 \neq 2f(v,v) = \sum_{i=1}^{s} \operatorname{Tr}(f(v,v_i)),$$

so there is an i such that $\operatorname{Tr}(f(v,v_i)) \neq 0$. Choosing $\lambda = \frac{f(v,v)}{\operatorname{Tr}(f(v,v_i))} \in \mathbb{F}_{q_0}$ we get that

$$\begin{aligned} f(v-\lambda v_i, v-\lambda v_i) &= f(v,v) - \lambda f(v_i,v) - \lambda f(v,v_i) \\ &= f(v,v) - \lambda \operatorname{Tr}(f(v,v_i)) = 0. \end{aligned}$$ ∎

THEOREM 4.20: *Let us assume that $G = \mathrm{SU}(V)$ and let c be the constant in Lemma 4.18. Then we have $\ell_X(\mathcal{T}) \leq O(n^{2\log_2 c})$.*

Proof. Let $v \in V$ be any nonzero singular vector. We need to generate T_v from X. As in the symplectic case, since ${}_V X$ spans V, there exist $a_1, \ldots, a_k \in {}_V X$ such that $k \leq n$ and $v = \sum_{i=1}^{k} a_i$, $T_{a_i} \subset X$ for every $1 \leq i \leq k$. Using the notation $l(k) = 2\log_2 k$, our goal is to prove that

$$\ell_X(T_v) \leq O(c^{l(k)}) = O(n^{2\log_2 c}).$$

Previously we proved this claim for $k \leq 3$. For an arbitrary $4 \leq k \leq n$ our goal is to write v as a sum of 3 singular vectors v_1, v_2, v_3 two of which are a linear combination of roughly $k/2$ many a_i's. First we consider the decomposition $v = u_1 + u_2$ with

$$u_1 = \sum_{i=1}^{\lceil k/2 \rceil} a_i \quad \text{and} \quad u_2 = \sum_{i=\lceil k/2 \rceil + 1}^{k} a_i.$$

Using Lemma 4.19 with, say, $i = 1$ (which we can assume), we get that $v_1 := u_1 - \lambda a_1$ is singular for some $\lambda \in \mathbb{F}_{q_0}$. Now, if $(1+\lambda)a_1 + u_2$ is singular, then we choose

$$v_2 = (1+\lambda)a_1 + u_2,\ v_3 = 0.$$

Finally, if $(1+\lambda)a_1 + u_2 = (1+\lambda)a_1 + \sum_{i=\lceil k/2 \rceil + 1}^{k} a_i$ is not singular, then we apply Lemma 4.19 again to write $(1+\lambda)a_1 + u_2 = v_2 + v_3$ with singular vectors $v_2 = \mu a_s$, $v_3 = (1+\lambda)a_1 + u_2 - \mu a_s$ for some $s \in \{1, \lceil k/2 \rceil + 1, \ldots, k\}$ and for some $\mu \in \mathbb{F}_{q_0}$.

Let $X' = X \cup T_{v_1} \cup T_{v_2} \cup T_{v_3}$. By our construction, each v_i is a linear combination of at most $\lceil k/2 \rceil + 1 < k$ many a_i's. Now, if k is bounded, then $\ell_X(T_v) = O(1)$ follows by a repeated application of Lemma 4.18.

So, we can assume that $k \geq 10 > 4/(\sqrt{2}-1)$. Then

$$\lceil k/2 \rceil + 1 \leq k/2 + 2 \leq k/\sqrt{2},$$

that is,

$$l(\lceil k/2 \rceil + 1) \leq l(k/\sqrt{2}) = 2\log_2(k/\sqrt{2}) = l(k) - 1.$$

Using an induction argument and Lemma 4.18, we get that

$$\ell_X(T_v) \leq \ell_X(X') \cdot \ell_{X'}(T_v) \leq O(c^{l(k)-1}) \cdot c = O(c^{l(k)}).$$

The result follows. ∎

In order to obtain our main result it is enough to see that $\ell_{\mathcal{T}}(G) = O(\frac{\log|G|}{\log|\mathcal{T}|})$. Since $\mathcal{T}$ is a conjugacy class, one can see that this holds by a result of Liebeck and Shalev [20]. On the other hand $\ell_{\mathcal{T}}(G) = O(n^2)$ can easily be proved using explicit Gaussian elimination-like algorithm.

4.5. Proof of Theorem 1.7. The proof of Theorem 1.7 follows easily from our previous proof. We only give a sketch here.

The only new thing we need is the following modification of Theorem 3.7 in our situation.

Proposition 4.21: *With the assumptions of Theorem 1.7, let* $(s,t) \in E(\mathcal{T})$ *and* $W = \langle s,t \rangle$. *Then* $H := \langle s^{K_0}, t \rangle \simeq \mathrm{Sp}(W) \oplus 1_{W^\perp}$ *or* $\mathrm{SU}(W) \oplus 1_{W^\perp}$, *moreover* $\mathrm{diam}(\mathrm{Cay}(H, \{s^{K_0}, t\})) = O(1)$.

Now, in order to prove Theorem 1.7, first we can apply a modification of the argument of Lemma 2.1 to get a generating set of transvections X for G (containing a transvections subgroup over K_0), in length $O(n^2)$ (for details, see the second paragraph of the proof of [12, Lemma 4.2]). After that, we can use the arguments of Section 4.1, to enlarge X in order to guarantee that the diameter of $\Gamma(X')$ is 2 for every $X \subset X' \subset \mathcal{T}$. Now, we need only use the first paragraph of the proof of Lemma 4.6 along with Proposition 4.21, to generate the K_0-closure of X, even in length $O(1)$. After that, we can use (a modification of) the arguments of Section 4.4 to get all the transvections.

Acknowledgements. We thank the anonymous referee for very helpful comments on a previous version of this paper.

Open access funding provided by Eötvös Loránd University

References

[1] E. Artin, *Geometric Algebra*, Interscience, New York–London, 1957.

[2] L. Babai and Á. Seress, *On the diameter of Cayley graphs of the symmetric group*, Journal of Combinatorial Theory. Series A **49** (1988), 175–179.

[3] L. Babai and Á.Seress, *On the diameter of permutation groups*, European Journal of Combinatorics **13** (1992), 231–243.

[4] J. Bajpai, D. Dona and H. A. Helfgott, *Growth estimates and diameter bounds for classical Chevalley groups*, `https://arxiv.org/abs/2110.02942`

[5] A. Biswas and Y. Yang, *A diameter bound for finite simple groups of large rank*, Journal of the London Mathematical Society **95** (2017), 455–474.

[6] E. Breuillard, B. Green and T. Tao, *Approximate subgroups of linear groups*, Geometric and Functional Analysis **21** (2011), 774–819.

[7] R. Brown and S. P. Humphries, *Orbits under symplectic transvections I*, Proceedings of the London Mathematical Society **52** (1986), 517–531.

[8] L. Di Martino, A. Previtali and R. Radina, *Sets of transvections generating subgroups isomorphic to special linear groups*, Communications in Algebra **33** (2005), 1663–1691.

[9] L. E. Dickson, *Linear Groups: With an Exposition of the Galois Field Theory*, Dover, New York, 1958.

[10] S. Eberhard and U. Jezernik, *Babai's conjecture for high-rank classical groups with random generators*, Inventiones Mathematicae **227** (2022), 149–210.

[11] D. Gorenstein, *Finite Groups*, Chelsea, New York, 1980.

[12] Z. Halasi, *Diameter of Cayley graphs of* $\mathrm{SL}(n,p)$ *with generating sets containing a transvection*, Journal of Algebra **569** (2021), 195–219.

[13] Z. Halasi, A. Maróti, L. Pyber and Y. Qiao, *An improved diameter bound for finite simple groups of Lie type*, Bulletin of the London Mathematical Society **51** (2019), 645–657.

[14] H. A. Helfgott, *Growth and generation in* $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$, Annals of Mathematics **167** (2008), 601–623.

[15] H. A. Helfgott, *Growth in* $\mathrm{SL}_3(\mathbb{Z}/p\mathbb{Z})$, Journal of the European Mathematical Society **13** (2011), 761–851.

[16] H. A. Helfgott and Á. Seress, *On the diameter of permutation groups*, Annals of Mathematics **179** (2014), 611–658.

[17] B. Huppert, *Endliche Gruppen. I.* Die Grundlehren der Mathematischen Wissenschaften, Vol. 134, Springer, Berlin–New York, 1967.

[18] I. M. Isaacs, *Character Theory of Finite Groups*, Dover, New York, 1994.

[19] P. Kleidman and M. W. Liebeck, *The Subgroup Structure of the Finite Classical Groups*, London Mathematical Society Lecture Note Series, Vol. 129, Cambridge University Press, Cambridge, 1990.

[20] M. W. Liebeck and A. Shalev, *Diameters of finite simple groups: sharp bounds and applications*, Annals of Mathematics **154** (2001), 383–406.

[21] A. Lubotzky. *Discrete Groups, Expanding Graphs and Invariant Measures*, Modern Birkhäuser Classics. Birkhäuser, Basel, 2010.

[22] B. H. Neumann, *A problem of Paul Erdős on groups*, Journal of the Australian Mathematical Society **21** (1976), 467–472.

[23] L. Pyber and E. Szabó, *Growth in finite simple groups of Lie type*, Journal of the American Mathematical Society **29** (2016), 95–146.

[24] A. Wagner, *Groups generated by elations*, Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg **41** (1974), 190–205.

[25] R. A. Wilson, *The Finite Simple Groups*, Graduate Texts in Mathematics, Vol. 251, Springer, London, 2009.