dc_650_12

MTA DOKTORI ÉRTEKEZÉS

Szabó Endre

2013

dc_650_12

DISSERTATION

Application of Algebraic Geometry in Combinatorics and in Group Theory

Szabó Endre
Rényi Alfréd Matematikai
Institute of Mathematics

2013

dc_650_12

dc_650_12

Contents

Overview

In this Dissertation we discuss the following three important, intertwined themes, and their applications. Here we list the three themes together with the most important results we prove about them.

1. *Bounding the number of incidences.*
   The basic problem is to find upper bound on the number of incidences among p points and q geometric figures. The main result of the Dissertation on this theme is Theorem 6, which generalises (among other things) the famous Szemerédi–Trotter theorem.
2. *How to find groups?*
   We study when does it happen that among three times n geometric figures we can find Cn^2 occurrences of certain three-figure configurations. We find that typically there is a large symmetry-group responsible for the too-many coincidences. The main results of the Dissertation on this theme are Theorems 14 and 15. They are related (among other things) to Hrushovski's Group Configuration Theorem [**75**] and the recent paper of Tao [**156**].
3. *Growth in groups.*
   For finite subsets α of a group we study how the size of its powers α^n vary in terms of n. We are especially interested in the structure of subsets with slowly growing α^n. The main results of the Dissertation on this theme are Theorems 16, 17 and 18. They are closely related (among many other things) to the Freiman–Ruzsa theorem, and the results of Helfgott [**72**], Breuillard–Green–Tao [**27**], Bourgain- Gamburd [**12**].

The first two of the themes are of combinatorial flavor, the last one belongs to group theory. Although it does not appear as a separate chapter, algebraic geometry has a crucial role in all three themes.

In the Dissertation there are several interesting applications of the above results.

4. *Applications in combinatorics.*
   - Corollary 19 improves significantly the best known exponent in a problem of Hirzebruch [**74**].
   - Theorem 20 answers a question of Erdős, Lovász and Vesztergombi [**51**].
   - Corollary 24 solves a conjecture of Székely (see [**42**, Conjecture 3.41]). Moreover, Theorem 23 is a vast generalisation of this conjecture.
   - Theorem 26 partially solves a variation of the so called Orchard problem (see [**83, 152**]).

5. *Applications in group theory.*
- Corollary 27 solves the conjecture of Babai [**7**] for simple groups of bounded rank.
- Theorem 29 solves the conjecture of Liebeck, Nikolov and Shalev [**99**] for simple groups of bounded rank. Theorem 30 is a variation on the same theme.
- Theorem 33 solves the conjecture of Weiss [**170**] in the class of BCP(r) groups.

The Dissertation's place among related areas of mathematics

In this section I will try to position the main results of this Dissertation in their mathematical vicinity.

The first theme is bounding various incidence numbers. Our starting point is the following famous theorem of Szemerédi–Trotter [**154**]. Among p points and q lines in the plane there are at most $\mathcal{O}\left(p^{2/3}q^{2/3}+p+q\right)$ incidences. Several generalisations of this result emerged since then, bounding the number of incidences among p points and q geometric figures (instead of lines). Let me list a few just to get a feeling for it. Pach–Sharir [**120**] (see Theorem 1) bounds incidence numbers for plane curves; Chazelle–Edelsbrunner–Guibas–Sharir [**31**] and Solymosi–Tao [**145**] studies hyperplanes in the euclidean space $\mathbb{R}^n$; Tóth Csaba [**162**] deals with complex lines in $\mathbb{C}^2$; Bourgain–Katz–Tao [**15**], and Bourgain [**10**] estimates incidence numbers concerning lines, and certain hyperbolas in the projective plane over the finite field $\mathbb{F}_p$ for any prime p. Theorem 6 in this Dissertation is a higher dimensional generalisation. It bounds the number of incidences among p points and q subvarieties of bounded degree in $\mathbb{R}^n$ or $\mathbb{C}^n$. It would be interesting to extend this result to finite fields as well, but this direction is still wide open. Theorem 6 is the most important link between our first and second theme, we use it to exclude certain kind of degeneracies during the construction of groups.

Our second theme is a very general phenomenon in mathematics. If in a geometric situation there are many unexpected coincidences then we may expect a large symmetry-group lurking in the background. There is a vast number of variations on this theme, here we mention only two results. In his Group Configuration Theorem [**75**](see also [**122**]) Hrushovski considers a model-theoretic scenario. He constructs symmetry-groups in extreme generality. Roughly it goes like this. Let us consider a *stable*[1] mathematical theory $\mathcal{T}$, and two *d-dimensional*[2] families[3] of "function-like" two-variable

[1]Stability of a mathematical theory $\mathcal{T}$ essentially means, that in the models of $\mathcal{T}$ there aren't too many "types" of elements.

[2] In model-theory, under appropriate conditions, one can define an extremely general notion of dimension similar to the Krull-dimension used in algebra.

[3] Members of a *family* are parametrised by an index set. The dimension of the family is the dimension of this index set.

relations[4] in a model of $\mathcal{T}$. The direct product of the two index sets is $2d$-dimensional, so we expect to get a $2d$-dimensional family of pairwise *compositions*.[5] If the compositions form instead a d-dimensional family,[6] then both families can be induced from a group in the following manner. In the theory $\mathcal{T}$ one can define a set X equipped with the action of a d-dimensional group G of symmetries, and both families "essentially look like" the family R_g of relations on X defined by the formula $R_g = \{(x, y) \in X \times X \mid y = gx\}$ (where g runs through the elements of G). The special case of the Group Configuration theorem where $\mathcal{T}$ is the theory of algebraically closed fields, which is reproduced in the Dissertation as Lemma 1.3.6, plays an important role in the proof of Theorem 14 and Theorem 15.

While the result of Hrushovski has a rather "continuous nature", in Theorem 14 and Theorem 15 we face a combinatorial situation, and find similar consequences. The precursor of this two theorems is a paper of Elekes–Rónyai [**45**] (where the group of symmetries does not yet appear explicitly). An interesting new development is the work of Tao [**156**], where he studies the phenomenon that, if A, B are "very large" subsets of a finite field, and P is a two-variate polynomial, then the set $P(A, B)$ is "typically" fills up almost the whole field. The exceptional polynomials, analogously to our Theorem 14 and Theorem 15, are just the reparametrisations of either the addition, or the multiplication, so they originate from the additive or the multiplicative group of the field. Tao even mentions in his blog [**157**] that the "Elekes–Szabó theory", as he puts it, may have an important role in the further investigation of his problem.

It is worth noting here that the finite point-configurations appearing in Theorem 14 and Theorem 15 show up also in the group obtained there, and one can see easily that they form a non-growing subset of that group. A finite subset α of a group is called *non-growing* if its third *power*,[7] denoted by α^3, has size at most $K|\alpha|$.

With this last comment we have arrived at our third theme, the study of non-growing subsets of groups. The theme is interesting for commutative as well as non-commutative groups, and both versions have plenty of applications within and outside of group-theory (see later). It is quite remarkable how the commutative and non-commutative worlds intertwine, and interact with each other in these investigations. Even though they study rather different-looking phenomena, they borrow a large number of ideas and methods from each other.

We begin with the commutative groups. The Freiman–Ruzsa theorem [**55**] (see also [**138**] for the proof given by Ruzsa) is a fundamental result in additive combinatorics. It is the following. If $\alpha \subseteq \mathbb{Z}$ is a finite

[4]It means that almost all x are related to a bounded number of y only.

[5]The *composition* of two relations R and S is the relation defined by the formula

$$\{(x, z) \mid \exists y : (x, y) \in R \text{ and } (y, z) \in S\}$$

[6] Typically all compositions are different, they form a $2d$-dimensional family. Sometimes there are coincidences, and we may get a smaller dimension. Our case is the "maximal degeneration".

[7] α^n denotes the set of all n-term products formed from the elements of α.

subset such that $\left|\alpha+\alpha\right| \leq K|\alpha|$ holds,[8] then α can be covered by a $d(K)$-dimensional generalised arithmetic progression of size $f(K)|\alpha|$. Later Green and Ruzsa [**65**] generalised the theorem for arbitrary abelian group. In this generality a non-growing subset α can be covered by the sum of a generalised arithmetic progression and a finite subgroup. (This kind of sums are called *coset progressions*.) The most important open question in this direction is, whether one can find a description of non-growing sets such that the parameters (like the $f(K)$ above) depend on K polynomially. For example, is it true that all non-growing sets $\alpha \subseteq \mathbb{Z}_2^n$ can be covered by at most CK^m cosets of a subgroup of size at most $|\alpha|$ (where C and m are constants independent of everything)?

After this detour on commutative structures let us return to the world of not necessarily commutative groups. Let α be a non-growing set in a group. What can we say about the structure of α? The first, and at the same time the most well-known result in this direction is the theorem of Gromov [**67**]. The size of α^n can be bounded from above[9] by a polynomial of n if and only if the subgroup generated by α is *virtually nilpotent*.[10] (Here the polynomial may depend on the group.)

The next breakthrough in the study of non-growing sets was the theorem of Helfgott [**72**]. Let α be a generating system in the *group* $SL(2,p)$,[11] (where p is an arbitrary prime). Then either α grows exponentially, i.e $\left|\alpha^3\right| \geq |\alpha|^{1+\varepsilon}$ for some constant ε independent of everything, or $\alpha^3 = G$ (so there is no room for further expansion).[12] A strong motivation for Helfgott was that his theorem implies immediately the Babai conjecture for the groups $SL(2,p)$ (see Corollary 27). Later it turned out that Helfgott's theorem can be significantly extended. According to the *Product theorem*, the same statement is valid in the groups $SL(n,q)$ for arbitrary prime-power q.[13] The importance of the Product theorem is indicated by the fact that it was proved independently at the same time by two different groups: Breuillard–Green–Tao [**25**] and Pyber–Szabó [**133**].

In the last few years there were a lot of advances in understanding the structure of non-growing sets. Here we mention two of them only. Breuillard–Green–Tao [**27**] studied non-growing subsets of arbitrary groups.

---

[8] It follows from the Plünecke–Ruzsa estimates that in this case $|\alpha+\alpha+\alpha| \leq K^2|\alpha|$, i.e. α is non-growing. In non-commutative groups this reasoning fails, this is why we insisted on bounding the size of α^3. Interestingly, in arbitrary groups, the size of the higher powers of α can be bounded in terms of $\left|\alpha^3\right|$.

[9] In Gromov's theorem we bound all powers of α, not just α^3 as above.

[10] A group is *virtually nilpotent* if it has a nilpotent subgroup of finite index.

[11] for a prime p, $SL(n,p)$ denotes the group of those $n \times n$ matrices of determinant 1, whose entries are taken from the p-element field $\mathbb{F}_p$ (i.e. the ring of remainder-classes modulo p); the group operation is the multiplication of matrices. More generally, if q is a power of a prime, and $\mathbb{F}$ is an arbitrary field, then $SL(n,q)$ and $SL(n,\mathbb{F})$ denote the groups of those $n \times n$ matrices of determinant 1, whose entries are taken from the q-element field $\mathbb{F}_q$ and the field $\mathbb{F}$ respectively.)

[12] Instead of $\alpha^3 = G$, Helfgott proved only that $\alpha^k = G$ with an appropriate value k.

[13] In fact the Product theorem deals with all simple subgroups of $SL(n,q)$, hence all finite simple groups but the alternating groups. The constant ε depends only on n (i.e. the rank of the group).

Their result is a common generalisation of Gromov's theorem and the Freiman–Ruzsa theorem. If α is a non-growing subset of a group (i.e. $\left|\alpha^3\right| \leq K|\alpha|$), then the subset $\alpha^{d(K)}$ contains a subgroup H for which, in the corresponding factor group[14] the image of α can be covered by $f(K)$ translates of an appropriate *nil-progression*.[15] Perhaps the only downside of their description is that their method does not give us bounds on the size of $d(K)$ and $f(K)$. For future applications in combinatorics and number theory however it would be important to know that $d(K)$ and $f(K)$ are polynomial functions of K at least for a certain classes of groups.[16] This was precisely the goal (to obtain polynomial bounds) we aimed at with László Pyber in our paper [**131**]. We proved (see Theorem 18) that if α is a *symmetric*[17] non-growing subset in the group $SL(n, \mathbb{F})$ (over an arbitrary field $\mathbb{F}$), then α^6 contains a subgroup H for which, in the corresponding factor group[14] the image of α can be covered by $f(K)$ translates of an appropriate soluble subgroup, where $f(K)$ is a polynomial function of K whose degree and coefficients depend on n.

So far the most impressive application of the Product theorem is the so-called "Bourgain–Gamburd expansion machine". The method was developed by Bourgain and Gamburd for the construction of *expander graphs*.[18] (The expander graphs have important applications, e.g., in computer science.) Bourgain and Gamburd proved in [**12**] that, for every girth g there is an $\varepsilon > 0$ for which, all those *Cayley-graphs*[19] of the groups $SL(2, p)$ having girth larger that g are ε-expander. When the paper [**12**] was born, Helfgott's theorem was the state of art. This is why they had to limit themselves to the groups $SL(2, p)$. Later with the same method, using the Product theorem, a large number of new expander families were constructed (see for example Breuillard–Green–Tao [**26**] and [**22**], Varjú [**165**], Bourgain–Varjú [**17**]), and also Golsefidy–Varjú [**62**]). Expander graphs are used in number theory in the so-called "affine sieve methods" (see, e.g., Bourgain–Gamburd–Sarnak [**14**]). In fact, the original motivation for [**12**] came from this kind of sieve methods.

The *Sum-product theorem*[20] (Erdős–Szemerédi [**53**]) and its numerous variations (see, e.g., Tao [**159**] and the references given there) form an important chapter in additive combinatorics. Even though the Sum-product type theorems do not appear in this Dissertation, still they are strongly related to some of our themes. Elekes [**40**] has shown that the Sum-product theorem follows from the Szemerédi–Trotter theorem (see also Solymosi [**144**]). The

---

[14] More precisely, in the quotient group of the normaliser of H by H.

[15] The *nil-progressions* are the counterparts of generalised arithmetic progressions living in nilpotent groups. Often it is enough to know that the image of α in that quotient group can be covered by at most $f(K)$ translates of a nilpotent subgroup.

[16] The Product theorem can also be rewritten in a similar form (valid for the class of finite simple groups of bounded rank), and indeed, in that version the constants depend on K polynomially . A number of existing applications depend crucially on this polynomiality.

[17] A subset α of a group is *symmetric*, if for each element $a \in \alpha$ we have $a^{-1} \in \alpha$.

[18] A graph on n vertices is called *ε-expander*, if any set X of vertices of size $|X| \leq \frac{n}{2}$ is adjacent to at least $\varepsilon|X|$ further vertices outside X.

[19] The *Cayley-graph* of a group G corresponding to a generating set α has vertex-set G, and two vertices $x, y \in G$ are connected with an edge if and only if $xy^{-1} \in \alpha$.

[20] If A is a finite set of real numbers then $\max\left(|A + A|, |A \cdot A|\right) \geq c|A|^{1+\varepsilon}$.

other way around, Bourgain–Katz–Tao [**16**] started with a Sum-product type theorem, and proved a Szemerédi–Trotter type theorem. Helfgott's theorem (on the group $SL(2,p)$ was originally proved in [**72**] using a Sum-product type theorem, and even today many researchers think of the Product theorem as a kind of "non-commutative Sum-product-like theorem". Actually, the proof of the Product theorem follows a different path, but the Sum-product type theorems still have an important role in the study of non-growing subsets (see, e.g., Gill–Helfgott [**59**]). This connection works in the other direction as well. (Commutative) Sum-product type theorems can be proved using the (non-commutative) Product theorem (see, e.g., Breuillard–Green–Tao [**25**, Chapter 8]).

It is worth mentioning a recent result of Bourgain [**10**], which is in close relation with our themes. He used the above mentioned "expansion-machine" methods to prove a Szemerédi–Trotter type bound for certain hyperbolas in a finite geometry.

The structure of the Dissertation

The Dissertation is based on seven articles, and each of these articles corresponds to one chapter of the Dissertation. At the moment when I'm writing, three of the articles ([**48**], [**46**], [**125**]) have already appeared in print, one of them ([**133**]) is submitted and another one ([**61**]) is already accepted for publication, and two of them ([**47**], [**131**]) are still in preprint form.

- [**48**] and [**47**] are joint papers with György Elekes, they correspond to Chapter 1 and Chapter 5 of the Dissertation.
- [**46**] is a joint paper with György Elekes an Miklós Simonovits, it corresponds to Chapter 4 of the Dissertation.
- [**133**] and [**131**] are joint works with László Pyber, they correspond to Chapter 2 and Chapter 3 of the Dissertation.
- [**125**] is a joint work with Cheryl Praeger, László Pyber and Pablo Spiga, it corresponds to Chapter 6 of the Dissertation.
- [**61**] is a joint work with Nick Gill, László Pyber, and Ian Short, it corresponds to Chapter 7 of the Dissertation.

The chapters of the Dissertation are essentially equivalent to the original papers, but I unified references, and tried to eliminate inconsistent notations. In those cases when one chapter uses a theorem proved in another chapter, instead of just giving a reference, I preferred to fully restate the theorem in the form most appropriate for the application. Therefore the chapters are self-contained, one can read them separately. Each chapter has its own introductory section where the history and the context is explained in detail.

In addition, the rest of this Overview serves as a (somewhat informal) guide to the main results of the Dissertation. It is organised along our main themes, as follows.

1. *Bounding the number of incidences.*
   Section 1.2 belongs here, which is part of the paper [**48**].
   Our main result in this area is Theorem 6.

2. *How to find groups?*
Section 1.1 and Section 1.4 belong here.
Our main results in this area are Theorem 14, and Theorem 15.

3. *Growth in groups.*
Chapter 2 and Chapter 3 belong here.
Our main results in this area are Theorem 16, and Theorem 18.

4. *Applications in combinatorics.*
Chapter 5, Chapter 4, and Section 1.5 belong here.
Our most important results in this area are Theorem 20, Theorem 23, Theorem 26, and Corollary 24.

5. *Applications in group theory.*
Chapter 7 and Chapter 6 belong here.
Our most important results in this area are Theorem 29, Theorem 30, and Theorem 33.

1. Bounding the number of incidences

Good upper bounds on the number of incidences play a central role in combinatorial geometry, and in the theory of geometric algorithms. (Recently they have shown up is additive combinatorics as well, see [**40, 44, 42**].) The first result of this type is the Szemerédi–Trotter theorem [**154**], which was later extended by Pach and Sharir for continuous plane curves.

Theorem 1 (Pach–Sharir [**120**])**.** *Let Γ be a family of* simple[21] *continuous plane curves such that any two have at most M points in common, and there are at most s curves of Γ passing through any point in the plane (i.e. Γ has s degrees of freedom). Then the number of incidences among p points and q curves of Γ is at most*

$$C\left(p^{s/(2s-1)}q^{(2s-2)/(2s-1)} + p + q\right), \tag{1.1}$$

where the constant factor C depends only on s and M. In the special case when Γ is the family of lines in the plane, we have $s = 2$, $M = 1$, hence we obtain the original Szemerédi–Trotter theorem.

We need the following notation.

Definition 2. Let X be an arbitrary set (it could be say $\mathbb{R}^N$, the N-dimensional space), $P \subseteq X$ be a subset, and Q be a collection of subsets of X. (We think of the elements of Q as if they were "geometric shapes" in X.)

- $I(P, Q)$ denotes the *number of incidences* in the (P, Q) system, i.e. the number of pairs $(p, q) \in P \times Q$ where the point p belongs to the subset q.
- For an arbitrary point $t \in P$ we denote by $Q_t \subseteq Q$ the collection of those members of Q that contain t.

Let us consider now the configuration in $\mathbb{R}^3$ that consists of p collinear points, and q planes containing all of the p points. The number of incidences in this configuration is pq. If we are looking for a bound similar to (1.1) that is valid for configurations in $\mathbb{R}^n$, then we will need some kind of non-degeneracy assumption to avoid this type of configurations. The following

[21] A curve is *simple* if it has no self-intersection.

definition refines this idea. It allows a few sub-configurations of this type, provided that they are small enough. The parameter b and the combinatorial dimension k regulates how many and how large "bad parts" do we allow in our configuration. Later in all of our upper bounds the constant factors will depend on both b and k, but the exponents may depend on k only. (It was an interesting problem on its own right to find a non-degeneracy condition that is sufficiently "generous" to be satisfied in a large number of interesting geometric situations.)

Definition 3 (Combinatorial dimension, recursive definition)**.** We fix a constant $b > 0$. Let X be an arbitrary set, $P \subseteq X$ a subset, and Q a collection of subsets of X. We say that $\mathrm{cdim}_b(P, Q) = 0$, if $|Q| \leq b$. In general, $\mathrm{cdim}_b(P, Q) \leq k$ (for integers $k \geq 1$), if there is a subset $P' \subseteq P$ such that

- $|P \setminus P'| \leq b$, i.e. P' is "almost the whole of" P, and
- $\mathrm{cdim}_b\big(P \setminus \{t\}, Q_t\big) \leq k-1$ for all $t \in P'$.

Remark 4. One can easily check that with an appropriate choice of b, the configuration (of p points and q curves) appearing in Theorem 1 has combinatorial dimension at most 2.

It seems rather hopeless to calculate the combinatorial dimension of a configuration directly from Definition 3. The next lemma (which is Lemma 1.2.13) shows that in configurations coming from geometry, the combinatorial dimension generally agrees with the geometric dimension.

Lemma 5. *Let A be a k-dimensional variety, let $\mathcal{H}$ denote the collection of all subvarieties of degree at most d. Then there is a value b depending on k and d only such that for arbitrary finite subset $P \subseteq A$* in general position[22] *we have* $\mathrm{cdim}_b(P, \mathcal{H}) \leq k$.

The next theorem is essentially Theorem 1.2.5 and a simplified version of Theorem 1.2.6 forged together.

Theorem 6. *Let $\mathcal{P}$ be a finite point-set in the N-dimensional complex projective space $\mathbb{CP}^N$, and $\mathcal{V}$ a finite collection of algebraic varieties (in the same projective space). Suppose that the combinatorial dimension of the $(\mathcal{P}, \mathcal{V})$ configuration is $\mathrm{cdim}_b(\mathcal{P}, \mathcal{V}) = k \geq 2$, and all members of $\mathcal{V}$ have degree at most d. Then there are constants α and β depending on k, N, and d only such that*

$$0 < \alpha, \beta < 1\,, \qquad k\alpha + \beta = k$$

and the number of incidences in the $(\mathcal{P}, \mathcal{V})$ configuration is

$$I(\mathcal{P}, \mathcal{V}) \leq C\Big(|\mathcal{P}|^\alpha |\mathcal{V}|^\beta + |\mathcal{P}| + |\mathcal{V}| \log(2|\mathcal{P}|)\Big),$$

where the constant C depends on the parameters N, b, k, d. In the special case when $\mathcal{V}$ consists of hyperplanes (i.e. $d = 1$), with any chosen value $0 < \varepsilon < \frac{k-1}{k(Nk-1)}$, one may use the following explicit values.

$$\alpha = \frac{N(k-1)}{Nk-1} - \varepsilon\,, \qquad \beta = \frac{k(N-1)}{Nk-1} + k\varepsilon\,.$$

[22] Here we say that P is in general position if each proper subvariety of degree at most d^k contains at most b points from P.

Remark 7. Theorem 6 was formulated in projective spaces in order to be able to talk about the *degree* of an algebraic variety. Of course an analogue statement holds for algebraic subsets in the affine space $\mathbb{C}^n$, but the lack of a standard notion of degree makes it more cumbersome to formulate precisely what the constant C depends on.

It is worth comparing Theorem 6 and the Pach-Sharir theorem (see Theorem 1). Our result is more general in the sense that instead of plane curves we study higher dimensional varieties, and our result is valid in complex geometry as well. On the other hand, this generality has a price to pay (at the moment). The Pach–Sharir theorem allows arbitrary continuous curves, and the exponents in the upper bound are optimal, while our result deals with algebraic varieties only, and the exponents we obtain aren't optimal at all. (Note that in Theorem 1.2.5 and Theorem 1.2.6 the exponents α and β are explicitly given.)

2. How to find groups?

Now we summarise the main results of Section 1.1 and Section 1.4. There is a very general principle hiding in the background. If in a geometric situation we find a lot of unexpected coincidences then we should expect to discover a large group of symmetries. One of the most-known results in this direction is Hrushovski's Group Configuration Theorem in [**75**] (see also [**122**]).

Here we shall study a geometric–combinatorial situation. Below we define when an algebraic surface $V \subseteq \mathbb{C}^3$ is called *rich* (i.e. when are there too many coincidences on it). After introducing a few simple examples we will see, that the rich surfaces have a very special structure. If the surface V is rich, then either it is a cylinder built on a plane curve (see Example 13), or there is an algebraic group behind the scene, and V can be constructed from this group as in Example 12. The special case of this result, when the surface is given by an equation of the form $z = f(x, y)$, was obtained by György Elekes and Lajos Rónyai in their paper [**45**]. The extension to arbitrary algebraic surfaces as well as the higher dimensional generalisation (see Theorem 14 and Theorem 15) are joint results of György Elekes and myself (see [**48**]).

Definition 8 (Richness)**.**

(a) An algebraic surface $V \subset \mathbb{C}^3$ is said to be *rich*, if for infinitely many values of n there are n-element subsets $X, Y, Z \subset \mathbb{C}$ such that

$$\big|V \cap (X \times Y \times Z)\big| \geq \mathcal{C}n^2$$

with some constant $\mathcal{C} > 0$ independent of n.

(b) Let A, B, C be n-dimensional complex varieties ($m \geq 1$). A $2m$-dimensional subvariety $V \subset A \times B \times C$ is said to be *rich*, if for infinitely many values of n there are n-element subsets $X \subset A$, $Y \subset B$, $Z \subset C$ in "general position" (see Definition 1.2.12) satisfying the same bound

$$\big|V \cap (X \times Y \times Z)\big| \geq \mathcal{C}n^2$$

with some constant $\mathcal{C} > 0$ independent of n.

An *algebraic group* is a group whose underlying set is an algebraic variety and the group operation is given by polynomials.[23] The algebraic groups are studied intensively for a long time, one a great deal is known about their internal structure. Let us see some examples.

Example 9. Each one-dimensional complex algebraic group belongs to one of the following three types. There is a single group only that belongs to the first type, and a single group that belongs to the second type. On the other hand, there are infinitely many different groups that belong to the third type (which are all isomorphic to each other as topological groups).

(a) $\mathbb{C}$ — the additive group of the complex numbers.
(b) $\mathbb{C}^*$ — the multiplicative group of the non-zero complex numbers. The function $z \to e^{2\pi i z}$ shows that $\mathbb{C}^*$ is isomorphic to the factor group $\mathbb{C}/\mathbb{Z}$.
(c) Elliptic curves — these are the plane curves given by the equations $y^2 = x^3+ax+b$ (extended with a single point at infinity), where $4a^3+27b^2 \neq 0$. Each elliptic curve can also be written in the form of a quotient group $\mathbb{C}/\mathbb{L}$ where $\mathbb{L}$ is a parallelogram lattice containing the origin. (Non-congruent lattices result in different quotient groups.)

Example 10. Let us consider the "square root function". Strictly speaking the square root isn't really a function. Around each complex number $x_0 \neq 0$ it has two "continuous branches", and globally the situation gets even more complicated. If we consider the square roots of all non-zero complex numbers at the same time, the two branches get "mixed up". As we move continuously the value of x (in the complex plane) along a circle around 0, the two square roots change also continuously, but they get swapped as they return to their initial position. "Functions" analogous to the "square root function" we call *multi-valued algebraic functions*.

Definition 11. Let A and B be sets. A function F that assigns to each element of A a subset of B is called a *multi-valued function from A into B*.

(a) the *graph of F* is the following subset.
$$\Gamma_F = \Big\{ (a,c) \;\Big|\; a \in A,\, c \in F(a) \Big\} \subseteq A \times B\,.$$
(b) For all subsets $H \subseteq A$ and all points $b \in B$ we define
$$F(H) = \cup_{h\in H} F(h) \subseteq B\,, \qquad F^{-1}(b) = \big\{ a \in A \;\big|\; F(a) \ni b \big\}\,.$$
Clearly F^{-1} is a multi-valued function from B into A. If both $\max_{a\in A} \big|F(a)\big|$ and $\max_{b\in B} \big|F^{-1}(b)\big|$ are finite then let $\deg(F)$ be the larger of the two, otherwise we set $\deg(F) = \infty$.
(c) Now let A and B be algebraic curves. We say that the multi-valued function F is *algebraic* if its graph Γ_F is an algebraic curve on the surface $A \times B$ and $\deg(F) < \infty$.
(d) Assume now that A and B are m-dimensional varieties. We say that the multi-valued function F is *algebraic* if the closure of its graph Γ_F is an m-dimensional subvariety of the $2m$-dimensional $A\times B$ and $\deg(F) < \infty$.

[23] More precisely, the group can be covered by open dense subsets $\{U_i\}$ such that the multiplication map $(x,y) \to xy$ is a polynomial function on each $U_i \times U_j$ and the inverse-element map $x \to x^{-1}$ is a polynomial function on each U_i.

Example 12. Let $\mathcal{G}$ be a complex algebraic group. First we concentrate on the one-dimensional case, and with the help of the group operation we build a rich algebraic surface in $\mathbb{C}^3$. Afterwards we extend the construction to higher dimensional groups.

(a) At the moment let the group $\mathcal{G}$ still be arbitrary. Let $n = 2k+1$ an odd natural number. Consider the (algebraic) variety

$$\mathcal{G}_{sp} := \left\{ (x, y, z) \in \mathcal{G}^3 \;\middle|\; xyz = 1 \text{ in the group } \mathcal{G} \right\},$$

that we call the *special subvariety in* $\mathcal{G}^3$,or, for one-dimensional $\mathcal{G}$, we call it the *special surface in* $\mathcal{G}^3$. Choose an element $a \in \mathcal{G}$ of infinite order (such element exists whenever $\dim(\mathcal{G}) \geq 1$), and set

$$X = Y = Z := \{a^{-k}, a^{-(k-1)}, \ldots, a^{-1}, 1, a, \ldots, a^{(k-1)}, a^k\} .$$

It is easy to check that $\mathcal{G}_{sp}$ really contains at least $\lceil k^2/2 \rceil \geq \frac{1}{4}n^2$ points from the subset $X \times Y \times Z$, hence it is rich.

(b) Assume now that $\mathcal{G}$ is one-dimensional and let f, g, h be multi-valued algebraic functions from $\mathcal{G}$ into $\mathbb{C}$. Their direct product $F = f \times g \times h$ is also a multi-valued function from $\mathcal{G}^3$ into $\mathbb{C}^3$, and $\deg(F) = \deg(f)\deg(g)\deg(h)$. Consider the F-image of the special surface $\mathcal{G}_{sp}$, the subset $F(\mathcal{G}_{sp}) \subset \mathbb{C}^3$, its closure is an algebraic surface $V \subseteq \mathbb{C}^3$. Clearly the surface V contains at least $\lceil \frac{k^2}{2\deg(F)} \rceil = \mathcal{C}n^2$ points from the subset $F(X \times Y \times Z) = f(X) \times g(Y) \times h(Z)$, hence it is rich.

(c) The variety V has at most $\deg(V)$ irreducible components, hence some of the irreducible components must be rich.

(d) Consider now the general case, $\dim(\mathcal{G}) = m \geq 1$ is arbitrary. Let f, g, h be multi-valued functions from $\mathcal{G}$ into three m-dimensional varieties, A, B and C. The above argument can be repeated in this situation as well. $F = f \times g \times h$ is a multi-valued algebraic function from $\mathcal{G}^3$ into $A \times B \times C$ (both $\mathcal{G}^3$ and the product $A \times B \times C$ are $3m$-dimensional varieties). The closure of the subset $F(\mathcal{G}_{sp})$ is a $2m$-dimensional subvariety $V \subset A \times B \times C$, and in certain cases (for example when G is abelian) its irreducible components $V_0 \subseteq V$ are rich. It turns out that these are the "prototype" of rich subvarieties.

Example 13 (Cylinders)**.**

(a) An algebraic surface $V_0 \subset \mathbb{C}^3$ is called a *cylinder* if its equation depends on two variables only i.e. $F(x, y) = 0$, $F(x, z) = 0$ or $F(y, z) = 0$. Consider now the case $F(x, y) = 0$, and choose two sets $X = \{x_1, x_2, \ldots, x_n\}$, $Y = \{y_1, y_2, \ldots, y_n\}$ of complex numbers such that $F(x_i, y_i) = 0$ for all i. One can easily see that for arbitrary n-element subsets $Z \subset \mathbb{C}$ of numbers we have $\left|V_0 \cap (X \times Y \times Z)\right| \geq n^2$, hence V is rich.

(b) Let A, B, C be m-dimensional varieties. We say that a $2m$-dimensional subvariety $V \subset A \times B \times C$ *contains a cylinder* if the image of one of the projections $V \to A \times B$, $V \to B \times C$, or $V \to A \times C$ has dimension smaller than $2m$. Indeed, one can see that such a V contains a cylinder.

The following theorem is a simplified version of Theorem 1.1.3.

Theorem 14 (Rich surfaces in $\mathbb{C}^3$)**.** *Let $V \subset \mathbb{C}^3$ be an algebraic surface of degree d. Then there are constants η and n_0 depending on d only such that the following properties are equivalent.*

(a) For some value $n \geq n_0$ there are n-element subsets $X, Y, Z \subset \mathbb{C}$ such that

$$\big|V \cap (X \times Y \times Z)\big| \geq n^{2-\eta} .$$

(b) V has an irreducible component V_0 which is either a cylinder (see Example 13), or it is one of the V_0 constructed in Example 12 (based on some one-dimensional complex algebraic group). In the latter case the degrees of the multi-valued functions needed in the construction can be bounded in terms of d.

(c) Let $\mathbb{D} \subset \mathbb{C}$ denote the unit disc. Either V contains a cylinder (see Example 13), or there are one-to-one analytic functions $f, g, h : \mathbb{D} \to \mathbb{C}$ with analytic inverses such that

$$V \supseteq \Big\{ \big(f(x), g(y), h(z)\big) \in \mathbb{C}^3 \;\big|\; x, y, z \in \mathbb{D},\ x + y + z = 0 \Big\} .$$

(d) V has an irreducible component V_0 such that all open subsets of V_0 are rich.

A (small) positive constant η appears in the theorem . We did not specify any explicit value for η, since we think that our present bounds for the exponents are far from being optimal. In fact it is still possible that the theorem holds for arbitrary value $0 < \eta < 1$ — see Problem 1.1.4.

The following theorem is a simplified version of Theorem 1.4.2.

Theorem 15 (Rich subvarieties in higher dimension)**.** *For all positive integers m there is a positive real number η with the following property. Let A, B, C be m-dimensional projective varieties, and $V \subset A \times B \times C$ such a $2m$-dimensional subvariety that does not contain a cylinder (see Example 13). The following properties are equivalent.*

(a) For some "sufficiently large" value n there are n-element subsets $X \subset A$, $Y \subset B$, $Z \subset C$ of "general type" such that

$$\big|V \cap (X \times Y \times Z)\big| \geq n^{2-\eta} .$$

(b) V has an irreducible component V_0 which is one of the V_0 constructed in Example 12 (based on some m-dimensional complex algebraic group).

(c) V has an irreducible component V_0 such that all open subsets of V_0 are rich.

For the precise meaning of the phrases "sufficiently large", and "general position", which appear in (a) above, consult Theorem 1.4.2, and Definition 1.2.12.

3. Growth in groups

We are given a finite set α of $n \times n$ matrices. We shall study for which sets α will α^3 be much larger than α and when will they have comparable size.

What happens in algebraic groups? In Chapter 2 we study the structure of algebraic groups. Using algebraic geometry and group theoretic methods we succeeded in proving two theorems about non-growing subsets. In the later chapters this two theorems plays a central role in the study of growth.

Theorem 2.1.4 describes growth properties of subsets in finite simple groups. Its importance is indicated by the fact that this result has a long list of authors: Breuillard–Green–Tao [**24**] and Pyber–Szabó [**132**].

Theorem 16 (Product theorem)**.** *Let G be a simple subgroup of the group $SL(n,q)$ (for some prime power q),*[24] *and α a system of generators in G. Then either $\alpha^3 = G$, or else we have*

$$\left|\alpha^3\right| \geq |\alpha|^{1+\varepsilon}$$

where $\varepsilon > 0$ depends on n only.

Corollary 2.13.4 talks about arbitrary (not necessarily finite) linear groups. Here we state a simplified version.

Theorem 17. *Let $\mathbb{F}$ be an arbitrary field, $K \geq 1$ a real number, and α a finite subset in the group $SL(n,\mathbb{F})$ such that*

$$\left|\alpha^3\right| \leq K|\alpha| \ .$$

Then there is a constant $m = m(n)$, and a virtually soluble[25] *subgroup Δ such that α can be covered by at most K^m cosets of Δ.*

What do we gain from Group theory? In Chapter 3 we combine Theorem 17 with group theoretic methods and with Theorem 16. We obtain a much more precise picture about the structure of non-growing sets of matrices. Theorem 3.6.13, which contains the Product theorem as a special case, is the following.

Theorem 18. *Let $\mathbb{F}$ be an arbitrary field, $K \geq 1$ a real number, and α a finite subset in $SL(n,\mathbb{F})$ such that for each element $a \in \alpha$ we have also $a^{-1} \in \alpha$, and*

$$\left|\alpha^3\right| \leq K|\alpha| \ .$$

Then the subgroup generated by α has two normal subgroups $P \leq \Gamma$ such that α^3 contains a coset of P, Γ/P is soluble, and α can be covered by at most $K^{c(n)}$ cosets of Γ, where $c(n)$ depends on n only.

4. Applications in Combinatorics

We are given n non-degenerate conics in the (real or complex) plane, no three of them are tangent to each other at the same point. Hirzebruch [**74**] asked if there is an upper bound of the form $Cn^{2-\varepsilon}$ on the number of tangencies among them. With Gábor Megyesi we answered the question positively in [**110**]. However, our bounds can be improved significantly with the help of Theorem 6. In Corollary 1.5.1 we prove the following.

[24] Each finite simple group can be embedded into some of the $SL(n,q)$, where n is very close to the rank of the group.

[25] A group is *virtually soluble* if it has a soluble subgroup of finite index.

Corollary 19. *In the above configuration of n conics the number of tangencies is at most $Cn^{\frac{139}{79}}$.*

We are given three centres in the (real) plane, and around each of them a concentric family of n circles — a so called *circle grid* We call a point P a *triple point* if each of the three circle families has a member passing through P. Erdős, Lovász and Vesztergombi [**51**] asked the following question. For which centre configurations is it possible for infinitely many values of n to choose the circles so that there be at least cn^2 triple points? György Elekes [**39**] have found such examples. On the other hand, in Theorem 1.5.3 we show that for most centre configurations there are no such families of circles.

Theorem 20. *There is an absolute constant $\eta > 0$ and a bound $n_1 \in \mathbb{Z}$ with the following property. If $n > n_1$ and three families of concentric circles as above have at least $n^{2-\eta}$ triple points then the three centres of the families are collinear.*

These are the basic ideas of the proof. First we reformulate the problem. Three circles meet in a common point if and only if their radii satisfy a certain polynomial equation. So we have to decide whether the zero locus of this equation, which is an algebraic surface $V \subset \mathbb{R}^3$, is rich or not. This surface V is rich if and only if its equation satisfies a certain partial differential equation constructed with the help of Theorem 14. Finally, checking that partial differential equation is a matter of some algebraic juggling.

Instead of circles we may study more general continuous curves. In place of the n concentric circles we select n members from a "continuously parametrised family of continuous curves". For simplicity here we restrict ourselves to families of curves that can be defined as level-curves of a polynomial function — an "algebraic family of curves" can always be written as a union of such families.

Definition 21. Let $G \subseteq \mathbb{R}^2$ be an open subset in the plane, $\overline{G}$ denote its closure.

(a) An *algebraic family of curves* in $\overline{G}$ is a collection $\Gamma = \{\gamma^{(t)} \subset \overline{G} \ : \ t \in [0,1]\}$ of continuous curves which can be defined via a 3-variable polynomial p as follows.

$$\gamma^{(t)} = \left\{(x,y) \in \overline{G} \;\middle|\; p(x,y,t) = 0\right\} .$$

Note that the polynomial p is not unique.[26] The *degree* of the family Γ is the smallest possible value of $\deg(p)$.

(b) The family Γ is *explicitly parametrised*, if there is a single member of Γ passing through each point of $\overline{G}$, and the implicit function defined by the equation $p\big(x,y,f(x,y)\big) = 0$ is analytic in G and continuous on $\overline{G}$.

(c) A continuous curve $\mathcal{E} \subset \overline{G}$ is an *envelope* for the family Γ if it has a tangent line at each of its points, it has no common arc with any member of the family Γ, and for each point $P \in \mathcal{E}$ there is a member $\gamma^{(t)} \in \Gamma$ that is tangent to $\mathcal{E}$ at the point P.

[26] For example all powers of p define the same family Γ.

Let $\alpha^{(r)}$, $\beta^{(s)}$, $\gamma^{(t)}$ be algebraic families of curves in the plane. The loci in $\mathbb{R}^3$ of the triples (r_0, s_0, t_0) for which the curves $\alpha^{(r_0)}$, $\beta^{(s_0)}$, $\gamma^{(t_0)}$ pass through a common point is an algebraic surface $V \subset \mathbb{R}^3$.

Definition 22. We choose n curves from each of the three families. A *triple point* of this configuration is a point P in the plane such that each of the three families have a chosen curve passing through P.

If the surface V is not rich (this is the typical situation) then the argument after Theorem 20 implies that there are at most $n^{2-\eta}$ triple points. We cannot decide in full generality whether V is rich or not, but with the help of Theorem 14 we may get useful geometric criteria. The following theorem is a simplified version of Theorem 4.4.1.

Theorem 23. *Let $G \subset H$ be open subsets in the plane, Γ_1, Γ_2 explicitly parametrised algebraic families of curves in $\overline{H}$, and Γ_3 an explicitly parametrised algebraic family of curves of in $\overline{G}$, Let d denote the largest among the degrees of the three families. Assume that Γ_3 has an envelope $\mathcal{E}$ which belongs to H, and $\mathcal{E}$ has no common arc with any member of the the other two families. If we pick n curves from each families (n is sufficiently large) then this configuration has at most $n^{2-\eta(d)}$ triple points in G, where the constant $\eta(d) > 0$ depends on d only.*

An immediate corollary is Theorem 4.5.1, which had been conjectured earlier by László Székely (see [**42**, Conjecture 3.41]).

Corollary 24. *We are given three points in the plane. We draw n unit circles through each of them. If n is large enough then this configutation has at most $n^{2-\eta}$ triple points. Here $\eta > 0$ is an absolute constant.*

Finally we discuss the following variation on a classical theme (the so-called Orchard problem, see [**83, 152**]).

Problem 25. Fix a constant $C > 0$, and find all such subsets $\mathcal{H}$ of n points in the plane for which there are at least Cn^2 lines intersecting $\mathcal{H}$ in three or more points.

Our general philosophy suggests that the subset $\mathcal{H}$ should be closely related to some kind of "symmetry group", and indeed, all known constructions can be described using groups. However, at the moment we cannot yet find the group in this generality. In Theorem 5.2.2 we solve the problem in a special case that can be handled using algebraic geometry.

Theorem 26. *Let $\mathcal{H}$ be a finite set of points in the plane such that there are at least $c|\mathcal{H}|^2$ lines passing through three or more points of $\mathcal{H}$. Assume that an algebraic curve of degree at most d contains all points of $\mathcal{H}$. If $|\mathcal{H}|$ is large enough (in terms of c and d) then the curve must have degree three.*

It is worth noting here that Green–Tao [**66**] have given sharp upper bound, valid in full generality, on the size of $\mathcal{H}$.

5. Applications in group theory

Conjecture of Babai. Babai [**7**] conjectured, that all Cayley graphs of all non-abelian finite simple groups L have diameter at most $C\big(\log|L|\big)^c$,

where c and C are absolute constants (see Conjecture 2.1.1). The Product theorem (see Theorem 16) implies immediately that the conjecture of Babai holds for finite simple groups of bounded rank.

Corollary 27. *If L is a non-abelian finite simple group of bounded rank,*[27] *and α is a symmetric generating set in L, then the Cayley graph $\Gamma(L,\alpha)$ has diameter at most $C\big(\log|L|\big)^c$, where c and C are absolute constants.*

Product decompositions. Let α be a subset of some group. The *conjugates of α* are the subsets of the form

$$g^{-1}\alpha g = \big\{g^{-1}ag \;\big|\; a\in\alpha\big\}$$

where g is an arbitrary element of the group. The starting point of Chapter 7 is the following conjecture of Liebeck, Nikolov and Shalev [**99**].

Conjecture 28. *Let G be a non-abelian finite simple group and $\alpha\subseteq G$ a subset of at least two elements. Then G can be written as the product of at most $c\log|G|/\log|\alpha|$ conjugates of α, where c is a universal constant.*

Note that this bound (if true) is optimal, since the number of, say, $\frac{1}{2}\log|G|/\log|\alpha|$ term products of elements from α cannot be more than $\sqrt{|G|}$. Conjecture 28 is the extension of a deep (and useful) theorem of Liebeck and Shalev [**104**]. They have shown that Conjecture 28 holds in the case when α is a conjugacy class.

If we bound the rank of the group G that appears in Conjecture 28 then, combining Theorem 16 with a surprising combinatorial argument, we may handle arbitrary subsets α. This is the content of Theorem 7.1.3.

Theorem 29. *Let G be a non-abelian finite simple group of rank*[27] *r and $\alpha\subseteq G$ a subset of at least two elements. Then G can be written as the product of at most $c(r)\log|G|/\log|\alpha|$ conjugates of α, where the constant $c(r)$ depends on r only.*

Theorem 7.1.4 transforms this result into a theorem on growth.

Theorem 30. *Let G be a non-abelian finite simple group of rank*[27] *r and $\alpha\subseteq G$ an arbitrary subset. Then either there is a conjugate α' of the set α such that $|\alpha\alpha'|\geq|\alpha|^{1+\varepsilon}$, where $\varepsilon(r)>0$ is a constant depending on r only, or else $\alpha^3=G$.*

By analogy we transform Conjecture 28 into a conjecture about growth. In Section 7.6 and Section 7.4 we generalise the classical Plünecke–Ruzsa type inequalities for arbitrary (not necessarily commutative) groups, and with the help of these new inequalities we prove that the original Conjecture 28 implies our new conjecture concerning growth.

Conjecture 31. *There is a real constant $\varepsilon>0$ and an integer constant $b>0$ with the following property. In each finite simple group G for all subsets α either there is a conjugate α' such that $|\alpha\alpha'|\geq|\alpha|^{1+\varepsilon}$, or G is equal to the product of b conjugates of α.*

It is possible that Conjecture 31 holds with $b=3$. On the other hand, there are counterexamples to $b=2$.

[27]The *rank* of a finite simple group G is roughly equal to the smallest value n such that G is isomorphic to a subgroup of $SL(n,q)$ for some prime power q.

Permutation groups. A graph Γ is said to be *G-vertex-transitive* if G is a subgroup of $\mathrm{Aut}(\Gamma)$ acting transitively on the vertex set of Γ. We say that a G-vertex-transitive graph Γ is *G-locally primitive* if the stabiliser G_α of the vertex α induces a primitive permutation group on the set of vertices adjacent to α. (Because of the transitivity this holds either for all vertices, or for none of them.) In 1978 *Richard Weiss* [**170**] *conjectured* that, for a finite connected G-vertex-transitive, G-locally primitive graph Γ, the size of G_α is bounded above by some function depending only on the valency of α. (By the transitivity, all vertices have equal valencies, and the stabiliser subgroups are all isomorphic to each other.)

In Chapter 6 we study the Weiss conjecture. The reduction theorems in [**129, 126**] show that it is enough to bound the size of the H_α stabiliser subgroups in certain H-vertex-transitive graphs, where H is a composition factor of the group G. As H is a simple group, we may use the Product theorem (see Theorem 16) for studying the H-vertex-transitive graphs. With this method we are able to deduce the Weiss conjecture in the case when the composition factors of G have bounded rank.

Definition 32. Define $\mathrm{BCP}(r)$ to be the class of finite groups G which have no section H/K, where $K < H \leq G$ and K is normal in H, that is isomorphic to the alternating group $\mathrm{Alt}(r+1)$.[28]

The class of $\mathrm{BCP}(r)$-groups was first considered by Babai, Cameron and Pálfy [**4**]. They showed that primitive $\mathrm{BCP}(r)$-groups of degree n have order at most $n^{f(r)}$. This result is an essential ingredient of many polynomial time algorithms for permutation groups related to the graph isomorphism problem [**86**]. The $\mathrm{BCP}(r)$-groups play also a very important role in the theory of subgroup growth of residually finite groups (see [**107**]).

Theorem 6.1.2 states that the Weiss conjecture holds in the class of $\mathrm{BCP}(r)$ groups. The full Weiss conjecture asks then whether the function g below can be chosen not to depend on r.

Theorem 33. *There exists a function $g : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ such that, for Γ a connected G-vertex-transitive, G-locally primitive graph of valency at most d, if G is a $\mathrm{BCP}(r)$-group, then a vertex stabiliser in G has size at most $g(r,d)$.*

[28] It is easy to see that all composition factors of a $\mathrm{BCP}(r)$-group must be either a sporadic simple group, or a finite simple group of rank at most r.

dc_650_12

CHAPTER 1

How to find groups

1.1. Introduction

This chapter is essentially equivalent to our joint paper [**48**] with György Elekes. The germs of the paper were two earlier manuscripts: "How to find groups?" by myself and our joint work "Triple points of circle grids". They have been circulated as sort of "technical reports" for several years. We decided to publish the method based upon the two of them as one article, since it is the interaction of the two points of view that makes the ideas work.

The philosophy of our main results (and also of their applications) is a general principle of geometry: whenever we find a lot of unexpected coincidences, then somewhere in the background there lurks a large group of symmetries. There are infinitely many variations on this theme, both continuous and discrete, and we shall only touch a few of them. We focus on algebraic geometry, with applications to Erdős geometry. To state precise results, we have to measure the amount of coincidences a certain geometric configuration has. In the discrete case we can simply count them while in the continuous case we measure the dimension of the parameter space instead.

As for the discrete versions, we shall usually consider finite Cartesian products $X \times Y \times Z = \{(x, y, z) \mid x \in X, y \in Y, z \in Z\}$, where, in the simplest case, $X, Y, Z \subset \mathbb{C}$, or in a more general setting, for some varieties A, B, C, we have $X \subset A$, $Y \subset B$, $Z \subset C$, and thus $X \times Y \times Z \subset A \times B \times C$. (In what follows, n will denote a large positive integer, usually $n = |X| = |Y| = |Z|$. Moreover, there also appear some constants like $c > 0$ or natural numbers d, k, r which remain fixed while $n \to \infty$.)

Geometric questions which involve Euclidean distances often lead to polynomial relations of type $F(x, y, z) = 0$ for some $F \in \mathbb{R}[x, y, z]$. Several problems of Combinatorial Geometry can be reduced to studying such polynomials which have many zeroes on $n \times n \times n$ Cartesian products. The special case when the relation $F = 0$ can be re–written as $z = f(x, y)$, for a polynomial or rational function $f \in \mathbb{R}(x, y)$, was considered in [**45**]. Our main goal is to extend the results found there to full generality (and also to show some geometric applications, e.g. one on "circle grids").

The main result of Chapter 1 concerns low–degree algebraic sets F which contain "too many" points of a (large) $n \times n \times n$ Cartesian product. Then we can conclude that, in a neighborhood of almost any point, the set F must have a very special (and very simple) form. Roughly speaking, then either F is a cylinder over some curve, or we find a group behind the scene: F must

be the image of the graph of the multiplication function of an appropriate algebraic group (see Theorem 1.1.3 for the 3D special case and Theorem 1.4.2 in full generality).

The structure of Chapter 1. We first state Theorem 1.1.3, the three dimensional special case of our Main Theorem 1.4.2. Its proof – as well as its arbitrary dimensional version — can be found in Section 1.4. It relies upon two basic tools: incidence bounds and composition sets. The former are described in Section 1.2 while the latter are the subject of Section 1.3. Moreover, in Section 1.5, we give an immediate consequence of our incidence bounds which concerns a problem posed by Hirzebruch and was partially solved in [**110**]. Also an application of our three dimensional Theorem 1.1.3 can be found there.

The main result in $\mathbb{C}^3$ (and $\mathbb{R}^3$). In [**45**] those bivariate polynomials $F \in \mathbb{R}[x, y]$ were characterized whose graph (in $\mathbb{R}^3$) passes through at least cn^2 points of an $n \times n \times n$ Cartesian product $X \times Y \times Z \subset \mathbb{R}^3$, where $n = |X| = |Y| = |Z|$. It was shown there that F must be very special, provided that $n > n_0 = n_0(c, \deg(F))$. More precisely,

$$F(x, y) = \begin{cases} f\big(g(x) + h(y)\big); & \text{or} \\ f\big(g(x) \cdot h(y)\big), \end{cases}$$

and these types of polynomials really have graphs which are incident upon many points of appropriately chosen Cartesian products, e.g., if both $g(X)$ and $h(Y)$ are arithmetic/geometric progressions. (The reader may have observed the additive group $\langle \mathbb{R}, + \rangle$ and the multiplicative group $\langle \mathbb{R} \setminus \{0\}, \cdot \rangle$ in the background.)

We generalise the foregoing result several ways:

(a) instead of real variables, we consider complex ones;
(b) instead of graphs of bivariate polynomials, we allow algebraic varieties (surfaces) in $\mathbb{C}^3$;
(c) instead of cn^2 points, we only require that the surface in question passes through as few as $n^{2-\eta}$ points of a Cartesian product, for a sufficiently small positive η.

To state the Theorem in its simplest (lowest interesting dimensional) 3D form, we recall the notion of "connected one dimensional algebraic groups". A good reference for the list below: Excersise 11, 12, 13 in Chapter 1 §2 of [**117**]. In this case, the complex analytic structure completely determines the algebraic structure, so we describe these groups as analytic manifolds. The following three types of groups are called *complex connected one dimensional algebraic groups:*

(a) $\langle \mathbb{C}, + \rangle$;
(b) $\langle \mathbb{C} \setminus \{0\}, \cdot \rangle \cong \langle \mathbb{C}, + \rangle / \mathbb{Z}$;
(c) $\langle \mathbb{C}, + \rangle / \mathbb{L}$, where $\mathbb{L}$ is a parallelogram lattice (an affine image of $\mathbb{Z}^2$). Algebraically these occur e.g. as the usual groups on cubic curves in the plane.

The irreducible *real* one dimensional algebraic groups are appropriate subgroups of those above. Analytically they are all isomorphic to the real line

$\langle \mathbb{R}, + \rangle$, to the unit circle $\langle S^1, \cdot \rangle$ in the complex plane, or two copies of the unit circle $\mathbb{Z}_2 \oplus \langle S^1, \cdot \rangle$. However, in contrast to the complex case, several nonequivalent algebraic structures correspond to the same analytic group.

Example 1.1.1. If $\langle \mathcal{G}, \oplus \rangle$ is any of the foregoing — real or complex — algebraic groups (or even if it is a higher dimensional one) then it is easy to show examples of $n \times n \times n$ Cartesian products $X \times Y \times Z$ in $\mathcal{G}^3$ or in $\mathbb{C}^3$, and two dimensional subvarieties (surfaces) which contain $\approx n^2/8$ points of $X \times Y \times Z$, as follows.

(a) Without loss of generality, assume that n is odd, say $n = 2k + 1$, and pick an arbitrary non-torsion element $a \in \mathcal{G}$ (i.e. one of infinite order). Let

$$X = Y = Z := \{-ka, -(k-1)a, \ldots, -a, 0, a, \ldots, (k-1)a, ka\}$$

and define

$$\mathcal{G}_{sp} := \big\{(x, y, z) \;\big|\; x \oplus y \oplus z = 0 \in \mathcal{G}\big\},$$

which we call the *special subvariety in* $\mathcal{G}^3$. (Of course, in higher dimensional — usually non–Abelian — groups the multiplicative notation would be more appropriate.) It is easy to check that this $\mathcal{G}_{sp}$ will, indeed, contain $\geq \lceil k^2/2 \rceil \approx n^2/8$ points of $X \times Y \times Z$. Moreover, if $U \subset \mathcal{G}$ is any neighborhood of 0, then we can choose $X = Y = Z \subset U$ via choosing an a sufficiently close to 0.

(b) In (a) we have found a Cartesian product set in any neighborhood of $(0, 0, 0) \in \mathcal{G}^3$. We can improve on this: there are similar Cartesian product sets in any neighborhood of any point $(a, b, c) \in \mathcal{G}_{sp}$. Indeed, if $a \oplus b \oplus c = 0$ then we may define $X' = X \oplus a$, $Y' = b \oplus c \oplus Y \oplus a \oplus b$ and $Z' = c \oplus Z$ (these formulas work even if $\mathcal{G}$ is noncommutative). Again, $\mathcal{G}_{sp}$ contains a quadratic order of magnitude of points of the Cartesian product $X' \times Y' \times Z'$, but this Cartesian product lives in the neighborhood of (a, b, c).

(c) More generally, suppose we have a connected open set of the form $U = U_f \times U_g \times U_h \subseteq \mathcal{G}^3$ intersecting the subvariety $\mathcal{G}_{sp}$ of (a), nonconstant analytic functions $f : U_f \to \mathbb{C}$, $g : U_g \to \mathbb{C}$, $h : U_h \to \mathbb{C}$, and a surface $V \subset \mathbb{C}^3$ containing the $f \times g \times h$-image of $\mathcal{G}_{sp} \cap U$:

$$V \supseteq \Big\{ \big(f(x), g(y), h(z)\big) \in \mathbb{C}^3 \;\big|\; (x, y, z) \in U,\ x \oplus y \oplus z = 0 \in \mathcal{G} \Big\}.$$

We may assume, that the functions f, g, h are one-to-one, otherwise we replace their domain with appropriate subsets. Then we choose a Cartesian product $X' \times Y' \times Z' \subset U$ as in (b). Then V contains a quadratic order of magnitude of points of the Cartesian product $f(X') \times g(Y') \times h(Z')$.

Definition 1.1.2. Let U, V be open subsets in $\mathbb{C}$ or in a connected one-dimensional algebraic group $\mathcal{G}$. A multi-valued function $f : U \to V$ is an *analytic multi–function* if, except for a finite point set $H \subset U$, every $P \in U \setminus H$ has a neighborhood where f is the union of finitely many one-to-one analytic functions, called the *analytic branches* of f near P, and f has no values at points of H. The *complexity* of such a function is the larger

of $|H|$ and the maximum number of its branches. We note, that if U is connected, then the number of branches is the same everywhere.

The following is the three dimensional version of our main result. It asserts, among others, that if a variety contains an "almost–quadratic" number of points of an $n \times n \times n$ Cartesian product then it must look like those in Example 1.1.1.

It also involves a (rather small) positive constant η. We refrain from computing an explicit value since we believe that it is far from best possible. Actually, we cannot even exclude the possibility that the result holds for every $\eta < 1$ — see Problem 1.1.4 below.

Theorem 1.1.3 (Surface theorem)**.** *For any positive integer d there exist positive constants $\eta = \eta(d)$, $\lambda = \lambda(d)$ and $n_0 = n_0(d)$ with the following property.*
If $V \subset \mathbb{C}^3$ is an algebraic surface (i.e. each component is two dimensional) of degree d then the following are equivalent:

(a) For at least one $n > n_0(d)$ there exist $X, Y, Z \subset \mathbb{C}$ such that $|X| = |Y| = |Z| = n$ and
$$|V \cap (X \times Y \times Z)| \geq n^{2-\eta};$$
(b) V has an irreducible component V_0 which is either a cylinder over a curve $F(x,y) = 0$ or $F(x,z) = 0$ or $F(y,z) = 0$ or, otherwise, there exist a one–dimensional connected algebraic group $\mathcal{G}$ and analytic multi–functions $f, g, h : \mathcal{G} \to \mathbb{C}$ of complexity bounded by $\lambda(d)$, such that their inverses are also analytic multi–functions of complexity bounded by $\lambda(d)$, and V_0 is the closure of a component of the $f \times g \times h$-image of the special subvariety $\mathcal{G}_{sp}$.
(c) Let $D \subset \mathbb{C}$ denote the open unit disc. Then either V contains a cylinder over a curve $F(x,y) = 0$ or $F(x,z) = 0$ or $F(y,z) = 0$ or, otherwise, there are one-to-one analytic functions $f, g, h : D \to \mathbb{C}$ with analytic inverses such that V contains the $f \times g \times h$-image of a part of the special subvariety $\langle \mathbb{C}, + \rangle_{sp}$ near the origin:
$$V \supseteq \left\{ \big(f(x), g(y), h(z)\big) \in \mathbb{C}^3 \;\middle|\; x, y, z \in D,\ x + y + z = 0 \right\}.$$
(d) For all positive integers n there exist $X, Y, Z \subset \mathbb{C}$ such that $|X| = |Y| = |Z| = n$ and $|V \cap (X \times Y \times Z)| \geq (n-2)^2/8$.
(e) Both (c) *and* (d) *can be localized as follows. There is a finite subset $H \subset \mathbb{C}$ of size $|H| \leq 3\lambda(d)$ and an irreducible component $V_0 \subseteq V$ such that whenever $P \in V_0$ is a point whose coordinates are not in H and $P \in U \subseteq \mathbb{C}^3$ is any neighborhood of P, then one may require that in* (c) *$\big(f(0), g(0), h(0)\big) = P$, and the Cartesian product $X \times Y \times Z$ in* (d) *lies entirely inside U.*

If $V \subset \mathbb{R}^3$ then the equivalence of (a), (c), (d) *and* (e) *still holds true with* real *analytic functions f, g, h defined on the interval $(-1, 1)$.*

This will follow from our Main Theorem (Theorem 1.4.2), see the proof near the end of Section 1.4. (The question whether, in the real case, (b) with a one-dimensional *real* algebraic group is equivalent to the other properties is left open.)

This result indicates a significant "jump": comparing (a) to (d) and (e) one shows that, for a given V, there are two possibilities: either we cannot get close to n^2, or, if we can, then it is not just "close-to-quadratic", rather, even a *"proper quadratic"* order of magnitude can be attained. Moreover, this quadric order of magnitude is achieved locally, everywhere along a component of V.

Actually, we do not know any example V with $|V \cap (X \times Y \times Z)| \geq n^{1+\varepsilon}$ (with $\varepsilon > 0$ and for infinitely many n) which does not satisfy (b), (c) and (d) of Theorem 1.1.3.

Problem 1.1.4. Are (b), (c) and (d) of Theorem 1.1.3 implied by the (much weaker) assumption $|V \cap (X \times Y \times Z)| \geq n^{1.001}$ in place of (a) above — for n large enough?

1.2. Incidences

Bounds on incidences play a central role in many areas of Erdős geometry and the theory of Geometric Algorithms. (Recently they have been used in Additive Number Theory, too, see [**40, 44, 42**]) The first such result was a celebrated and widely applicable bound of [**154**], concerning incidences of points and straight lines. Later on Pach and Sharir extended it to families Γ of (continuous) curves of d degrees of freedom (roughly speaking, the dimension of Γ as a variety is $\leq d$ and the curves are irreducible, see [**120**]). Then the number of incidences between p points and q curves of Γ is

$$I(p,q) = \mathcal{O}\Big(p^{d/(2d-1)}q^{(2d-2)/(2d-1)} + p + q\Big). \tag{1.2.1}$$

Specifically, if we are given such a family, and also n points in $\mathbb{R}^2$, then the number $f(m)$ of curves which pass through m or more points satisfies

$$f(m) = \mathcal{O}\Big(\frac{n^d}{m^{2d-1}} + \frac{n}{m}\Big). \tag{1.2.2}$$

In higher dimensions one must assume some non-degeneracy since, as the example of p points on a line and q planes containing this line shows, there are no nontrivial estimates in general. (One interesting problem was to find the right notion of non-degeneracy, which is weak enough to hold in interesting geometric situations.) We exclude those configurations where the intersection of a large number of our varieties contains a large number of our points. This is the essence of our notion of "combinatorial dimension", which is a invariant of the "incidence graph" defined below.

First we fix a constant b that we shall use throughout this section. Let $G \subseteq \mathbb{S} \times \mathbb{T}$ be a *bipartite graph.*[1] For all subsets $S \subseteq \mathbb{S}$, $T \subseteq \mathbb{T}$, and for each vertex $s \in S$ we denote by T_s the set of neighbors of s, and similarly by S_t the set of neighbors of the vertex $t \in T$.

Definition 1.2.1 (combinatorial dimension)**.** As we agreed above, b is a fixed constant throughout this section. Let $G \subseteq \mathbb{S} \times \mathbb{T}$ be a *bipartite graph.*[1] For all subsets $S \subseteq \mathbb{S}$, $T \subseteq \mathbb{T}$ we define by induction the *combinatorial dimension* $\operatorname{cdim}_b(S,T)$. We say that $\operatorname{cdim}_b(S,T) = 0$ if S has at most b

[1] I.e. G is a graph whose vertex set is the disjoint union of $\mathbb{S}$ and $\mathbb{T}$, and edges are only allowed between these sets, but not within any individual set.

vertices. In general, $\operatorname{cdim}_b(S,T) \le k$ for some $k \ge 1$, if there is a subset $T' \subseteq T$ such that

(A) T' is "almost the whole" of T, i.e. $|T \setminus T'| \le b$, and
(B) $\operatorname{cdim}_b\bigl(S_t, T' \setminus \{t\}\bigr) \le k-1$ for all $t \in T'$.

Finally, we set $\operatorname{cdim}_b(S,T) = \infty$ if the above induction does not assign any finite value to $\operatorname{cdim}_b(S,T)$.

Remark 1.2.2. This notion is more general than just excluding complete bipartite graphs. Actually, it is the prime feature of our definition that we *do* allow such subgraphs — but, of course, not too large ones.

Proposition 1.2.3. *If the bipartite graph $G \subseteq \mathbb{S} \times \mathbb{T}$ contains no $K_{u,v}$ (i.e, a complete bipartite subgraph with u vertices in $\mathbb{S}$ and v vertices in $\mathbb{T}$) and $u \le b+1$ then $\operatorname{cdim}_b(S,T) \le v$ for arbitrary subsets $S \subseteq \mathbb{S}$ and $T \subseteq \mathbb{T}$.* ∎

Definition 1.2.4. In geometry we often deal with configurations which consist of a collection of points, say P, and a collection of subsets, say Q, in some base space.

– The *incidence graph* of this configuration is the subset $G \subseteq P \times Q$ consisting of those pairs (p,q) where p is a point of q. By definition this is a bipartite graph.
– The *number of incidences* in this configuration, denoted by $I(P,Q)$, is the number of edges in the incidence graph.
– Finally $\operatorname{cdim}_b(P,Q)$, the *combinatorial dimension* of this configuration, is just the combinatorial dimension in the incidence graph.

We believe that, for families Γ of algebraic sets parametrised by a d dimensional variety and a set P of n points with combinatorial dimension $\operatorname{cdim}_b(P,\Gamma) \le k$, the number $f(m)$ of members $V \in \Gamma$ which contain at least m of the n points satisfies

$$f(m) = \mathcal{O}\Bigl(\frac{n^d}{m^{(kd-1)/(k-1)}} + \frac{n}{m}\Bigr), \tag{1.2.3}$$

which would be a generalisation of (the dual of) the Pach–Sharir bound (1.2.2). The following results show that, on the one hand, our expectations are "almost justified" for hyperplanes of (real) Euclidean spaces, when (1.2.3) and also the corresponding incidence bound like (1.2.1) will hold, with an arbitrary small error $\varepsilon > 0$ in the exponents. On the other hand, even for arbitrary algebraic sets and parameter variety, similar bounds can be established, with a (larger) constant D in place of d.

Theorem 1.2.5. *Let there be given a family $\mathcal{H}$ of hyperplanes in $\mathbb{R}^d$ and a finite point set $\mathcal{P}$ with combinatorial dimension $\operatorname{cdim}_b(\mathcal{P},\mathcal{H}) = k$. Moreover, let ε be any value such that*

$$0 < \varepsilon < \frac{k-1}{k(dk-1)},$$

and put

$$\alpha \stackrel{\text{def}}{=} \frac{d(k-1)}{dk-1} - \varepsilon;$$
$$\beta \stackrel{\text{def}}{=} k(1-\alpha) = \frac{k(d-1)}{dk-1} + k\varepsilon.$$

Then

$$I(\mathcal{P},\mathcal{H}) = \mathcal{O}\Big(|\mathcal{P}|^{\alpha}|\mathcal{H}|^{\beta} + |\mathcal{P}| + |\mathcal{H}|\log(2|\mathcal{P}|)\Big),$$

and the constant in this big–Oh expression depends on b, k, d and ε only.

Theorem 1.2.6. *Let there be given a family of algebraic subsets of a complex projective space $\mathbb{CP}^N$, parametrised by an algebraic set Y (of some other projective space). Let $\mathcal{V}$ be a finite subcollection from this family, and $\mathcal{P}$ a finite point set with combinatorial dimension $\operatorname{cdim}_b(\mathcal{P},\mathcal{V}) = k$. Then there exists a constant $D = D(\dim(Y)) > 0$ such that, for any ε with*

$$0 < \varepsilon < \frac{k-1}{k(Dk-1)}$$

and values

$$\alpha := \frac{D(k-1)}{Dk-1} - \varepsilon;$$

$$\beta := k(1-\alpha) = \frac{k(D-1)}{Dk-1} + k\varepsilon,$$

we have

$$I(\mathcal{P},\mathcal{V}) = \mathcal{O}\Big(|\mathcal{P}|^{\alpha}|\mathcal{V}|^{\beta} + |\mathcal{P}| + |\mathcal{V}|\log(2|\mathcal{P}|)\Big).$$

The constant of this big–Oh expression depends on b, k, ε, $\dim(Y)$, $\deg(Y)$, N, and the maximum degree of the members of the family (which is finite in each algebraic family).

Remark 1.2.7. We formulated the above theorem for projective space, so we could talk about degrees of algebraic subsets. Of course, the theorem remains valid for algebraic subsets of $\mathbb{C}^N$, but it is harder to formulate the precise dependence of the big–Oh expression.

Remark 1.2.8. Brass and Knauer found the upper bound $\mathcal{O}\Big(\big(|\mathcal{P}||\mathcal{V}|\big)^{d/(d+1)}\Big)$ in [**18**], under the assumption that the incidence graph contains no complete bipartite subgraph $K_{t,t}$ of $t+t$ vertices (for a fixed t). Their assumption is stronger than our condition of "bounded combinatorial dimension".

Basic properties of the combinatorial dimension. For the proof of the foregoing incidence results we need some preliminaries.

Proposition 1.2.9. *Let $G \subseteq S \times T$ be a finite bipartite graph such that $\operatorname{cdim}_b(S,T) = k \geq 1$. Then:*

(A) in each subgraph $G' \subseteq S' \times T'$ of the graph G we have $\operatorname{cdim}_b(S',T') \leq k$.
(B) each complete bipartite subgraph of G has at most $\mathcal{O}(|S| + |T|)$ edges, and
(C) G has at most $\mathcal{O}\left(|S| + |S|^{1-\frac{1}{k}}|T|\right)$ edges.

The constants in these big–Oh expressions depend on k and b, but not on the graph G.

Proof

(A) is obvious. To prove (B) one can show via a straightforward induction on k that either $|S| \leq b$ or $|T| \leq k(b+1)$. This is left to the reader.

Let's prove (C). It is clear for $k = 1$, otherwise we use induction. For arbitrary subsets $X \subseteq S$, $Y \subseteq T$ we denote by $G(X,Y)$ the subgraph of G

spanned by X and Y. Let $T' \subseteq T$ be the subset defined in Definition 1.2.1. The subgraph $G(S, T \setminus T')$ has at most $b|S|$ edges. Hence it is enough to estimate the number of edges of the subgraph $G(S, T')$, which we denote by E. To estimate E we add up the number of edges in the graphs $G(S_t, T')$ for each $t \in T'$. On the one hand for some positive C this sum is

$$\#\leq \sum_{q\in T'} C\left(|S_t|+|S_t|^{1-\frac{1}{k-1}}|T'|\right)\leq$$

$$\leq CE + C|T||T|^{\frac{1}{k-1}}\left(\sum_{q\in T'}|S_t|\right)^{1-\frac{1}{k-1}} = CE + C|T|^{\frac{k}{k-1}}E^{1-\frac{1}{k-1}}.$$

On the other hand we can count these edges according to their endpoints in S. Then for each point $s \in S$ we count the pairs of edges starting from s. Therefore

$$\#\geq \sum_{p\in S}|T'_s|^2 \geq \frac{1}{|S|}\left(\sum_{p\in S}|T'_s|\right)^2 = \frac{E^2}{|S|}$$

Comparing the two inequalities we get

$$\left(1-\frac{C|S|}{E}\right)E^{\frac{k}{k-1}} \leq C|S||T|^{\frac{k}{k-1}}$$

whence either $E \leq 2C|S|$ or $E \leq (2C)^{1-\frac{1}{k}}|S|^{1-\frac{1}{k}}|T|$. This proves the required upper bound for the number of edges in G. ∎

Proof of Theorems 1.2.5 and 1.2.6. The two results will be demonstrated along a common, almost identical line of reasoning, which follows that of [**31**]. We present both of them simultaneously, and mark with *(proof of 1.2.5.)* and *(proof of 1.2.6.)* the differences in the proofs.

Whenever, during the proof, we say that something is "bounded", it will mean that it is bounded in terms of b, k, ε, $\dim(Y)$, $\deg(Y)$, the dimension of the ambient space, and the maximum degree of the given subvarieties (which is finite in each algebraic family, and 1 for hyperplanes). Our goal is to exhibit a constant factor C' that is bounded in the aforementioned sense but sufficiently large to fit in the big–Oh notations in the claimed incidence bounds in the statements of the two Theorems.

Step I. To start with, we first dualize the situation as follows.
(proof of 1.2.5.): We assign points of the dual space $Y = (\mathbb{R}^d)^*$ to hyperplanes and, conversely, hyperplanes of Y to points.
(proof of 1.2.6.): We represent the algebraic sets in $\mathcal{V}$ by the corresponding points of the parameter space Y. For the other direction, let y^* denote the algebraic set parametrised by $y \in Y$. Then to each point $p \in \mathcal{P}$ we assign the set of those $y \in Y$ which satisfy $p \in y^*$, this is an algebraic subset of Y (of bounded degree).
In either case we denote by S the set of hyperplanes/algebraic subsets assigned to points in $\mathcal{P}$ and by T the set of points of Y assigned to the original hyperplanes/algebraic subsets. By definition we have $\mathrm{cdim}_b(S, T) = k$.

Step II. We are going to make use of the following two "cutting lemmata" in the two situations, respectively.

(proof of 1.2.5.): Given s hyperplanes in $\mathbb{R}^d$ and any positive integer $r < s$, the space can be subdivided into $\leq r^d$ parts such that each part is cut by $\mathcal{O}(s/r)$ of the hyperplanes. (Here "cutting a part" means "intersecting it but not containing it", see [**108**].)
(proof of 1.2.6.): Given s real-algebraic subsets in $\mathbb{R}^d$ with $d > 1$ and any integer $r < s$ large enough, the space can be subdivided into $\leq r^{2d-2}$ parts such that each part is cut by $\mathcal{O}(s \log r/r)$ of the algebraic subsets. ("Cutting", again, is used as in (a), see [**31**].)

Step III. We put $s = |S|$ and $t = |T|$. Moreover, we fix a sufficiently large r (to be specified later) and apply the foregoing "cutting lemmata":
(proof of 1.2.5.): we set $D \stackrel{\text{def}}{=} d$ and use Step II*(proof of 1.2.5.)* for the s hyperplanes in $Y = (\mathbb{R}^d)^*$ and get a decomposition into r^D parts;
(proof of 1.2.6.): Y is an algebraic set in some complex projective space $\mathbb{CP}^N$. By choosing appropriate coordinates, we can achieve, that the hyperplane at ∞ avoids all points of T. Then we can throw ∞ out, all of the incidences will happen in the complementary $\mathbb{C}^N$, which we identify with $\mathbb{R}^{2N}$. We write $d \stackrel{\text{def}}{=} \dim(Y)$ and project Y to $\mathbb{R}^{2d-1}$ in a generic manner (i.e., no incidences be lost), the algebraic sets in S will turn into real-algebraic sets of dimension at most $2d-2$. If $d > 1$ then we set $D = 4d-4$ and use Step II*(proof of 1.2.6.)* for $2d-1$ in place of d to get a decomposition of the underlying (real) space into r^D parts. If $d \leq 1$ then we set $D = d$. The algebraic sets in S are finite subsets of bounded size, hence we can use Step II*(proof of 1.2.5.)* and decompose the space again into r^D parts.
In all cases we have a decomposition into r^D parts (but of course with different D values), and each part is cut by $\mathcal{O}(s \log r/r)$ of the hyperplanes/algebraic subsets. In case of *(proof of 1.2.5.)* we could spare the $\log r$ factor, but we won't trouble with that: from now on, the two proofs will be identical.

Step IV. We first show the validity of the two assertions in two extreme cases, when s is very small or very large.
(1) If $s \leq r$ then we have $I(s,t) \leq st \leq rt$, so it is enough to chose $C' \geq r$.
(2) Next we deal with the case of very large s, when we assume that $r^{\frac{D}{1-\alpha}} s \geq t^k$. It is shown in Proposition 1.2.9(C) that $I(s,t) \leq \mathcal{O}(s + s^{1-\frac{1}{k}} t) \leq \mathcal{O}\left(s + s^{1-\frac{1}{k}} (r^{\frac{D}{\beta}} s^{\frac{1}{k}})\right) = \mathcal{O}\left(1 + r^{D/\beta}\right) s$, and this is exactly what we wanted, provided that, again, C' is large enough, as compared to r.

Step V. As for the general (not too large and not too small) values of s, we use induction, based upon the decomposition(s) found in Step III. If $s > r$ and $s < r^{\frac{-D}{1-\alpha}} t^k$ then $s < r^{-D} s^\alpha t^{k-k\alpha} = r^{-D} s^\alpha t^\beta$. We distribute the given points and algebraic subsets among the parts found above. Assign each point to the part containing it, and to each part we assign all those algebraic subsets which cut it. So one hyperplane/algebraic subset belongs to many parts. Let t_i and s_i denote the number of points and algebraic subsets assigned to the i-th part, and M denote the number of parts. Then $M \leq r^D$, $\sum t_i = t$ and $s_i \leq Bs \log r/r$ for for a bounded constant B and all i.

If a given point and a given algebraic subset meet each other then there are two possibilities. Either they are assigned to the same part, or the point lies in a part which is contained entirely in the algebraic subset. The number of incidences of the first kind is at most $\sum I(s_i, t_i)$, it can be estimated by the induction hypothesis. On the other hand, we count the second kind on each part separately. By Proposition 1.2.9(B), at most $\mathcal{O}(s+t)$ such incidences occur in each part, so there are all together at most $r^D\mathcal{O}(s+t)$ of them. Moreover, if r is large enough then $B\log(r)/r < 1/2$, hence its logarithm is less than 1 (we use base 2 logarithm). Hence we have the following chain of inequalities:

$$\begin{aligned}
I(s,t) &= \sum_{i=1}^{M} I(s_i,t_i) + r^D\mathcal{O}(s+t) < \\
&< C'\sum_{i=1}^{M}\left(s_i^\alpha t_i^\beta + s_i + t_i\log(2s_i)\right) + r^D\mathcal{O}(s+t) \leq \\
&\leq \left[C'B\left(\frac{s}{r}\right)^\alpha (\log r)^\alpha \sum_{i=1}^{M} t_i^\beta + C'MB\left(\frac{s}{r}\right)\log r + \right. \\
&\qquad \left. + C't\log\left(2B\frac{s\log r}{r}\right)\right] + r^D\mathcal{O}(s+t) \leq \\
&\leq \left[C'B\frac{\log r}{r^\alpha}s^\alpha M^{1-\beta}\left(\sum_{i=1}^{M} t_i\right)^\beta + C'Br^D\frac{\log r}{r}\left(r^{-D}s^\alpha t^\beta\right) + \right. \\
&\qquad \left. + C't\log(2s) - C't\right] + r^D\mathcal{O}(s+t) \leq \\
&\leq C'B\frac{\log r}{r^\alpha}s^\alpha\left(r^{D(1-\beta)}\right)t^\beta + C'B\frac{\log r}{r}s^\alpha t^\beta + C't\log(2s) + \\
&\qquad + (\mathcal{O}(r^D) - C')t + r^D\mathcal{O}(s) \leq \\
&\leq C'\left[\left(B\frac{\log r}{r^{\alpha - D(1-\beta)}}\right)s^\alpha t^\beta + \left(B\frac{\log r}{r}\right)s^\alpha t^\beta + t\log(2s) + \mathcal{O}\left(\frac{r^D}{C'}\right)s\right] + \\
&\qquad + \left(\mathcal{O}(r^D) - C'\right)t
\end{aligned}$$

So it is enough to pick r large enough so that

$$B\frac{\log r}{r^{\alpha - D(1-\beta)}} + B\frac{\log r}{r} < 1$$

— which is clearly possible since the exponent $\alpha - D(1-\beta) = D(k-1) - (Dk-1)\alpha$ is positive — and then choose $C' \geq \mathcal{O}(r^D)$.
Theorems 1.2.5 and 1.2.6 are proved now. ∎

Combinatorial dimension in geometric settings. As usual, a subset $X \subseteq \mathbb{C}^n$ is *algebraic* if it can be described in terms of polynomial equations. Of course, there is an analogous theory of algebraic subsets in the projective space $\mathbb{CP}^n$, and there are even more general versions. Most of the definitions below can be easily generalised for all of these. If we wants to emphasise that X is a subset of $\mathbb{C}^n$ or $\mathbb{CP}^n$ then we call X an *affine algebraic*

set or a *projective algebraic set*. In the following definitions, for simplicity, we restrict our attention to affine algebraic sets.

Let P be a point of an algebraic set X, in a neighbourhood of P we choose defining equations for the set such a way that they generate a radical ideal (e.g., in case of a hypersurface this means that in the prime decomposition of the defining polynomial each prime occurs at most once). Then the rank of the Jacobi matrix of the defining equations at P is independent of the equations chosen. Moreover, the rank is a semi-continuous function of P. We call an algebraic set *smooth*, if this rank is locally constant. More generally, we say that P is a smooth point of the algebraic set, or the set is smooth at P, if the rank is constant in a neighbourhood of P. In each algebraic set the smooth points form an open and dense subset.
The minimum value c of the above rank is the *codimension* of X, and $n-c$ is the *dimension* of X, denoted by $\dim(X)$.
An algebraic set X is *reducible* if it there is no nontrivial decomposition of X into a finite union of algebraic subsets. X is *irreducible* if it is not reducible. A *variety* is a reducible algebraic set. If we wants to emphasise that the variety lives in $\mathbb{C}^n$ or $\mathbb{CP}^n$ then we call it an *affine variety* or a *projective variety*.

Definition 1.2.10. We call a subset $\mathcal{P} \subseteq X$ of an algebraic set X *constructible* if it is the Boolean combination of finitely many algebraic subsets $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_t \subseteq X$, i.e., it can be expressed from the $\mathcal{P}_i$, using unions and differences (no need for the operation of intersection since the family of algebraic sets is closed for it).

The following notion of "constructible–parametrisation" extends usual "algebraic parametrisations".

Definition 1.2.11 (Family of algebraic sets)**.** Let there be given an algebraic set A and an algebraic "parameter set" $\mathcal{P}$. We say that a constructible (resp. algebraic) set $A_0 \subset \mathcal{P} \times A$ gives a *constructible* (resp. *algebraic*) *parametrisation* of a family $\mathcal{H}$ of algebraic subsets of A if, for each $p \in \mathcal{P}$, the corresponding subset $\mathcal{H}_p \stackrel{\text{def}}{=} A_0 \cap \big(\{p\} \times A\big)$ is either empty, or belongs to $\mathcal{H}$, and $\mathcal{H} = \{H_p \mid p \in \mathcal{P}, \mathcal{H}_p \neq \emptyset\}$. In such cases we shall also say that $\mathcal{H}$ is a *constructible family* (resp. an *algebraic family*). $\mathcal{P}$ is called the *parameter space* of the family and the sets $\mathcal{H}_p$ are the *members* of the family.

Examples of constructible families:

(1) all algebraic subsets of given degree in a given projective variety — can be parametrised via the coefficients of their equations, have to leave out those, which have multiple components;
(2) smooth subvarieties of given degree in a given projective variety — subtract those which are not smooth (they can be described by finitely many vanishing determinants) and those which are reducible (they can be described as unions of algebraic sets of smaller degree);
(3) smooth subvarieties of given degree in a given projective variety going through a number of given points (substituting the given points into the equations gives linear equations for the coefficients);
(4) smooth subvarieties of given degree in a given projective variety going through a number of given points with prescribed tangent directions

(tangent directions give linear equations for the coefficients of the derivatives of the equations of the variety).

The following lemma shows that the combinatorial dimension of the incidence structure of some points of an algebraic set A of bounded dimension and a constructible family $\mathcal{H}$ also obeys the same bound, provided that we exclude "too many" points on the intersection of "too many" members of the family.

Definition 1.2.12 (general position)**.** We call the intersection of infinitely many members of $\mathcal{H}$ a *forbidden set*, if the intersection is infinite (i.e., at least one dimensional), and denote the family of forbidden sets by $\mathcal{F} = \mathcal{F}_{\mathcal{H}}$. Moreover, for a fixed constant b, we say that a point set $P \subset A$ is in $\mathcal{F}$*–general position*, if $|P \cap F| \leq b$, for all $F \in \mathcal{F}$.

We note that if $k = \dim(A) = 2$ and each member of $\mathcal{H}$ is irreducible then $\mathcal{F}$ is empty, hence any subset is in $\mathcal{F}$-general position automatically.

Lemma 1.2.13. *Let A be a k-dimensional projective variety, and $\mathcal{H}$ a constructible family of algebraic subsets. As above, denote by $\mathcal{F} = \mathcal{F}_{\mathcal{H}}$ the family of forbidden subsets. Then one may choose b sufficiently large so that for any finite set $S \subseteq A$ of points in $\mathcal{F}$–general position we have* $\mathrm{cdim}_b(S, \mathcal{H}) \leq k$.

In fact it is easier to prove more. We shall get a common b for a whole family of A-s:

Lemma 1.2.14. *Let $\mathcal{A}$ and $\mathcal{H}$ be constructible families of algebraic subsets of a given variety, assume that each member of $\mathcal{A}$ has dimension at most k. As above, denote by $\mathcal{F} = \mathcal{F}_{\mathcal{H}}$ the family of forbidden subsets for $\mathcal{H}$. Then one may choose b sufficiently large so that for any finite set S of points in $\mathcal{F}$–general position lying on the same member of $\mathcal{A}$ we have* $\mathrm{cdim}_b(S, \mathcal{H}) \leq k$.

Proof The proof goes by induction on k. In any family there is an upper bound on the number of irreducible components of the members. Hence we may replace $\mathcal{A}$ with the family of the irreducible components, which is again a constructible family — at the end we simply multiply b with the maximum number of components. So we assume that each member of $\mathcal{A}$ is irreducible. We may subdivide the parameter space of $\mathcal{A}$ into a bounded number of parts, and prove the lemma separately for each part — then we simply add up the b-s we get for the parts. Hence we may assume, that each member of $\mathcal{A}$ is irreducible and has the same dimension. If this dimension is smaller than k, then the induction hypothesis takes care of the situation. If $k = 0$, then each member of $\mathcal{A}$ has just one point, we may take any $b \geq 1$. So we assume that each member of $\mathcal{A}$ has dimension $k > 0$. We note here that also the forbidden sets form a constructible family. We subdivide $\mathcal{A}$ further: on the family of those members which are contained in some forbidden set, we may take any $b \geq 1$, and get that the combinatorial dimension is at most 1. So we may restrict $\mathcal{A}$ to the complement: the family of those which are not contained in any forbidden set. Then there is a common upper bound α such that each member of $\mathcal{A}$ is contained in at most α members of $\mathcal{H}$. Let $\mathcal{B}$ denote the family of all proper intersections of a member of $\mathcal{A}$ with a member of $\mathcal{H}$. Then $\mathcal{B}$ is a constructible family of algebraic sets of dimension smaller than k, hence the induction hypotheses applies to $\mathcal{B}$ and $\mathcal{H}$ with any

bound $b \geq b'$. We claim that the Lemma holds for $\mathcal{A}$ and $\mathcal{H}$ with each $b \geq \max(\alpha, b')$. Let S be a finite set of points in $\mathcal{F}$–general position lying on the same member A of $\mathcal{A}$ — we must calculate the combinatorial dimension $\mathrm{cdim}_b(S, \mathcal{H})$. Now we define $\mathcal{H}' \subseteq \mathcal{H}$ by throwing out those members which contain A (there are at most α of them). If $H \in \mathcal{H}'$ is any of the remaining members, then $S \cap H \subseteq A \cap H$ is a member of $\mathcal{B}$, hence the combinatorial dimension $\mathrm{cdim}_b(S \cap H, \mathcal{H}')$ is at most $k-1$. This holds for all H, so the combinatorial dimension $\mathrm{cdim}_b(S, \mathcal{H})$ is at most k. This proves the lemma. ∎

The number of algebraic subsets with "very many" points is always linear. The following application is a generalisation to complex numbers and to higher dimensions of [**45**], Lemma 15.

We are given n^k points on a k-dimensional variety, then under some mild non-degeneracy conditions (e.g. irreducible for $k = 2$) there can be at most a constant times n subvarieties of any given degree passing through at least n^{k-1} of the points. The attractive feature is that the exponents are always independent of the degree.

Corollary 1.2.15. *We are given a projective variety (over the complex numbers), and positive integers b, d, k. Let $\mathcal{P}$ be a set of n^k points (for some integer n) on the variety, and $\mathcal{V}$ a collection of algebraic subsets of degree at most d, each containing at least $\Omega(n^{k-1})$ of the points. Suppose that the combinatorial dimension $\mathrm{cdim}_b(\mathcal{P}, \mathcal{V})$ is k. Then there are at most $O(n)$ of these subvarieties.*

Proof By assumption, each member of $\mathcal{V}$ contains at least λn^{k-1} of the points of $\mathcal{P}$ for some positive λ. Then $|\mathcal{P}| = n^k$ and we write $|\mathcal{V}| = Cn$, we want a bound on C. We can assume that $C \geq 1$. Just using Theorem 1.2.6 and $I(\mathcal{P}, \mathcal{V}) \geq (\lambda n^{k-1})(Cn) = C\lambda n^k$ yields

$$C\lambda n^k \leq C' \left(n^{k\alpha}(Cn)^{k-k\alpha} + n^k + Cn \log n^k \right) \leq C' n^k \left(C^\beta + 1 + Ckn^{1-k} \log n \right)$$

whence

$$C \left(\lambda - \frac{kC' \log n}{n^{k-1}} \right) \leq C'(C^\beta + 1) \leq 2C'C^\beta.$$

We can assume that n is large enough to make the left hand side positive. Then we get an upper bound on C and the corollary is proved. ∎

1.3. Compositions

Throughout this section we work with quasi-projective algebraic sets defined over an algebraically closed field. The goal of this section is to state and prove the Composition Lemma. Actually, it is a special case of the very general Group Configuration Theorem of Hrushovski [**75**] (for a self–contained explanation, see e.g. [**122**]).

But in the literature Hrushovski's theorem is always stated in the language of logic (model theory), and it would not be easy for us even to explain the words in his statement. So we prefer to give here a complete proof for our special case, and leave it for the interested reader to compare it with Hrushovski's theorem.

We study families of multi-valued, generically finite-to-finite functions, called "multi–functions". The technical difficulty comes from the possibility

that the graph of such a function might be reducible, and different components may behave differently. We call two families "related" if they have a common component. We shall use Galois theory to control how the composition of two multi–functions can break up into irreducible components.

Definition 1.3.1. A *multi–function* $(F : A \to B)$ between two irreducible projective varieties A and B is a nonempty closed algebraic subset $F \subset A \times B$ such that the two projections of F are generically finite and surjective. Then A, B and F necessarily have equal dimensions. If the projections $\pi_A : F \to A$ and $\pi_B : F \to B$ have degrees α and β, respectively, then we say that the *degree of the multi–function* F is $\max(\alpha, \beta)$. Such an F defines a multi-valued function $\pi_B \circ \pi_A^{-1} : A \to B$, for most points it has α values. In particular, *compositions* and *inverses* of multi–functions are defined as usual, and generically finite morphisms can be treated as multi–functions (given by their graphs). Compare this notion to the analytic multi–functions in the sense of Definition 1.1.2: here the branches are algebraic (and not just analytic), we define the function even in the "wrong" points (where the number of values is different from α), and we allow constant branches.
A *family of multi–functions* $(F_t : A \to B, t \in T)$ parametrised by the irreducible variety T is a closed algebraic subset $F \subset A \times B \times T$ such that the generic fiber $F_t \subset A \times B$ is a multi–function (here we only work with irreducible parameter spaces).

It is often useful to consider all subvarieties of a given variety. One would like to organize them into a nice algebraic family, which contains each of them exactly once. It is even nicer if this family has some universal properties. To do this for varieties is rather straightforward, but complications arise when one wants to include degenerations (i.e. reducible algebraic sets, possibly with multiple components). There are several ways to solve problems of these kind. For our purpose we could use any of them — we choose the Hilbert scheme. The *Hilbert scheme* of a projective variety is a universal parameter space for the family of all algebraic subsets (subschemes, to be more precise) – the precise definition, as well as the construction, can be found e.g. in section I.1. of [**90**]. An easier introduction can be found e.g. in Lecture 21 of [**69**].

Definition 1.3.2 (dimension, equivalence, common component)**.** A family $(F_t : A \to B, t \in T)$ of multi–functions induces a rational map from T to the Hilbert scheme of $A \times B$ by sending each point t from some dense open subset of T into the class of F_t. (The rational map does not depend on this open set, and in general, one cannot extend it continuously to the whole T.) We say that F is a *k-dimensional family* if the image of T is k-dimensional. Moreover, F is called *equivalent* to another family $\left(\hat{F}_u : A \to B, u \in U\right)$ if the rational images of T and U in the Hilbert scheme have the same closure, i.e. if they parametrise essentially the same set of multi–functions. Clearly every family is equivalent to one which parametrise each multi–function only once. We say that F and another family $(G_s : A \to B, s \in S)$ have a *common component* if there are families $\left(\hat{F}_u : A \to B, u \in U\right)$ and $\left(\hat{G}_u : A \to B, u \in U\right)$ equivalent to them, parametrised by the same U, such

that for all $u \in U$ the algebraic subsets $\hat{F}_u, \hat{G}_u$ of $A \times B$ have a common component.

Example 1.3.3. Let x, y denote the coordinates in the plane $\mathbb{C} \times \mathbb{C}$. We consider the family of plane curves given by the equations

$$\Big\{ux^2 + vxy + wy^2 = 0 \;\Big|\; (u, v, w) \in \mathbb{C}^3, w \neq 0\Big\}.$$

The parameter space of this family is $\mathbb{C}^3 \setminus \{w = 0\}$, and each member is the graph of the multi–function

$$F_{u,v,w} : \; x \mapsto y = x\frac{-v \pm \sqrt{v^2 - 4uw}}{2w}$$

(If $w = 0$ then the vertical axis is a component of the graph, hence it is not generically finite — we do not get a multi–function then.) These multi–functions are generically two valued, i.e. most of them sends all nonzero x to two different y values (the exception is the parameter locus $\{v^2 = 4uw\}$, where each x has only one image). They all have degree 2. The inverse multi–function is obtained simply by swapping x with y, i.e. for $w \neq 0 \neq u$ we have $F^{-1}_{u,v,w} = F_{w,v,u}$. The composition of any two of these multi–functions is a four-valued multi–function:

$$\Big(F_{p,q,r} \circ F_{u,v,w}\Big) : x \to y = F_{p,q,r}\left(x \cdot \frac{-v \pm \sqrt{v^2 - 4uw}}{2w}\right)$$
$$= x \cdot \frac{-v \pm \sqrt{v^2 - 4uw}}{2w} \cdot \frac{-q \pm \sqrt{q^2 - 4pr}}{2r}$$

One can easily get a polynomial equation for this multi–function by eliminating the square roots.

Although the parameter space of our family $F_{u,v,w}$ is 3-dimensional, this family is "over-parametrised": as we shall see soon, each member appears infinitely many times. The zero set of an equation does not change, if one multiplies the equation with a nonzero number, hence parallel parameter vectors correspond to equal multi–functions. Via identifying these parameter values, we arrive to a family parametrised by the lines through the origin, i.e. by the $w \neq 0$ part of the projective plane. Since the members of this new family are already pairwise different, we see that our family is 2 dimensional. In fact, in this case the Hilbert scheme is just the space of all nonzero quadric equations modulo multiplying by numbers, i.e. the 5 dimensional projective space (quadric equations have 6 coefficients). Our projective plane is a linear subspace of it, and the map $\mathbb{C}^3 \setminus \{w = 0\} \to \mathbb{P}^2 \subset \mathbb{P}^5$ is the above map to the Hilbert scheme.

It is useful to have a more general description of multi–functions.

Definition 1.3.4. A *generalised multi–function* $(F : A \to B)$ is an algebraic set F together with morphisms $F \to A$ and $F \to B$ such that the closure of the image of F in $A \times B$ is a multi–function. This is the multi–function *represented by* F. If $(G : B \to C)$ is another generalised multi–function, then the fiber product $H = F \times_B G$ has natural projections to A and C, and it is easy to see that $(H : A \to C)$ is a generalised multi–function representing the composition of F and G.

Definition 1.3.5 (standard family)**.** Let Γ be a connected algebraic group acting on a variety V. Then the *standard family* of multi–functions corresponding to this group action is the family $(F_\gamma : V \to V, \gamma \in \Gamma)$ where F_γ is the graph of the automorphism γ. We say that a family $(G_s : A \to B, s \in S)$ *is related to* the standard family F *along the multi–functions* $(\alpha : A \to V)$ and $(\beta : B \to V)$ if G has a common component with the family $(\beta^{-1} \circ F_\gamma \circ \alpha : A \to B, \gamma \in \Gamma)$.

The following lemma is a special case of the Group Configuration Theorem of Hrushovski [**75**].

Lemma 1.3.6 (Composition Lemma)**.** *Suppose there are two k-dimensional families of multi–functions, $(F_t : A \to B, t \in T)$ and $(G_s : B \to C, s \in S)$, such that the family of compositions $\big(G_s \circ F_t : A \to C, (t,s) \in T \times S\big)$ have a common component with a k-dimensional family. Then there is a k-dimensional connected algebraic group Γ acting on a variety V, and multi–functions $(\alpha : A \to V)$, $(\beta : B \to V)$ and $(\gamma : C \to V)$ such that the family F is related to the standard family corresponding to the Γ-action on V along α and β, and the family G is related to it along β and γ. Moreover, the degrees of α, β and γ can be bounded in terms of the degrees of the generic members F_t and G_s.*

Remark 1.3.7. It is easy to see that the components of the compositions form an at least k-dimensional family, so we treat here the most degenerate (i.e. most interesting) case.

Proof of the Composition Lemma. Through several steps we shall reduce the problem to simpler and simpler special cases, and finally solve the last (and simplest) case. During the process we can replace the families with equivalent families, in particular we can shrink S and T any time. The method is the following. In each step we choose one of the varieties, say B, and build a multi–function ρ from B to another variety B'. Then we replace B with B', and replace F_t, G_s with a component of the composition $\rho \circ F_t$ and $G_s \circ \rho^{-1}$. We treat A and C similarly. Clearly it is enough to deal with this new situation. Moreover, the degrees of these ρ will be visibly bounded, so one could get an explicit bound on the final degrees. We leave the degree calculations to the careful readers.

First Step. To begin with, we throw away unneeded components, so we may assume, that A, B, C, F, G are irreducible. After possibly switching to an equivalent family each F_t and G_s are irreducible as well. First we reduce the problem to the special case when the compositions $G_s \circ F_t$ are irreducible for almost all pairs (s,t), hence they themselves move in a k-dimensional family. We shall achieve this goal via some Galois theory. For any variety V we shall denote by $K(V)$ its function field. See Lecture 7 of [**69**] for the connection between generically finite rational maps and finite field extensions, and [**84**] for Galois theory.

First we look at the finite field extension $K(B \times T) \leq K(F)$, let $\tilde{F}$ denote the normalization of F in the Galois closure of this extension. Similarly, let $\tilde{G}$ denote the normalization of G in the Galois closure of the field extension

$K(B\times S) \leq K(G)$. $\tilde{F}_t$ and $\tilde{G}_s$ will denote the fibers of the natural morphisms $\tilde{F} \to T$ and $\tilde{G} \to S$. Clearly $\tilde{F}_t$ and $\tilde{G}_s$ are generalised multi–functions representing F_t and G_s. Now we pick general $s \in S$ and $t \in T$. We know that the fields $K(\tilde{G}_s)$ and $K(\tilde{F}_t)$ are Galois extensions of $K(B)$, so we can identify them with their unique copy in the algebraic closure of $K(B)$. Let $L = K(\tilde{G}_s) \bigcap K(\tilde{F}_t)$. Since s and t were general, L is in fact contained in all other $K(\tilde{G}_{s'})$ and all $K(\tilde{F}_{t'})$.

Suppose first that $L = K(B)$. Then we claim, that the composition $G_s \circ F_t$ is in fact irreducible. To see this, let $K(B) \leq M$ be the smallest field extension containing $K(\tilde{F}_t)$ and $K(\tilde{G}_s)$, and let $\tilde{B}$ be the normalization of B in M. The degree of the projection $\tilde{B} \to B$ is the product of the degrees of $\tilde{F}_t \to B$ and $\tilde{G}_s \to B$, and one can see at once that $\tilde{B} = \tilde{F}_t \times_B \tilde{G}_s$ at least over a dense open subset of B. Hence $(\tilde{B} : A \to C)$ is a generalised multi–function representing the composition $G_s \circ F_t$. On the other hand $\tilde{B}$ is irreducible by construction.

In general let B' be the normalization of B in L, then the maps $\tilde{G}_s \to B$ and $\tilde{F}_t \to B$ will factor through $\varphi : B' \to B$ for all s, t. Let F' and G' denote the image of $\tilde{F}$ and $\tilde{G}$ in $A \times B' \times T$ and $B' \times C \times S$, respectively. Clearly these are families of multi–functions $({F'}_t : A \to B', t \in T)$ and $({G'}_s : B \to C, s \in S)'$ such that $G_s = {G'}_s \circ \varphi^{-1}$ and $F_t = \varphi \circ F'_t$ for all s, t. (Here we treat φ as a multi–function.)

We see from the construction that $L = K(B')$ is Galois over $K(B)$, so the components of the multi–function $\varphi^{-1}\circ\varphi$ are just the relative automorphisms of B' over B (one for each element of the group $\mathrm{Gal}(B'/B)$). But then $G_s \circ F_t = G'_s \circ (\varphi^{-1}\varphi) \circ F'_t$ also splits according to the above Galois group, and one of these components, corresponding to a certain automorphism, moves in a k-parameter family. By composing each multi–function G'_s with this automorphism we arrive to the situation, that one component of $G'_s \circ F'_t$ moves in a k-parameter family. Now we replace B with B', F with F' and G with G', and we achieved that $K(\tilde{G}_s) \bigcap K(\tilde{F}_t) = K(B)$ for general s, t. This implies that $G_s \circ F_t$ is irreducible, and therefore not just a component, but the entire composition moves in a k-dimensional family.

Second Step. Let R be the k-dimensional subvariety of the Hilbert scheme of $A \times C$ parameterizing (almost all of) the compositions $G_s \circ F_t$. As we go along, we shall freely shrink R to dense open subsets. Then there is a universal family $(H_r : A \to C, r \in R)$ and the composition of the original families defines a rational map $\mu : T \times S \to R$ such that $G_s \circ F_t = H_{\mu(t,s)}$ for almost all pairs $(s, t) \in S \times T$. After shrinking R we shall assume that each H_r is irreducible.

Next we pick a generic $r \in R$, let Q_r be a component of the inverse image $\mu^{-1}(r)$. Clearly $Q_r \subset S \times T$ is a k-dimensional algebraic subset and $G_s \circ F_t = H_r$ for almost all $(s, t) \in Q_r$. Then G_s is a component of the multi–function $G_s \circ F_t \circ F_t^{-1} = H_r \circ F_t^{-1}$, which is independent of s. Hence for general F_t there are at most finitely many G_s with the property $G_s \circ F_t = H_r$. Therefore the projection $Q_r \to S$ is generically finite, hence surjective. Then for almost all $s \in S$ there is a $(t, s) \in Q_r$ and similarly for almost all $t \in T$ there is a $(t, s) \in Q_r$.

Third Step. This step will reduce the problem to the case when all of the multi–functions G_s and F_t^{-1} are graphs of rational maps $g_s : B \to C$ and $f_t : B \to A$. (i.e. they are single–valued functions.)

Let δ and γ denote the degree of the projections $H_r \to C$ and $G \to C \times S$. Then for general s the degree of $G_s \to C$ is also γ. Let $C^{(\delta)}$ and $C^{(\gamma)}$ denote the symmetric powers of C, then we can represent the multi–functions H_r and G_s by rational maps $h_r : A \to C^{(\delta)}$ and $g_s : A \to C^{(\gamma)}$, i.e. for general $a \in A$ let $h_r(a)$ be the δ-tuple of H_r-images of a, and $g_s(a)$ is defined similarly.

Let $X \subset C^{(\gamma)}$ denote the set of those γ-tuples which are contained in some δ-tuple in the image of h_r, it is an algebraic subset. Let $n = \dim(A) = \dim(B) = \dim(C)$, then the image of h_r is also n-dimensional. Every δ-tuple contains only finitely many γ-tuple, hence X is also n-dimensional. If $(t, s) \in Q_r$ then the image of g_s must be an n-dimensional subvariety of X, thus it is a component, and this must hold for almost all g_s. Since g_s moves in an irreducible family, the images of g_s must all be the same component $C^* \subseteq X$. But then we can replace C with C^* (possibly shrinking S to a dense open subset), and the multi–functions G_s become the graphs of the rational maps $g_s : B \to C^*$. Then we can turn around and repeat the whole argument for the compositions $F_t^{-1} \circ G_s^{-1} \supseteq H_r^{-1}$, and replace A with some A^*, and F_t^{-1} with rational maps $f_t : B \to A^*$. So from now on we assume that G_s and F_t^{-1} are graphs of the rational maps g_s and f_t for all s, t.

Fourth Step. We can look at finite (branched) covers $A' \to A$ such that each f_t factors through it. Each f_t has only finitely many factorizations hence we can find a maximal A'. We replace A with this cover, so from now on each such cover $A' \to A$ is an isomorphism. Similarly we can assume that C has no nontrivial finite cover which is a factor of each g_s.

Fifth Step. Next we reduce the problem to the case when all f_t and g_s are birational. For all $(t, s) \in T \times S$ the composite multi–function $G_s \circ F_t$ is just the closure of the image of the function $(f_t, g_s) : B \to H_{\mu(t,s)} \subset A \times C$. Then each f_t factors through each $H_r \to A$ hence each $H_r \to A$ is birational, and so is $H_r \to C$. Hence each H_r is the graph of a birational map $A \to C$. Now we fix origins $0 \in S$ and $0 \in T$, and replace A and C with $H_{\mu(0,0)}$ (using the projection maps). Then $H_{\mu(0,0)}$ becomes the identity multi–function. The map $\big(f_0(x), g_0(x)\big) \to \big(f_0(x), g_s(x)\big)$ is birational, hence the family $\{g_s\}$ is obtained from g_0 via composition with a k-dimensional family of birational automorphisms $\Gamma_s : C \to C$. Similarly $f_t = \Phi_t \circ f_0$ for another k-dimensional family of birational automorphisms $\Phi_t : A \to A$. But then $G_s \circ F_t = \Gamma_s \circ g_0 \circ f_0^{-1} \circ \Phi_t^{-1} \supset \Gamma_s \circ \Phi_t^{-1}$, and these are in fact equal because of the irreducibility. This implies that we can replace B with $H_{\mu(0,0)}$ using the map (f_0, g_0). So from now on we can assume that $A = B = C$, and F_t, G_s are graphs of the birational automorphisms Φ_t^{-1}, Γ_s.

Sixth Step. Now we replace the parameter spaces T, S, R with their images in the Hilbert scheme of $A \times A$, so we can compare them. Let Ψ_r denote the automorphism whose graph is H_r, then $\Phi_t^{-1} \circ \Gamma_s = \Psi_{\mu(t,s)}$. Since $\Phi_0^{-1} \circ \Gamma_s = \Gamma_s$ must belong to the Ψ family we find that $S \subseteq R$. But they are

both irreducible and k-dimensional hence they are equal. Similarly T is also equal to them. But then μ is an associative operation on S. For fixed t the compositions $\Phi_t^{-1} \circ \Gamma_s$ are all different, and form a k-dimensional irreducible family contained in Ψ. Hence $\mu(0, -)$ is generically one-to-one and onto. So μ has an inverse operation on the right hand side and similarly on the left hand side as well. Therefore it is a rational group structure. The family Γ_s defines a rational action of this group on A. It is proved in [**168**] that up to birational equivalence we have a standard family now. ∎

1.4. The main result in arbitrary dimension

Convention Throughout this section we study the following configuration: A, B, C are projective varieties and $F \subset A \times B \times C$ is a subvariety with the property that the projections $F \to A \times B$, $F \to B \times C$, $F \to C \times A$ are surjective and generically finite. It follows from this that $\dim(A) = \dim(B) = \dim(C) = \frac{1}{2}\dim(F)$. Let us fix a constant b, so we can use Definition 1.2.12, and talk about points in general position. The constants of the big–Oh and big–Omega expressions of this section will depend on b, $\dim(A)$ and the degrees of A, B, C, F, but are independent of the varieties themselves. We say that the degree of a variety is *bounded* if the degree is at most $\mathcal{O}(1)$, i.e. if there is an upper bound depending only on the above parameters.

Definition 1.4.1. If G is an algebraic group, then the *special subvariety* G_{sp} of the three-fold product G^3 is the set

$$G_{sp} = \left\{(a, b, c) \in G^3 \;\middle|\; abc = 1\right\}.$$

We say that our F is a *special subvariety* of the product $A \times B \times C$ if there is an algebraic group G and there are multi–functions $(\alpha : G \to A)$, $(\beta : G \to B)$ and $(\gamma : G \to C)$ such that F is a component of the $(\alpha \times \beta \times \gamma)$-image of the special subvariety $G_{sp} \subset G^3$. Here $\alpha \times \beta \times \gamma$ is the multi–function naturally induced by α, β, γ between G^3 and $A \times B \times C$.

Theorem 1.4.2 (Main Theorem)**.** *There is positive constant η depending only on* $\dim(A)$*, and bounded positive constants n_0, d with the following property: Suppose we choose $n > n_0$ points on each variety: $X = \{\mathbf{a}_1, \ldots, \mathbf{a}_n\} \subset A$, $Y = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\} \subset B$ and $Z = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\} \subset C$ in general position with respect to the family of all algebraic subsets of degree not exceeding d and dimension less than* $\dim(A)$ *(i.e., any such algebraic set only contains at most b of them, see the Convention above, or Definition 1.2.12). Assume that $|F \cap (X \times Y \times Z)| \geq n^{2-\eta}$. Then F must be a special subvariety. Moreover, the degrees of the multi–functions relating A, B and C to the group are bounded.*

The most general example of "many points" that we have found comes from a nilpotent group and the special subvariety F corresponding to it, when there indeed are at least $\Omega(n^2)$ points in $F \cap (X \times Y \times Z)$ for a good choice of X, Y and Z. This can be shown as follows.

Example 1.4.3. Let $\mathcal{G}$ be a group, $a_1, \ldots, a_t \in \mathcal{G}$ arbitrary elements and s a positive integer. Consider the set X_s of the at-most-s-term products of

terms $a_1, \dots, a_t, a_1^{-1}, \dots, a_t^{-1}$. According to a celebrated theorem of Gromov [**67**],

> *the growth of $|X_s|$ is polynomial in s iff the subgroup generated by the a_i has a nilpotent subgroup of finite index.*

Consequently, in such groups, $|X_s \cdot X_s| = |X_{2s}| \leq 2^{\lambda}|X_s|$, for an exponent λ, whence at least $|X_s|^2 \geq 2^{-2\lambda}|X_{2s}|^2$ three–tuples $\langle x_i, x_j, x_k \rangle \in X_{2s} \times X_{2s} \times X_{2s}$ satisfy $x_i x_j x_k = 1 \in \mathcal{G}$. (E.g. those where $x_i, x_j \in X_s$.)

We don't have any other example with at least $n^{1+\gamma}$ such points on F (cf. Problem 1.1.4).

Proof of the Main Theorem. In the proof we use part (b) of the following lemma several times.

Lemma 1.4.4. *(a) Let $\tilde{A}$ be an algebraic set, $\tilde{X} \subset \tilde{A}$ a finite subset in general position with respect to the family of all algebraic subsets of bounded degree and dimension smaller than $\dim(\tilde{A})$, and $U = \tilde{A}^r$ the product of finitely many copies of $\tilde{A}$ (hence $\tilde{X}^r \subset U$). Moreover, let $V \subseteq U$ be a subvariety of bounded degree and "small" dimension: assume that $\dim(V) < (t+1)\dim(\tilde{A})$, for a positive integer t. Then*

$$|V \cap \tilde{X}^r| = \mathcal{O}\big(|\tilde{X}|^t\big).$$

(b) If $X \subset A$, $Y \subset B$, $Z \subset C$ as in the Main Theorem 1.4.2, U is the product of r terms, each one of A, B or C, and S is the corresponding r–term product of X's, Y's and Z's (hence $S \subset U$), then $|V \cap S| = \mathcal{O}(n^t)$ holds for any $V \subseteq U$ of bounded degree and dimension smaller than $(t+1)\dim(A)$.

Proof (a) We use induction on $\dim(U)$. If $r \leq t$ then we are done. Otherwise U can be written as a product: $U = \tilde{A} \times U'$. Look at the projection $\pi : V \to U'$, let $Z \subseteq U'$ be the locus of those points whose inverse image is $\dim(\tilde{A})$ dimensional. Clearly $\dim(\pi(V)) \leq \dim(V) < (t+1)\dim(\tilde{A})$ and $\dim(Z) \leq \dim(V) - \dim(\tilde{A}) < t\dim(\tilde{A})$. We apply the induction hypothesis to $\pi(V) \subseteq U'$ and to $Z \subseteq U'$. If $p \in V \cap \tilde{X}^r$, then $\pi(p) \in \pi(V) \cap \tilde{X}^{r-1}$, and by the induction hypothesis there are at most $\mathcal{O}(|\tilde{X}|^t)$ possible values for $q = \pi(p)$. If $q \in \pi(V) \setminus Z$ is a possible value then V intersects the fiber $\pi^{-1}(q) = \tilde{A}$ in an algebraic subset of bounded degree, whose dimension is smaller than $\dim(\tilde{A})$, hence there are only a bounded number of possible p values in each $\pi^{-1}(q)$, and all together there are at most $\mathcal{O}(|\tilde{X}|^t)$ such points. On the other hand, by the induction hypothesis there are at most $\mathcal{O}(|\tilde{X}|^{t-1})$ special q values in Z, and in each $\{q\} \times \tilde{A}$ there are exactly $|\tilde{X}|$ possible p in U, these are again at most $\mathcal{O}(|\tilde{X}|^t)$ points. Part (a) of the lemma is proved.
(b) follows easily if we use $A \cup B \cup C$ in place of $\tilde{A}$ and $X \cup Y \cup Z$ in place of $\tilde{X}$ in part (a). ∎

Proof of Theorem 1.4.2. By assumption A, B, C and F are irreducible. For each $p \in C$ let $(F_p : A \to B)$ denote the the multi–function whose graph is the intersection of F with $A \times B \times \{p\}$. It is indeed a multi–function away from a proper closed subset, and we shall simply restrict C to the

complement of it. We loose at most $\mathcal{O}(1)$ from the points $\mathbf{c}_i$, so there are more than $n/2$ of them remains, and we loose at most $\mathcal{O}(n)$ points of the form $(\mathbf{a}_i, \mathbf{b}_j, \mathbf{c}_k)$ on F, so at least $\frac{1}{2}n^{2-\eta}$ remains. So from now on each F_p is a multi–function.

We look at the family of compositions $F_{st} = F_s^{-1} \circ F_t$, and for convenience we define $\mathcal{F}_l = F_{\mathbf{c}_l}$ and $\mathcal{F}_{lm} = F_{\mathbf{c}_l \mathbf{c}_m}$. If it happens to have a component moving in a k-dimensional family then we apply the Composition Lemma. We get an algebraic group G acting on a variety V, and multi–functions $(\alpha : V \to A)$, $(\beta : V \to B)$ and $(\gamma : G \to C)$ such that F is contained in the $\alpha \times \beta \times \gamma$ image of the graph of the G-action on V. The assumption that the projections of F are generically finite imply that this group action is transitive. Since $\dim(V) = \dim(G) = k$, V must be the quotient of G by a finite subgroup, and we can simply replace V with G itself. This means that our F is special.

Let's look at the other case when each component of the compositions F_{st} forms an at least $(k+1)$-dimensional family, we want to get a contradiction in this case. After possibly shrinking C, each component of the composition defines a map φ from $C \times C$ to the Hilbert scheme of $A \times A$. Let U denote the union of those fibers of φ which are less than $\dim(A)$ dimensional (it is a dense open subset), and V denote the complement of U. Clearly V contains at most $\mathcal{O}(n)$ points of the form $(\mathbf{c}_l, \mathbf{c}_m)$, we shall call the corresponding $\mathcal{F}_{lm}$ *forbidden*, and call the others *ordinary*. On the other hand the fibers of the restriction $\varphi|U$ contain at most $\mathcal{O}(1)$ of these points. Hence among all components of all $\mathcal{F}_{lm}$ there are at most $\mathcal{O}(n)$ forbidden, and there are $\mathcal{O}(n^2)$ ordinary components, each repeated at most $\mathcal{O}(1)$ times. Moreover, any two of the ordinary components intersect each other in at most $\mathcal{O}(1)$ points.

We shall estimate the number of points in the set

$$\mathcal{H} = \left\{ (\mathbf{a}_i, \mathbf{a}_j, \mathbf{c}_l, \mathbf{c}_m) \;\middle|\; (\mathbf{a}_i, \mathbf{a}_j) \in \mathcal{F}_{lm} \right\}$$

From the forbidden components of $\mathcal{F}_{lm}$-s we get at most $\mathcal{O}(n^2)$ elements in $\mathcal{H}$. On the other hand we can use Theorem 1.2.6 to get an upper bound on the number of elements of $\mathcal{H}$ arising from the ordinary components. In our case we have $k = 2$, hence $\alpha+\beta = k-(k-1)\alpha < k-(k-1)(1-\frac{1}{k}) = 2-\frac{1}{k} = \frac{3}{2}$. Hence we get

$$|\mathcal{H}| \leq \mathcal{O}\left(n^{2(\alpha+\beta)}\right) = \mathcal{O}\left(n^{3-\eta'}\right)$$

for some positive η'. Next we want to get a lower estimate and compare with it. To begin with, let

$$\mathcal{H}' = \left\{ (\mathbf{a}_i, \mathbf{a}_j, \mathbf{b}_k, \mathbf{c}_l, \mathbf{c}_m) \;\middle|\; (\mathbf{a}_i, \mathbf{b}_k) \in \mathcal{F}_l, (\mathbf{a}_j, \mathbf{b}_k) \in \mathcal{F}_m \right\}.$$

We want to compare the two sets via the natural projection $\mathcal{H}' \to \mathcal{H}$ (i.e. the forgetting of the $\mathbf{b}$ coordinate). Therefore we study the natural projection

$$\begin{aligned} \psi : W = \left\{ (\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{c}, \mathbf{c}') \;\middle|\; (\mathbf{a}, \mathbf{b}, \mathbf{c}), (\mathbf{a}', \mathbf{b}, \mathbf{c}') \in F \right\} \to \\ \to \left\{ (\mathbf{a}, \mathbf{a}', \mathbf{c}, \mathbf{c}') \;\middle|\; \exists \mathbf{b} : (\mathbf{a}, \mathbf{b}, \mathbf{c}), (\mathbf{a}', \mathbf{b}, \mathbf{c}') \in F \right\} \end{aligned}$$

As before, let $V \subset W$ be the union of the $\dim(B)$-dimensional fibers, and U it's complement. As before, the fibers of the map $\mathcal{H}' \bigcap U \to \mathcal{H}$ are bounded, hence $|H' \bigcap U| < \mathcal{O}(|\mathcal{H}|)$. Moreover, $\dim(V) < \dim(W) = 3\dim(A)$, hence $|\mathcal{H}' \bigcap V| < \mathcal{O}(n^2)$. Putting these together we get

$$|\mathcal{H}'| < \mathcal{O}\left(|\mathcal{H}| + n^2\right) < \mathcal{O}(n^{3-\eta'})$$

On the other hand we have

$$|\mathcal{H}'| = \sum_{\mathbf{b}_k} \Big| \left\{(\mathbf{a}_i, \mathbf{a}_j, \mathbf{b}_k, \mathbf{c}_l, \mathbf{c}_m) \;\Big|\; (\mathbf{a}_i, \mathbf{b}_k, \mathbf{c}_l) \in F, (\mathbf{a}_j, \mathbf{b}_k, \mathbf{c}_m) \in F\right\} \Big|$$

$$= \sum_{\mathbf{b}_k} \Big| \left\{(\mathbf{a}_i, \mathbf{b}_k, \mathbf{c}_l) \in F\right\} \Big| \cdot \Big| \left\{(\mathbf{a}_j, \mathbf{b}_k, \mathbf{c}_m) \in F\right\} \Big|$$

$$= \sum_{\mathbf{b}_k} \Big| \left\{(\mathbf{a}_i, \mathbf{b}_k, \mathbf{c}_l) \in F\right\} \Big|^2 \geq \frac{1}{n} \left(\sum_{\mathbf{b}_k} \Big| \left\{(\mathbf{a}_i, \mathbf{b}_k, \mathbf{c}_l) \in F\right\} \Big| \right)^2 = \frac{1}{n}|F|^2$$

Out of this we get, that $|F| < \mathcal{O}(n^{2-\eta'/2}) < \frac{1}{2}n^{2-\eta}$ if $\eta < \eta'/2$ and n is large enough. This is in contradiction with our assumption. The proof of the theorem is now complete. ∎

Proof of Theorem 1.1.3. First part of (e) ⇒ (c) and **Second part of (e) ⇒ (d) ⇒ (a)** are obvious if n is large enough.
(a) ⇒ (b) is a special case of the Main Theorem 1.4.2 just proven.
(b) ⇒ (c) and first part of (e). With the notation of (b), let us assume that V_0 is not a cylinder, and let $H \subset \mathbb{C}$ denote the union of the finite exceptional sets of the analytic multi–functions f^{-1}, g^{-1} and h^{-1}. Let $P = (x, y, z) \in V_0$ be as in (e), i.e. such that $x, y, z \notin H$. Then the value set $f^{-1}(x) \times f^{-1}(y) \times f^{-1}(z)$ intersects the special subvariety, i.e. there is a point $(a, b, c) \in G_{sp}$ such that $f(a)$, $g(b)$ and $h(c)$ are defined, and the value set $f(a) \times g(b) \times h(c)$ contains P. Then the translated functions $x \to f(x \oplus a)$, $y \to g(y \oplus b)$ and $z \to h(z \oplus c)$ are all defined in $0 \in \mathcal{G}$. According to Definition 1.1.2, we can choose their analytic branches $\hat{f}$, $\hat{g}$ and $\hat{h}$ defined in a common neighborhood $0 \in W \subset \mathcal{G}$ so that the $\hat{f} \times \hat{g} \times \hat{h}$-image of 0 is just P, and the image of $\mathcal{G}_{sp} \cap W^3$ is contained in V_0. As we have seen, all the three types of one-dimensional algebraic groups are quotient groups of $\langle \mathbb{C}, + \rangle$, so we can find a surjective homomorphism $\varphi : \langle \mathbb{C}, + \rangle \to \mathcal{G}$. By rescaling $\mathbb{C}$, we can achieve that $\varphi(D) \subseteq W$, the restriction $\varphi|_D$ is one-to-one, and $\varphi(0) = 0$. We note here that by Definition 1.1.2 the analytic branches $\hat{f}$, $\hat{g}$ and $\hat{h}$ are also one-to-one. Now it is clear, that the compositions $\hat{f} \circ \varphi$, $\hat{g} \circ \varphi$ and $\hat{h} \circ \varphi$ satisfy the conditions in (c), and also the strengthening in (e).
(c) ⇒ (d) and (e). Let W denote the open set $W \stackrel{\text{def}}{=} f(D) \times g(D) \times h(D)$, we shall find our Cartesian product $X \times Y \times Z$ inside W. Let us choose any point $P \in W \cap V_0$. By Example 1.1.1 (c), we can find $n \times n \times n$ Cartesian product $X \times Y \times Z$ arbitrarily close to P (i.e. inside W) such that V_0 intersects it in at least $(n-2)^2/8$ points. This proves (d), and that implies (a), (b) and the first part of (e), as we have seen above. We still have to prove the other part of (e). By the (already established) first part, we may assume that $W = f(D) \times g(D) \times h(D) \subset U$. But then the above

construction gives us a Cartesian product $X \times Y \times Z \subset W \subset U$, hence the whole (e) is proven.

Finally, the real special case. The implications $\big\{(\mathrm{c}) + \text{first part of } (\mathrm{e})\big\} \Rightarrow \big\{(\mathrm{d}) + (\mathrm{e})\big\} \Rightarrow$ (a) hold just as in the complex case. To show (a) $\Rightarrow \big\{(\mathrm{c}) + \text{first part of } (\mathrm{e})\big\}$, note that any $V \subset \mathbb{R}^3$ and $X \times Y \times Z \subset \mathbb{R}^3$ can be considered as subsets of $\mathbb{C}^3$. Therefore, the original complex version of the implication provides the complex analytic functions f, g and h. We expand into power series the inverse functions f^{-1}, g^{-1} and h^{-1} around the points $f(0)$, $g(0$ and $h(0)$ (i.e. the coordinates of P) respectively. The three power series — perhaps with complex coefficients — have no constant terms, and since they are invertible, their linear terms must be nonzero. First we restrict ourselves to real variables. It would be natural to consider the real parts of the three power series, but these may not be invertible (or even one or more of these may be identically zero) — and the same may happen to the imaginary parts. That is why we consider a "generic" $\lambda \in [0,1]$ and change each coefficient c to $\lambda\Re(c) + (1-\lambda)\Im(c)$. Moreover, if $c \neq 0$, then the new coefficient cannot vanish for more than one value of λ. We choose λ so that these new (real) power series also have nonzero linear terms. Finally we invert the three real power series, and rescale $\mathbb{R}$ so that the inverses are convergent on $(-1,1)$, and have nonzero derivative on the whole interval. ∎

1.5. Applications

A problem of Hirzebruch. Suppose we are given an arrangement of n non-degenerate conic sections in the projective plane over some field of characteristic different from 2, and assume that no three of them are tangent to each other at the same point. Hirzebruch [**74**] asked for a non-trivial (sub-quadratic) bound on the number of tangencies among such families of curves. In [**110**] a bound of $\mathcal{O}(n^{2-1/7633})$ was shown for characteristic at least 3. (It is shown in [**110**] that in characteristic 2 there are two infinite families of conic sections such that each conic from the first family touches all conics from the other family, hence a quadric number of tangencies is possible.) Here — as an application of the notion of "combinatorial dimension" using Proposition 1.2.9 and Lemma 1.2.13 — we prove the following.

Corollary 1.5.1. *The number of tangencies is $\mathcal{O}(n^{9/5})$. If the characteristic of the base field is zero, then the exponent can be improved to any number greater than $\frac{139}{79}$.*

Proof Let X be the five dimensional projective space parameterizing conics in the plane. For each non-degenerate conic $p \in X$ the locus $T_p \subset X$ of the conics tangent to p is a hypersurface of degree six in X, whose equation can be determined as follows: A general conic q has a two-variable quadric equation with 6 unknown coefficients (6 parameters, these are the coordinates on X). The intersection $p \cap q$ can be calculated from the resultant R of the equations of p and q, which is a one-variable degree 4 polynomial, the coefficients of R are expressions of the 6 parameters. Now q is tangent to p iff two of the four intersection points coincide, i.e. iff R has a double root. Double roots are detected by the vanishing of the discriminant. The

discriminant $D(R)$ is an expression of the coefficients of R, hence a polynomial of the 6 parameters. The locus $T_p \subset X$ is precisely the zero set of $D(R)$. The precise degree of $D(R)$ is not important for us, only that it is bounded independently from p, which is clear from the description.

Let $\mathcal{H}$ denote the collection of these T_p as $p \in X$ varies. The locus of those conics which are tangent to a given line at a given point is a three dimensional linear subspace of the projective space X, these subspaces are our forbidden varieties. It is proved in [**110**] that if two infinite sets of conics have the property that each conic in one set is tangent to all conics in the other set, then in fact all these conics are tangent to the same line at the same point, i.e. they are in a forbidden variety.

Our arrangement of n conics gives us a subset $\mathcal{P} \subseteq X$ of n points in general position, and n corresponding members of $\mathcal{H}$. Lemma 1.2.13 implies that for absolute constant b the combinatorial dimension $\mathrm{cdim}_b(\mathcal{P}, \mathcal{H})$ is 5, and Proposition 1.2.9 gives us the bound $\mathcal{O}(n^{9/5})$. In characteristic zero we use Theorem 1.2.6. In this case $k = \dim(Y) = 5$, hence $D = 4d - 4 = 16$, and our exponent is

$$\alpha + \beta = \frac{D(k-1) + k(D-1)}{Dk - 1} + 4\varepsilon = \frac{139}{79} + 4\varepsilon\,.$$

This proves the corollary. ∎

Triple Points of Circle Grids. Let $n \geq 1$ be an integer and P_1, P_2, P_3 three distinct points in the Euclidean plane $\mathbb{R}^2$; finally, let $\mathcal{R}$ be a (finite) set of distinct positive reals (they will be radii of circles). For the set of those points P which are covered by three circles around the P_i of radii $r_i \in \mathcal{R}$, we shall use the notation

$$T(P_1, P_2, P_3, \mathcal{R}) \stackrel{\mathrm{def}}{=} \{P \in \mathbb{R}^2 \;\big|\; \forall i = 1, 2, 3 \;\; |P_i P| \in \mathcal{R}\}$$

and call these points *triple points* with respect to the $n + n + n$ circles (see Figure 1.5.1 for a configuration with many triple points).

Also, we denote

$$t(P_1, P_2, P_3, n) \stackrel{\mathrm{def}}{=} \max_{|\mathcal{R}|=n} \big|T(P_1, P_2, P_3, \mathcal{R})\big|.$$

Question 1.5.2. Is $t(P_1, P_2, P_3, n) \geq cn^2$ possible for a fixed $c > 0$, three non-collinear points P_i, and infinitely many n?

This problem — without the non-collinearity condition — was raised by Erdős, Lovász and Vesztergombi [**51**]. As for collinear P_i, some structures called "circle grids" demonstrate that t can be as large as $n^2/2$ ([**39**], Lemma 5). The construction given there is simple: let $P_1 = (-1, 0)$, $P_2 = (0, 0)$, $P_3 = (1, 0)$ and $r_i = \sqrt{i}$ (see Figure 1.5.1). We shall call such patterns of $n + n + n$ circles *circle grids* since they look like (bent) grid lines and diagonals of a square or rectangular grid.

Here we prove that t has a significant "jump". More specifically, we show that, in case of non-collinear centers, a circle grid cannot have more than $n^{2-\eta}$ triple points, provided that n is large enough.

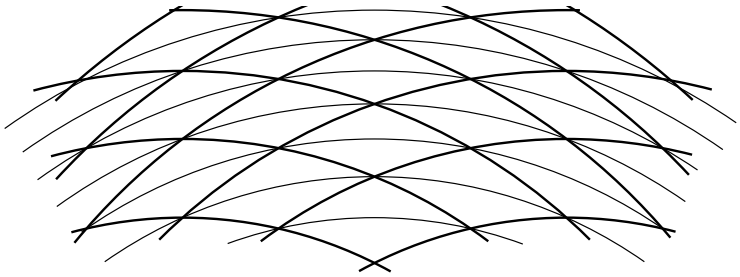

FIGURE 1.5.1. Part of a circle grid with collinear centers.

Theorem 1.5.3. *There exist absolute constants $\eta > 0$ and a bound $n_1 \in \mathbb{N}$ such that, for any non-collinear triple $P_1, P_2, P_3 \in \mathbb{R}^2$ and all $n > n_1$,*

$$t(P_1, P_2, P_3, n) \leq n^{2-\eta}.$$

Proof We show that $\eta = \eta(2)$ and $n_1 \stackrel{\text{def}}{=} n_0(2)$ found in Theorem 1.1.3 satisfy the statement of Theorem 1.5.3, too.

Assume that

$$t(P_1, P_2, P_3, n) > n^{2-\eta}, \tag{1.5.1}$$

for three points P_1, P_2, P_3 and $n > n_1 = n_0(2)$ radii. Without loss of generality we may assume that $P_1 = (1, 0)$ and $P_2 = (-1, 0)$, otherwise we can rotate and shrink/magnify the configuration. Also, put $P_3 = (a, b)$. We must show $b = 0$.

For any $P(u, v) \in \mathbb{R}^2$, consider the *squares* of the distances $\overline{P_i P}$ and write

$$\begin{aligned} X &\stackrel{\text{def}}{=} (u-1)^2 + v^2 \\ Y &\stackrel{\text{def}}{=} (u+1)^2 + v^2 \\ Z &\stackrel{\text{def}}{=} (u-a)^2 + (v-b)^2. \end{aligned} \tag{1.5.2}$$

Of course, they are not independent of each other; using a computer algebra system one can easily find a quadratic polynomial relation which these distances satisfy (and it is also not too difficult to calculate it by hand):

$$\begin{aligned} F(X, Y, Z) = &\big((a+1)^2 + b^2\big)X^2 + \big((a-1)^2 + b^2\big)Y^2 + 4Z^2 \\ &- 2(a^2 + b^2 - 1)XY - 4(a+1)XZ + 4(a-1)YZ \\ &+ 4(a-1)\big((a+1)^2 + b^2\big)X - 4(a+1)\big((a-1)^2 + b^2\big)Y \\ &- 8(a^2 + b^2 - 1)Z + 4\big(b^4 + 2a^2b^2 + (a^2-1)^2 + 2b^2\big) \\ &= 0. \end{aligned}$$

If $b \neq 0$ then the quadratic polynomial F is irreducible. Indeed, no linear relation $\alpha X + \beta Y + \gamma Z + \delta = 0$ can be satisfied by X, Y and Z as in (1.5.2). (Considering the v–term shows $\gamma = 0$; moreover, $\alpha = \beta$ by the u–term while $\alpha = -\beta$ from the u^2–term.) Hence F cannot be factored into linear terms.

On the other hand, if $b = 0$, then the equation $F = 0$ reduces to (the square of)

$$(1 + a)X + (1 - a)Y - 2Z - 2(1 - a^2) = 0,$$

thus, in this case, it is not difficult to find appropriate radii for which $t \geq cn^2$, for a suitable $c > 0$. E.g.,

$$\begin{aligned} X &= \left\{ \frac{1}{1+a}, \frac{2}{1+a}, \ldots, \frac{n}{1+a} \right\}, \\ Y &= \left\{ \frac{1}{1-a}, \frac{2}{1-a}, \ldots, \frac{n}{1-a} \right\}, \\ Z &= \big\{1, 2, \ldots, n\big\} + a^2 - 1, \quad \text{i.e.,} \\ &= \big\{a^2, a^2 + 1, \ldots, a^2 + n - 1\big\} \end{aligned}$$

will work well.

For our purposes it is no use writing F explicitly; rather, we shall consider the surface

$$S \stackrel{\text{def}}{=} \{(X, Y, Z) \in \mathbb{R}^3 \mid F(X, Y, Z) = 0\},$$

parametrised by u and v as in (1.5.2).

Denote by T the set $T = \{r_i^2 \mid r_i \in \mathcal{R}\}$, i.e., the squares of the n radii which demonstrate (1.5.1). Then S contains more than $n^{2-\eta}$ points of $T \times T \times T$. Of course, we want to apply Theorem 1.1.3. In order to do so, we first note that F really depends on all three variables X, Y and Z. Therefore, S is not a cylinder over any curve and thus, by Theorem 1.1.3 applied to $V = S$, there exists one-to-one analytic functions $f, g, h : (-1, 1) \to \mathbb{R}$ such that

$$S \supset \Big\{ \big(f(x), g(y), h(z)\big) \in \mathbb{R}^3 \mid -1 < x, y, z < 1,\ x + y + z = 0 \Big\}.$$

But the right hand side is a surface too, hence the two surfaces must agree in some open set. So there exist an open set $U \subset \mathbb{R}^3$ such that, after a linear change of variable in the functions f, g and h (in order to get a small portion of the surface on the right hand side), we get

$$S \cap U = \Big\{ \big(f(x), g(y), h(z)\big) \in \mathbb{R}^3 \mid -1 < x, y, z < 1,\ x + y + z = 0 \Big\}.$$

Equivalently, $(X, Y, Z) \in S \cap U$ iff

$$Z = Z(X, Y) = h\big(-f^{-1}(X) - g^{-1}(Y)\big), \tag{1.5.3}$$

and this holds in a nonempty open subset of the X, Y plane.

In order to show that this is only possible if $b = 0$, we follow the method used in [**45**]. The Lemma below will form the basis of this test.
In what follows we put

$$q(X, Y) \stackrel{\text{def}}{=} \frac{Z'_X}{Z'_Y}.$$

Lemma 1.5.4. *(i) With the above notation, if Z is as in* (1.5.3)*, then*

$$q = (f^{-1})'(X)/(g^{-1})'(Y),$$

and, as a consequence, it is the product (quotient) of two analytic functions which only depend on X and Y, respectively.

(ii) $q_1 \stackrel{\text{def}}{=} (\log|q|)'_X$ *only depends on* X.

(iii) $q_2 \stackrel{\text{def}}{=} (q_1)'_Y = (\log|q|)''_{XY} = 0$ *identically, on a non–empty open subset of* $\mathbb{R}^2$.

Proof (i) is obvious and so are (i)⇒(ii) and (ii)⇒(iii). ∎

For the forthcoming numeric computations, we consider X, Y, Z parametrised by u, v, as in (1.5.2). We must first find the partial derivatives of u, v and Z by X and Y. If Y is a constant then, by differentiating (1.5.2),

$$\begin{aligned} 1 &= 2(u-1)u'_X + 2vv'_X \\ 0 &= u'_X(u+1) + v'_X v \\ Z'_X &= 2(u-a)u'_X + 2(v-b)v'_X \end{aligned}$$

whence

$$u'_X = -\frac{1}{4}, \quad v'_X = \frac{1}{4}\frac{u+1}{v}, \quad Z'_X = -\frac{1}{2}\frac{-va - v + bu + b}{v}. \tag{1.5.4}$$

Similarly, if X is a constant, then

$$u'_Y = \frac{1}{4}, \quad v'_Y = -\frac{1}{4}\frac{u-1}{v}, \quad Z'_Y = \frac{1}{2}\frac{-va + v + bu - b}{v}. \tag{1.5.5}$$

Hence

$$q = -\frac{-va - v + bu + b}{-va + v + bu - b}$$

and, after suitable substitutions and simplifications, we have

$$q_1 = -\frac{1}{2}\frac{b(u^2+v^2-(a+1)u - bv + a)}{v(bu-(a-1)v-b)(bu-(a+1)v+b)} = \frac{-bu^2+\dots}{2b^2u^2v+\dots}.$$

Finally, in the numerator of $q_2 = (q_1)'_Y$, there will be exactly one term which contains u^5. It comes from

$$-(-bu^2+\dots)\cdot(2b^2u^2v+\dots)'_Y = (2b^3u^4v'_Y+\dots) = -2b^3u^4\cdot\left(-\frac{1}{4}\frac{u-1}{v}\right)+\dots,$$

and thus this term is $2b^3u^5$. Since, by Lemma 1.5.4(iii), q_2 must vanish on a non–empty open set, we can really infer that $b = 0$. ∎

dc_650_12

CHAPTER 2

Growth in finite simple groups of Lie type

2.1. Introduction

The *diameter*, $\mathrm{diam}(X)$, of an undirected graph $X = (V, E)$ is the largest distance between two of its vertices.

Given a subset A of the vertex set V the expansion of A, $c(A)$, is defined to be the ratio $|\sigma(A)|/|A|$ where $\sigma(A)$ is the set of vertices at distance 1 from A. A graph is a *C-expander* for some $C > 0$ if for all sets A with $|A| < |V|/2$ we have $c(A) \geq C$. A family of graphs is an *expander family* if all of its members are C-expanders for some fixed positive constant C.

Let G be a finite group and S a symmetric (i.e. inverse-closed) set of generators of G. The *Cayley graph* $\Gamma(G, S)$ is the graph whose vertices are the elements of G and which has an edge from x to y if and only if $x = sy$ for some $s \in S$. Then the diameter of Γ is the smallest number d such that $S^d = G$.

The following classical conjecture is due to Babai [**7**]

Conjecture 2.1.1 (Babai)**.** *For every non-abelian finite simple group L and every symmetric generating set S of L we have* $\mathrm{diam}\big(\Gamma(L, S)\big) \leq C\big(\log |L|\big)^c$ *where c and C are absolute constants.*

In a spectacular breakthrough Helfgott [**72**] proved that the conjecture holds for the family of groups $L = PSL(2, p)$, p a prime. In recent major work [**73**] he proved the conjecture for the groups $L = PSL(3, p)$, p a prime. Dinai [**36**] and Varjú [**165**] have extended Helfgott's original result to the groups $PSL(2, q)$, q a prime power.

We prove the following.

Theorem 2.1.2. *Let L be a finite simple group of Lie type of rank r. For every symmetric set S of generators of L we have*

$$\mathrm{diam}\big(\Gamma(L, S)\big) < \big(\log |L|\big)^{c(r)}$$

where the constant $c(r)$ depends only on r.

This settles Babai's conjecture for any family of simple groups of Lie type of bounded rank.

A key result of Helfgott [**72**] shows that generating sets of $SL(2, p)$ grow rapidly under multiplication. His bound on diameters is an immediate consequence.

Theorem 2.1.3 (Helfgott)**.** *Let $L = SL(2, p)$ and A a generating set of L. Let δ be a constant, $0 < \delta < 1$.*
a) Assume that $|A| < |L|^{1-\delta}$. Then

$$|A^3| \gg |A|^{1+\varepsilon}$$

where ε and the implied constant depend only on δ
b) Assume that $|A| > |L|^{1-\delta}$. Then $A^k = L$ where k depends only on δ.

It was observed in [**113**] that a result of Gowers [**63**] implies that b) holds for an arbitrary simple group of Lie type L with $k = 3$ for some $\delta(r)$ which depends only on the Lie rank r of L (see [**6**] for a more detailed discussion). Hence to complete the proof of our theorem on diameters it remains to prove an analogue of the (rather more difficult) part a) as was done by Helfgott for the groups $SL(3, p)$ in [**73**].

We prove the following.

Theorem 2.1.4 (Product theorem)**.** *Let L be a finite simple group of Lie type of rank r and A a generating set of L. Then either $A^3 = L$ or*

$$|A^3| \gg |A|^{1+\varepsilon}$$

where ε and the implied constant depend only on r.

We also give some examples which show that in the above result the dependence of ε on r is necessary. In particular we construct generating sets A of $SL(n, 3)$ of size $2^{n-1} + 4$ with $|A^3| < 100|A|$ for $n \geq 3$.

Thre Product theorem was first announced in [**132**]. The same day similar results were announced by Breuillard, Green and Tao [**24**] for finite Chevalley groups. It is noted in [**24**] that their methods are likely to extend to all simple groups of Lie type, but this has not yet been checked. On the other hand in [**24**] various interesting results for complex matrix groups were also announced.

Somewhat earlier Gill and Helfgott [**60**] had shown that small generating sets (of size at most $p^{n+1-\delta}$ for some $\delta > 0$) in $SL(n, p)$ grow.

Helfgott's work [**72**] has been the starting point and inspiration of much recent work by Bourgain, Gamburd, Sarnak and others. Let $S = \{g_1, g_2, \ldots, g_k\}$ be a symmetric subset of $SL(n, \mathbb{Z})$ and $\Lambda = \langle S \rangle$ the subgroup generated by S. Assume that Λ is Zariski dense in $SL(n)$. According to the theorem of Matthews-Vaserstein-Weisfeiler [**109**] there is some integer m_0 such that $\pi_m(\Lambda) = SL(n, \mathbb{Z}/m\mathbb{Z})$ assuming $(m, m_0) = 1$. Here π_m denotes reduction mod m.

It was conjectured in [**105**], [**14**] that the Cayley graphs $\Gamma\big(SL(n, \mathbb{Z}/m\mathbb{Z}), \pi_m(S)\big)$ form an expander family, with expansion constant bounded below by a constant $c = c(S)$. This was verified in [**12**], [**11**], [**14**] in many cases when $n = 2$ and in [**13**] for $n > 2$ and moduli of the form p^d where $d \to \infty$ and p is a sufficiently large prime.

In [**13**] Bourgain and Gamburd also prove the following

Theorem 2.1.5 (Bourgain, Gamburd)**.** *Assume that the analogue of Helfgott's theorem on growth holds for $SL(n, p)$, p a prime. Let S be a symmetric finite subset of $SL(n, \mathbb{Z})$ generating a subgroup Λ which is Zariski dense in $SL(n)$. Then the family of Cayley graphs $\Gamma(SL(n, p), \pi_p(S))$ forms an expander family as $p \to \infty$. The expansion coefficients are bounded below by a positive number $c(S) > 0$.*

By the Product theorem (2.1.4) the condition of this theorem is satisfied hence the above conjecture is proved for prime moduli.

For $n = 2$ Bourgain, Gamburd and Sarnak [**14**] proved that the conjecture holds for square free moduli. This result was used in [**14**] as a building block in a combinatorial sieve method for primes and almost primes on orbits of various subgroups of $GL(2, \mathbb{Z})$ as they act on $\mathbb{Z}^m$ (for $m \geq 2$).

Recently, extending Theorem 2.1.5 P. Varjú [**165**] has shown that if the analogue of Helfgott's theorem holds for $SL(n, p)$, p a prime, then the above conjecture holds for square free moduli and Zariski dense subgroups of $SL(n)$. Hence our results constitute a major step towards obtaining a generalisation to Zariski dense subgroups of $SL(n, \mathbb{Z})$ and to other arithmetic groups.[1]

Simple groups of Lie type can be treated as subgroups of simple algebraic groups. In fact, instead of concentrating on simple groups, we work in the framework of arbitrary linear algebraic groups over algebraically closed fields. We set up a machinery which can be used to obtain various results on growth of subsets in linear groups. In particular, we prove the following extension of the Product theorem (2.1.4), valid for finite groups obtained from connected linear groups over $\overline{\mathbb{F}_p}$, which produces growth within certain normal subgroups (for the terminology see Definition 2.11.1).

Theorem 2.1.6. *Let G be a connected linear algebraic group over $\overline{\mathbb{F}_p}$ and $\sigma : G \to G$ a Frobenius map. Let G^σ denote the subgroup of the fixpoints of σ and $1 \in S \subseteq G^\sigma$ a symmetric generating set. Then for all $1 > \varepsilon > 0$ there is an integer $M = M_{\text{main}}\big(\dim(G), \varepsilon\big)$ and a real K depending on ε and the numerical invariants of G (notably $\dim(G)$, $\deg(G)$, $\text{mult}(G)$ and $\text{inv}(G)$, see Definition 2.5.1) with the following property. If $\mathcal{Z}(G)$ is finite and*

$$K \leq |S| \leq |G^\sigma|^{1-\varepsilon}$$

then there is a connected closed normal subgroup $H \triangleleft G$ such that $\deg H \leq K$, $\dim(H) > 0$ and

$$|S^M \cap H| \geq |S|^{(1+\delta)\dim(H)/\dim(G)}$$

where $\delta = \frac{\varepsilon}{128\dim(G)^3}$.

Consider the groups G^σ for simply connected simple algebraic groups G. Central extensions of all but finitely many simple groups of Lie type are obtained in this way (see [**149**]) and the centres $\mathcal{Z}(G^\sigma)$ have bounded order. Hence Theorem 2.1.6 implies the Product theorem (2.1.4) for both twisted and untwisted simple groups of Lie type in a unified way.

The proof of Theorem 2.1.6 relies basically on two properties of the finite groups G^σ. First, if G^σ is large enough then $\mathcal{C}_G(G^\sigma) = \mathcal{Z}(G)$. Second, if a σ-invariant connected closed subgroup of G is normalised by G^σ then it is in fact normal in G. In this generality Theorem 2.1.6 depends on Hrushovski's twisted Lang-Weil estimates [**76**]. In the proof of the Product theorem (2.1.4) this can be avoided (see Remark 2.11.6). Hence the constants in this theorem are explicitly computable.

We believe that Theorem 2.1.6 and the general results concerning algebraic groups involved in its proof will have many applications to investigating growth in linear groups. Here we first prove (using Theorem 2.1.6) the following partial extension of the Product theorem (2.1.4):

[1] Finally the conjecture has very recently been proved by Bourgain and Varjú [**17**].

Theorem 2.1.7. *Let S be a symmetric subset of $GL(n,p)$ satisfying $|S^3| \leq K|S|$ for some $K \geq 1$. Then $GL(n,p)$ has two subgroups $H \geq P$, both normalised by S, such that P is perfect, H/P is soluble, P is contained in S^6 and S is covered by $K^{c(n)}$ cosets of H where $c(n)$ depends on n.*

Understanding the structure of symmetric subsets S of $GL(n,p)$ (or more generally of $GL(n,q)$, q a prime-power) satisfying $|S|^3 \leq K|S|$ is mentioned by Breuillard, Green and Tao as a difficult open problem in [**24**].

Subgroups of $GL(n,p)$ generated by elements of order p were investigated in detail by Nori [**115**] and Hrushovski-Pillay [**78**]. As a byproduct of the proof of Theorem 2.1.7 we obtain the following.

Theorem 2.1.8. *Let $P \leq GL(n,p)$, p a prime, be a perfect subgroup which is generated by its elements of order p. Let S be a symmetric set of generators of P. Then*

$$\operatorname{diam}\big(\Gamma(P,S)\big) \leq \big(\log|P|\big)^{M(n)}$$

where the constant $M(n)$ depends only on n.

Theorem 2.1.8 is a surprising extension of the fact (included in Theorem 2.1.2) that simple subgroups of $GL(n,p)$ (n bounded) have polylogarithmic diameter.

Combining Theorem 2.1.8 with results of Aldous [**1**] and Babai [**3**] we immediately obtain the following corollary.

Corollary 2.1.9. *Let $\Gamma = \Gamma(P,S)$ be a Cayley graph as in Theorem 2.1.8. Then Γ is a C-expander with some*

$$C \geq \frac{1}{1+\big(\log|P|\big)^{M(n)}} \,.$$

Equivalently, if A is a subset of P of size at most $|P|/2$, then we have

$$|A \cdot S| \geq (1+C)|A| \,.$$

For a very recent unexpected application in arithmetic geometry of the above corollary see [**49**].

To indicate the generality of our methods we derive the following consequence.

Theorem 2.1.10. *Let $\mathbb{F}$ be an arbitrary field and $S \subseteq GL(n,\mathbb{F})$ a finite symmetric subset such that $\big|S^3\big| \leq K|S|$ for some $K \geq \frac{3}{2}$. Then there are normal subgroups $H \leq \Gamma$ of $\langle S\rangle$ and a bound m depending only on n such that $\Gamma \subseteq S^6H$, the subset S can be covered by K^m cosets of Γ, H is soluble, and the quotient group Γ/H is the product of finite simple groups of Lie type of the same characteristic as $\mathbb{F}$. (In particular, in characteristic 0 we have $\Gamma = H$.) Moreover, the Lie rank of the simple factors appearing in Γ/H is bounded by n, and the number of factors is also at most n.*

This theorem may be viewed as a common generalisation of the Product theorem (2.1.4) and a result of Hrushovski [**77**] obtained by model-theoretic tools. It would be most interesting to obtain a result that would also imply Theorem 2.1.7.

The first result of this type was obtained by Elekes and Király [**43**]. In characteristic 0 the above theorem was first proved by Breuillard, Green and

Tao [**25**]. Actually in that case they have a stronger conclusion: one can even require $\Gamma = H$ to be nilpotent.

Methods used in Chapter 2. The proofs of Helfgott combine group theoretic arguments with some algebraic geometry, Lie theory and tools from additive combinatorics such as the sum-product theorem of Bourgain, Katz, Tao [**16**]. Our argument relies on a deeper understanding of the algebraic group theory behind his proofs and an extra trick, but not on additive combinatorics.

We prove various results which say that if L is a "nice" subgroup of an algebraic group G generated by a set A then A grows in some sense. These were motivated by earlier results of Helfgott [**72**], [**73**] and Hrushovski-Pillay [**78**].

To illustrate our strategy we outline the proof of the Product theorem (2.1.4) in the simplest case, when A generates $L = SL(n, q)$, q a prime-power. Assume that "A does not grow" i.e. $|AAA|$ is not much larger than $|A|$. Using an "escape from subvarieties" argument it is shown in [**73**] that if T is a maximal torus in L then $|T \cap A|$ is not much larger than $|A|^{1/(n+1)}$. This is natural to expect for dimensional reasons since $\dim(T)/\dim(L) = (n-1)/(n^2-1) = 1/(n+1)$.

We use a rather more powerful escape argument. The first part of Chapter 2 is devoted to establishing the necessary tools in great generality (in particular Theorem 2.6.8).

Now T is equal to $L \cap \bar{T}$ where $\bar{T}$ is a maximal torus of the algebraic group $SL(n, \overline{\mathbb{F}_q})$. Let T_r denote the set of regular semisimple elements in T. Note that $T \setminus T_r$ is contained in a subvariety $V \subsetneq \bar{T}$ of dimension $n-2$. By the above mentioned escape argument $\left|(T \setminus T_r) \cap A\right|$ is not much larger than

$$|A|^{\dim(V)/\dim(L)} = |A|^{1/(n+1)-1/(n^2-1)} .$$

By [**73**] or by our escape argument A does contain regular semisimple elements. If a is such an element then consider the map $SL(n) \to SL(n)$, $g \to g^{-1}ag$. The image of this map is contained in a subvariety of dimension $n^2 - 1 - (n-1)$ since $\dim\left(\mathcal{C}_{SL(n)}(a)\right) = n-1$. By the escape argument we obtain that for the conjugacy class $\mathrm{cl}(a)$ of a in L, $\left|\mathrm{cl}(a) \cap A^{-1}aA\right|$ is not much larger than $|A|^{(n^2-n)/(n^2-1)}$. Now $\left|\mathrm{cl}(a) \cap A^{-1}aA\right|$ is at least the number of cosets of the centraliser $C_L(a)$ which contain elements of A . It follows that $\left|AA^{-1} \cap C_L(a)\right|$ is not much smaller than $|A|^{1/(n+1)}$. Of course $C_L(a)$ is just the (unique) maximal torus containing a.

Let us say that A *covers* a maximal torus T if $\left|T \cap A\right|$ contains a regular semisimple element. We obtain the following fundamental dichotomy (see Lemma 2.9.2):

Assume that a generating set A does not grow

i) If A does not cover a maximal torus T then $\left|T \cap A\right|$ is not much larger than $|A|^{1/(n+1)-1/(n^2-1)}$.

ii) If A covers T then $\left|T \cap AA^{-1}\right|$ is not much smaller than $|A|^{1/(n+1)}$. In this latter case in fact $\left|T_r \cap AA^{-1}\right|$ is not much smaller than $|A|^{1/(n+1)}$.

It is well known that if A doesn't grow then $B = AA^{-1}$ doesn't grow either hence the above dichotomy applies to B.

Let us first assume that B covers a maximal torus T but does not cover a conjugate $T' = g^{-1}Tg$ of T for some element g of L . Since A generates L we have such a pair of conjugate tori where g is in fact an element of A. Consider those cosets of T' which intersect A. Each of the, say, t cosets contains at most $|B \cap T'|$ elements of A i.e. not much more than $|B|^{1/(n+1)-1/(n^2-1)}$ which in turn is not much more than $|A|^{1/(n+1)-1/(n^2-1)}$. Therefore $|A|$ is not much larger than $t|A|^{1/(n+1)-1/(n^2-1)}$.

On the other hand $A\big(A^{-1}(BB^{-1})A\big)$ has at least $t\big|T \cap BB^{-1}\big|$ elements which is not much smaller than $t|A|^{1/(n+1)}$. Therefore $A\big(A^{-1}(AA^{-2}A)A\big)$ is not much smaller than $|A|^{1+1/(n^2-1)}$ which contradicts the assumption that A does not grow.

We obtain that B covers all conjugates of some maximal torus T. Now the conjugates of the set T_r are pairwise disjoint (e.g. since two regular semisimple elements commute exactly if they are in the same maximal torus). The number of these tori is $|L : N_L(T)| > c(n)|L : T|$ for some constant which depends only on n. Each of them contains not much less than $|B|^{1/(n+1)}$ regular semisimple elements of BB^{-1}. Altogether we see that $|A|$ is not much smaller than $q^{n^2-n}|A|^{1/(n+1)}$ and finally that $|A|$ is not much less than $|L|$. In this case by [**113**] we have $AAA = L$.

The proof of Theorem 2.1.6 follows a similar strategy. However there is an essential difference; maximal tori have to be replaced by a more general class of subgroups called CCC-subgroups (see Definition 2.8.6). These subgroups were in fact designed to make the argument work in not necessarily simple (or semisimple) algebraic groups. In Sections 2.8, 2.9 and 2.10 we establish the basic properties of these subgroups and justify that they indeed play the role of maximal tori in general algebraic groups. The proof of Theorem 2.1.6 is completed in Section 2.13.

In [**115**] Nori showed that if p is sufficiently large in terms of n, there is a correspondence between subgroups of $GL(n, p)$ generated by elements of order p and a certain class of closed subgroups of $GL(n, \overline{\mathbb{F}_p})$. Note that the bounds in [**115**] are ineffective. Using this correspondence Theorem 2.1.7 is proved for perfect p-generated groups by a short induction argument based on a slight extension of Theorem 2.1.6. The general case can be reduced to this by applying various known results on finite linear groups.

Theorem 2.1.10 follows by combining some of the ingredients of the proof of Theorem 2.1.7 in a rather more direct way.

Examples given in Section 2.14 show that in the Product theorem (2.1.4) we must have $\varepsilon(r) = O(1/r)$. We believe that this is the right order of magnitude.

2.2. Notation

Throughout this chapter $\overline{\mathbb{F}}$ denotes an arbitrary algebraically closed field. For a prime number p we denote by $\mathbb{F}_p$ and $\overline{\mathbb{F}_p}$ the finite field with p elements and its algebraic closure. Similarly, $\mathbb{F}_q$ denotes the finite field with q elements, where q is a prime power. The letters N and Δ will always be used for an upper bound for dimensions and degrees respectively, K is used for a lower bound on the size of certain finite sets. When we study growth,

M will denote the length of the products we allow. In several lemmas we use a parameter ε, it is the error-margin we allow in the exponents when we count elements in certain subsets.

2.3. Dimension and degree

We use affine algebraic geometry i.e. all occurring sets will be subsets of some *affine space* $\overline{\mathbb{F}}^m$ for some integer $m > 0$, and we define all of them via m-variate polynomials whose coefficients belong to $\overline{\mathbb{F}}$. Below we make this more precise.

Definition 2.3.1. A subset $Z \subseteq \overline{\mathbb{F}}^m$ is *Zariski closed*, or simply *closed*, if it can be defined as the common zero set of some m-variate polynomials. This defines a topology on $\overline{\mathbb{F}}^m$, each subset of $\overline{\mathbb{F}}^m$ inherits this topology, called the *Zariski topology*. This is the only topology that we use in Chapter 2, so we omit the adjective Zariski. The complements of closed subsets are called *open*, The intersection of a closed and an open subset is called *locally closed*. If we do not use explicitly the ambient affine space then locally closed subsets are called *algebraic sets* and closed subsets are called *affine algebraic sets*. For an arbitrary subset $X \subseteq \overline{\mathbb{F}}^m$ we denote by $\overline{X}$ the *closure* of X.

Our algebraic sets are subsets of the affine space $\overline{\mathbb{F}}^m$. One can define algebraic subsets in more general spaces, e.g. in the projective space $\overline{\mathbb{F}}\mathbb{P}^m$. However, in this chapter, we do not use such generality.

Note, that algebraic sets are always equipped (by definition) with an ambient affine space, even if it is not explicitely given. This is one reason for choosing the name "algebraic set" instead of "variety".

Definition 2.3.2. An algebraic set X is called *irreducible* if it has the following property. Whenever X is contained in the union of finitely many closed subsets, it must be contained in one of them.

Definition 2.3.3. Let X be an algebraic set. Then there are finitely many closed subsets $X_i \subseteq X$ which are irreducible, and maximal among the irreducible closed subsets of X. Then $X = \bigcup_i X_i$ is the *irreducible decomposition* of X and these X_i are called the *irreducible components* of X.

Definition 2.3.4. Let $Z \subseteq \overline{\mathbb{F}}^m$ be an algebraic set. We consider chains $Z_0 \subsetneq Z_1 \subsetneq \cdots \subsetneq Z_n$ where the Z_i are nonempty, irreducible closed subsets of Z. The largest possible length n of such a chain is called the *dimension* of Z, denoted by $\dim(Z)$.

Definition 2.3.5. Let $X \subseteq \overline{\mathbb{F}}^m$ be an algebraic set. An *affine subspace* of $\overline{\mathbb{F}}^m$ is a translate of a linear subspace. If X is irreducible then we consider all affine subspaces $L \subseteq \overline{\mathbb{F}}^m$ such that $\dim(X) + \dim(L) = m$ and $X \cap L$ is finite. The *degree* of X is the largest possible number of intersection points:

$$\deg(X) = \max_L |X \cap L| .$$

In general, the degree of X is defined as the sum of the degrees of its irreducible components.

Remark 2.3.6. Let X be an algebraic set. Then $\dim(X) = 0$ iff X is finite. A finite subset $X \subset \overline{\mathbb{F}}^m$ is always closed, and satisfies $\deg(X) = |X|$.

Definition 2.3.7. Let $X \subseteq \overline{\mathbb{F}}^m$ and $Y \subseteq \overline{\mathbb{F}}^n$ be algebraic sets. A function $f : X \to Y$ is called a *morphism*if it is the restriction to X of a map $\phi : \overline{\mathbb{F}}^m \to \overline{\mathbb{F}}^n$ whose n coordinates are m-variate polynomials. Then the *graph of* f, denoted by $\Gamma_f \subseteq X \times Y \subseteq \overline{\mathbb{F}}^{m+n}$, is locally closed. We define the *degree* of f to be $\deg(f) = \deg(\Gamma_f)$.

Remark 2.3.8. Algebraic sets form a category with the above notion of morphism. Isomorphic algebraic sets have equal dimensions and isomorphisms respect the irreducible decomposition. In contrast, the degrees of isomorphic algebraic sets may not be be equal.

In Chapter 2 we work mainly in the category of algebraic sets and morphisms. To obtain explicit bounds we need to estimate the degrees of all appearing objects. If one is satisfied with existence results only then one can avoid all these calculations by simply noticing that all of our constructions can be done simultaneously in families of algebraic sets. (Such proofs a priori do not give explicit constants, but with careful examination, in principle they can be made explicit.) In fact this technique is really used e.g. in the proof of Proposition 2.12.8.

The following fact is standard:

Fact 2.3.9. *Let $X, Y \subseteq \overline{\mathbb{F}}^m$ be locally closed sets.*

(a) The dimension and the degree of X are equal to the dimension and the degree of its closure $\overline{X}$.

(b) Any closed subset of X has dimension at most $\dim(X)$.

(c) The irreducible components $X_i \leq X$ satisfy

$$\dim(X_i) \leq \dim(X) = \max_j \big(\dim(X_j)\big) ,$$

$$\deg(X_i) \leq \deg(X) = \sum_j \deg(X_j) .$$

It follows that there are at most $\deg(X)$ components and at least one of them has the same dimension $\dim(X_i) = \dim(X)$.

(d) The sets $X \cap Y$, $\overline{X} \cup \overline{Y}$, $X \setminus \overline{Y}$ and $X \times Y$ are also locally closed with the following bounds:

$$\begin{aligned}
\dim(\overline{X} \cup \overline{Y}) &= \max\big(\dim(X), \dim(Y)\big) \\
\deg(\overline{X} \cup \overline{Y}) &\leq \deg(X) + \deg(Y) \\
\dim(X \cap Y) &\leq \min\big(\dim(X), \dim(Y)\big) \\
\deg(X \cap Y) &\leq \deg(X)\deg(Y) \\
\dim(X \setminus \overline{Y}) &\leq \dim(X) \\
\dim(X \times Y) &= \dim(X) + \dim(Y) \\
\deg(X \times Y) &= \deg(X)\deg(Y)
\end{aligned}$$

Note that we cannot estimate $\deg(X \setminus \overline{Y})$ in this generality.

(e) Suppose that X is irreducible. Then each nonempty open subset $U \subset X$ is dense in X with $\dim(X \setminus U) < \dim(X)$ (and we do not bound the degree of $X \setminus U$).

(f) The direct product of irreducible algebraic sets is again irreducible.

(g) *If X is the common zero locus of degree d polynomials, then it is the common zero locus of at most $(d+1)^m$ of them, and* $\deg(X) \leq d^m$. *On the other hand, a closed set X is the common zero locus of polynomials of degree at most* $\deg(X)$.

Most of this Fact is proved in [**71**, Chapters I.1 and II.3]. The bound on $\deg(X \cap Y)$ is (an appropriate version of) Bézout's theorem (see [**57**]) and (g) follows from [**91**, Section I.3].

We also need the following:

Fact 2.3.10. *Let X and Y be affine algebraic sets and $f : X \to Y$ a morphism. We define several (open, closed or locally closed) subsets of X and Y. Their dimension is at most* $\dim(X)$, *and we bound their degrees from above. We define the function* $\Phi(d) = (d+2)^{(d+1)^{\dim(X)+\deg(f)}2^d}$ *and the constant* $D = \Phi\big(\Phi\big(\dots\Phi\big(\deg(f)\big)\big)\dots\big)^{\dim(X)+\deg(f)}$ *where the function Φ is iterated* $\dim(X)+\deg(f)-1$ *times.*

(a) *There is a partition of $\overline{f(X)}$ into at most D locally closed subsets Y_i of degree at most D such that the closure of each Y_i is the union of partition classes and either $f^{-1}(Y_i) = \emptyset$ or* $\dim\big(f^{-1}(y)\big) = \dim(X) - \dim(Y_i)$ *for all $y \in Y_i$.*

(b) *We have* $\deg\big(\overline{f(X)}\big) \leq \deg(f)$. *The image $f(X)$ contains a dense open subset of $\overline{f(X)}$. If X is irreducible then so is $\overline{f(X)}$.*

(c) *For each $y \in f(X)$ the fibre $f^{-1}(y) \subseteq X$ is closed with* $\deg\big(f^{-1}(y)\big) \leq \deg(f)$. *For each closed set $T \subseteq Y$ the subset $f^{-1}(T)$ is also closed and its degree is at most* $\deg(T)\deg(f)$.

(d) *The degree of the closed complement $\overline{\overline{f(X)} \setminus f(X)}$ is at most D^2.*

(e) *Suppose that X is irreducible. For each $t \in X$ we have*

$$\dim\Big(f^{-1}\big(f(t)\big)\Big) \geq \dim(X) - \dim\big(\overline{f(X)}\big)\ .$$

Those $t \in X$ where equality holds form an open dense subset $X_{\min} \subseteq X$ and $\deg\big(X \setminus X_{\min}\big) \leq D^2\deg(f)$.

(f) *Let $S \subseteq X$ be a closed subset that is the intersection of X and a closed set of degree d. Then the degree of the restricted morphism $f\big|_S$ is at most $d \cdot \deg(f)$, hence* $\deg\big(\overline{f(S)}\big) \leq d \cdot \deg(f)$ *(see* (b)*). If S is an irreducible component of X then there are better bounds:* $\deg\big(f\big|_S\big) \leq \deg(f)$ *and* $\deg\big(\overline{f(S)}\big) \leq \deg(f)$.

Parts (b), (c) and (f) as well as the fact that $X_{\min}$ of (e) is open and dense follows easily using [**71**, Chapters I.1 and II.3] and Fact 2.3.9. Moreover, the closed complement considered in (d) is the union of a number of the locally closed subsets of (a), hence its degree bound follows immediately from (a). Similarly, the subset discussed in (e) is the inverse image of the union of a number of the locally closed subsets of (a), hence its degree is bounded by (a) and (c). So the only thing that remains to be proved is (a).

Proof [Sketch of the proof of (a)] Let $\overline{\mathbb{F}}^m \supseteq X$ and $\overline{\mathbb{F}}^n \supseteq Y$ be the ambient affine spaces, $\Gamma_f \subseteq \overline{\mathbb{F}}^m \times \overline{\mathbb{F}}^n$ the graph of f, and $\pi : \overline{\mathbb{F}}^m \times \overline{\mathbb{F}}^n \to \overline{\mathbb{F}}^n$ the linear projection to the second factor. Then Γ_f is isomorphic to X, hence it is enough to find an analogous partition of $\overline{\pi(\Gamma_f)} = \overline{f(X)}$ with respect to π and Γ_f (with the same bound D defined in terms of $\deg(f)$ and $\dim(X)$).

Let L denote the linear span of Γ_f and set $\tilde{\pi} = \pi\big|_L$. In general, for each variety V of degree at least 2, [**71**, Ex.I.7.7] constructs a cone containing V whose dimension is $\dim(V)+1$, and whose degree is strictly smaller that $\deg(V)$. By iterating this cone-construction we arrive, in at most $\dim(V)-1$ steps, at a variety of degree 1. By [**71**, Ex.I.7.6] this iterated cone is a linear subspace, i.e. the original V is contained in a linear subspace of dimension at most $\dim(V)+\deg(V)-1$. In particular, we have $\dim(L) \leq \dim(\Gamma_f)+\deg(\Gamma_f)-1 = \dim(X)+\deg(f)-1$. We need to find a partition of $\overline{\tilde{\pi}(\Gamma_f)} = \overline{f(X)}$ as in (a) with respect to $\tilde{\pi}$ and Γ_f (with the same bound D). We factor $\tilde{\pi}$ into $\dim(L) - \dim\big(\tilde{\pi}(L)\big) \leq \dim(X)+\deg(f)-1$ consecutive linear projections $\tilde{\pi}_j$, each with one-dimensional fibres. Our strategy is the following. First we partition $\overline{\tilde{\pi}_1(\Gamma_f)}$ via the next Claim 2.3.11. Then for each partition class $C \subseteq \overline{\tilde{\pi}_1(\Gamma_f)}$ we apply again Claim 2.3.11, and partition the closed image $\tilde{\pi}_2(\overline{C})$ We obtain various partitions on partially overlapping subsets of $\tilde{\pi}_2\big(\tilde{\pi}_1(\Gamma_f)\big)$. Let us consider the common refinement of them, it is a partition of $\tilde{\pi}_2\big(\tilde{\pi}_1(\Gamma_f)\big)$ into locally closed sets. We iterate this procedure, and obtain partitions of $\tilde{\pi}_j \circ \cdots \circ \tilde{\pi}_1(\Gamma_f)$ for each j. (Note that k in these applications of Claim 2.3.11 is always at most $\dim(X)+\deg(f)-2$.) In the last step we obtain a partition of $\overline{\tilde{\pi}(\Gamma_f)} = \overline{f(X)}$ as required.

Claim 2.3.11. *Let $Z \subseteq \overline{\mathbb{F}}^k$ be a locally closed set and Γ be the common zero locus inside $\overline{\mathbb{F}} \times Z$ of some polynomials of degree at most d.*

(a) Then Z has a partition into at most $(d+2)^{(d+1)^{k+2}-1}$ locally closed subsets Z_i and there are corresponding $(k+1)$-variate polynomials P_i of degree at most $d^{(d+1)^{k+1}2^d}$ such that

$$\Gamma \cap \big(\overline{\mathbb{F}} \times Z_i\big) = \Big\{(t,\underline{z}) \in \overline{\mathbb{F}} \times Z_i \;\Big|\; P_i(t,\underline{z}) = 0\Big\}$$

for all i, and the closures $\overline{Z_i}$ are defined via equations of degree at most $d^{(d+1)^{k+1}2^d}$ plus the equations of $\overline{Z}$.

(b) Those points $\underline{z} \in Z_i$ for which $\Gamma \cap \big(\overline{\mathbb{F}} \times \{\underline{z}\}\big)$ has any prescribed number of points (it can be $0, 1, \ldots d$ or ∞) form a locally closed subset that is defined (inside Z) via equations of degree at most $d^{(d+1)^{k+1}2^d}$, and the total number of these subsets is at most $(d+2)^{(d+1)^{k+2}}$.

(c) Moreover, one may require both partitions to have the following additional property: the closure in Z of each partition class is the union of partition classes.

Proof [Sketch of proof] The upper bounds and part (c) follow immediately from our construction, we leave them to the reader. Γ can be defined as the common zero locus inside $\overline{\mathbb{F}} \times Z$ of at most $(d+1)^{k+1}$ polynomials of degree at most d (see Fact 2.3.9.(g)). We prove (a) via induction on the number of defining polynomials. If $\Gamma = \overline{\mathbb{F}} \times Z$ then there is nothing to prove. Otherwise let g be one of the nonzero defining polynomials of Γ and $\Gamma' \subseteq \overline{\mathbb{F}} \times Z$ the common zero locus of the other defining polynomials. Applying the induction hypothesis to Γ' gives us a partition $\bigcup_j Z'_j = Z$ and corresponding polynomials P'_j. Our goal is to refine this partition, i.e. find

partitions $Z'_j = \bigcup_i Z'_{ji}$ and find appropriate polynomials P'_{ji}. We shall find the Z'_{ji} one by one with the following algorithm.

The portion of Γ that lies inside $\overline{\mathbb{F}} \times Z'_j$ is defined by the equations $P'_j(t, \underline{z}) = g(t, \underline{z}) = 0$ (besides the equations and inequalities defining Z'_j). We consider g and P'_j as polynomials in the variable t whose coefficients are polynomial functions of the parameter $\underline{z}$. Note that g and P'_j as well as all the polynomials P'_{ji} we construct below have t-degrees at most d. Our plan is to find the gcd of g and P'_j with respect to the variable t for all values of $\underline{z}$ simultaneously. In order to do so we try to run Euclid's algorithm simultaneously for all $\underline{z}$. There are two obstacles we have to overcome. First, for different values of $\underline{z}$ the algorithm needs a different number of steps to complete. Second, to do a polynomial division uniformly for several values of $\underline{z}$ we have to make sure that the degree of the divisor do not vary with $\underline{z}$ (i.e. we can talk about the leading coefficient). So before each polynomial division we construct also a partition of Z'_j, always refining the partition obtained in the previous step, so that the upcoming division can be done uniformly for values $\underline{z}$ lying in the same partition class.

To begin with, let Z'_{j0} and Z'_{j1} denote the loci of those $\underline{z} \in Z'_j$ where all coefficients of g or P'_j respectively vanish. We set $P'_{j0} = P'_j$ and $P'_{j1} = g$. Similarly, for each pair of integers $0 \leq a, b \leq d$ we consider the locus of those $\underline{z} \in Z'_j$ where the t-degrees of g and P'_j are just a and b. This is a partition of Z'_j into locally closed subsets, each defined via the vanishing or non-vanishing of a number of coefficients. For parameter values $\underline{z}$ lying in Z'_{j0} or Z'_{j1} the algorithm stops right away with gcd equal to P'_{j0} or P'_{j1}. On the other hand, for any other partition class $\tilde{Z} \subseteq Z'_j$ we can do the first polynomial division uniformly for all $\underline{z} \in \tilde{Z}$.

During the algorithm we do similar subdivisions again and again. Suppose that we completed a number of polynomial divisions and constructed the partition corresponding to the last completed division. Let $\tilde{Z}$ be a class of that partition and suppose that the algorithm is still running for $\underline{z} \in \tilde{Z}$ and $\tilde{g}$ and $\tilde{r}$ are the divisor and the remainder of the last completed polynomial division for all values $\underline{z} \in \tilde{Z}$. We consider the locus of those $\underline{z} \in \tilde{Z}$ where all coefficients of $\tilde{r}$ vanish (here $\tilde{g}$ does not vanishes). This will be our next Z'_{ji} (whatever i follows now). For $\underline{z} \in Z'_{ji}$ Euclid's algorithm stops at this stage, and we set $P'_{ji} = \tilde{g}$, the gcd we obtain. As before, we partition $\tilde{Z} \setminus Z'_{ji}$ according to the t-degree of $\tilde{r}$ (here the t-degree of $\tilde{g}$ is unimportant). Then we can do the polynomial division $\tilde{g} : \tilde{r}$ uniformly for values $\underline{z}$ lying in the same partition class. This way we obtain our new remainders (one for each partition class), and Euclid's algorithm continues.

It is clear that for each $\underline{z} \in Z'_j$ the gcd is found in at most $\deg(g) + 1 \leq d + 1$ steps, hence we obtain the promised partition $Z'_j = \cup_i Z'_{ji}$. The induction step is complete.

Part (b) follows from part (a). Indeed, the portion of Γ that lies inside $\overline{\mathbb{F}} \times Z_i$ is defined by the equation $P_i(t, \underline{z}) = 0$ (besides the equations of Z_i). For each $\underline{z} \in Z_i$ the number of points in $\Gamma \cap \big(\overline{\mathbb{F}} \times \{\underline{z}\}\big)$ is either ∞ (in case all t-coefficients of P_i are zero at $\underline{z}$), or equal to the t-degree of the polynomial $P_i(t, \underline{z})$ (which is at most d). The locus of those $\underline{z}$ which correspond to a

given degree can be defined via the vanishing or nonvanishing of a number of t-coefficients of $P_i(t, \underline{z})$. This proves the claim.

2.4. Concentration in general

Let $\alpha \subseteq \overline{\mathbb{F}}^m$ be a finite ordered subset. (We will explain later, after the proof of Lemma 2.4.3, why do we need to order this α.) An essential part of our general strategy is to find closed sets X which contain a large number of elements of α compared to their dimension. To measure the relative size of $\alpha \cap X$ we introduce the following:

Definition 2.4.1. For each subset $X \subseteq \overline{\mathbb{F}}^m$ with $\dim(\overline{X}) > 0$ we define the *concentration* of α in X as follows:

$$\mu(\alpha, X) \;=\; \frac{\log|\alpha \cap X|}{\dim(\overline{X})}$$

For simplicity, here and everywhere in Chapter 2, log stands for the natural logarithm. When $\alpha \cap X = \emptyset$, we set $\mu(\alpha, X) = -\infty$.

In this section we first show that the concentration in a closed set X does not decrease too much when we take an appropriate irreducible closed subset.

Proposition 2.4.2. *Let $X \subseteq Y \subseteq \overline{\mathbb{F}}^m$ be closed sets of positive dimension. Then for all finite ordered sets $\alpha \subseteq \beta \subset \overline{\mathbb{F}}^m$ with $\alpha \cap X \neq \emptyset$ we have:*

$$0 \leq \mu(\alpha, X) \leq \mu(\beta, X) \leq \tfrac{\dim(Y)}{\dim(X)} \cdot \mu(\beta, Y) \tag{2.4.1}$$

and for all integers $n > 0$ the n-fold direct products satisfy

$$\mu\Big(\prod{}^n\alpha, \prod{}^n X\Big) = \mu(\alpha, X)\ . \tag{2.4.2}$$

Proof Clear from the definition.

Lemma 2.4.3. *Let $Z \subseteq \overline{\mathbb{F}}^m$ be a closed set with $\dim(Z) > 0$ and $\alpha \subseteq \overline{\mathbb{F}}^m$ a finite ordered subset with $|\alpha \cap Z| > \deg(Z)$. Then there is an irreducible component $Z' \subseteq Z$ such that $\dim(Z') > 0$ and*

$$\mu(\alpha, Z') \geq \mu(\alpha, Z) - \log\big(\deg(Z)\big)\ . \tag{2.4.3}$$

Proof Since Z has at most $\deg(Z)$ irreducible components (see Fact 2.3.9.(c)) there is a component $Z' \subseteq Z$ with

$$\big|\alpha \cap Z'\big| \geq \frac{|\alpha \cap Z|}{\deg(Z)} > 1\ . \tag{2.4.4}$$

In particular we have $\dim(Z') > 0$. We take the logarithm of inequality (2.4.4), divide the two sides by $\dim(Z')$ and rewrite it in terms of concentrations. Using $\dim(Z') \leq \dim(Z)$ we obtain

$$\mu(\alpha, Z') \geq \frac{\dim(Z)}{\dim(Z')}\mu(\alpha, Z) - \frac{\log\big(\deg(Z)\big)}{\dim(Z')} \geq$$
$$\geq \mu(\alpha, Z) - \log\big(\deg(Z)\big)$$

as required.

The proof of Lemma 2.4.3 involves a choice. For proving Theorem 2.1.6 it will be important to use constructions that are uniquely determined. To this end we will use the order on α to make the choices unique. Of course, α-valued sequences and subsets of α can be ordered lexicographically.

In the rest of Chapter 2 we state several existence results. However, in the proofs we typically use explicit constructions. When we write that our construction of a subset (or a tuple of elements, etc.) is uniquely determined, we understand that the result of the construction depends uniquely on the input data (which usually involves an ordered set α).

Lemma 2.4.4. *For all $N > 0$ and $\Delta > 0$ there are reals $B = B_{\mathrm{irr}}(N, \Delta) \geq 0$ and $K = K_{\mathrm{irr}}(N, \Delta) \geq 0$ with the following property.*
Let $Z \subseteq \overline{\mathbb{F}}^m$ be a closed set and $\alpha \subseteq \overline{\mathbb{F}}^m$ an ordered finite subset. Suppose that $0 < \dim(Z) \leq N$, $\deg(Z) \leq \Delta$ and $|\alpha \cap Z| \geq K$. Then there is an irreducible closed subset $Z' \subseteq Z$ such that $\dim(Z') > 0$, $\deg(Z') \leq B$ and

$$\mu(\alpha, Z') \geq \mu(\alpha, Z) - \log(B) \ .$$

Moreover, our construction of Z' is uniquely determined.

Proof Let $B = \Delta^{(N+1)^N}$ and set $K > \Delta^{2N(N+1)^N}$. Then

$$\mu(\alpha, Z) \geq \frac{\log(K)}{N} > \log\left(\Delta^{2(N+1)^N}\right) \ . \tag{2.4.5}$$

We build by induction a sequence $Z = Z_0 \supset Z_1 \supset Z_2 \supset \cdots \supset Z_I$ of closed subsets such that

$$\begin{aligned} 0 < \dim(Z_{i+1}) &< \dim(Z_i) \ , \\ \deg(Z_{i+1}) &\leq \deg(Z_i)^{N+1} \leq \Delta^{(N+1)^{i+1}} \ , \\ \mu(\alpha, Z_i) &\geq \mu(\alpha, Z) - \log\left(\Delta^{i(N+1)^{i-1}}\right) \ . \end{aligned} \tag{2.4.6}$$

for all $0 \leq i < I$. Since the dimensions are strictly decreasing, such a sequence has length $I + 1 \leq N$. Suppose Z_i is already constructed. If it is irreducible, we stop the induction and set $Z' = Z_i$, the lemma holds in this case. Otherwise, it follows from (2.4.5) and (2.4.6) that $|\alpha \cap Z_i| > \Delta^{(N+1)^N} > \deg(Z_i)$ and we may apply Lemma 2.4.3. So there is an irreducible component $Z_i' \subseteq Z_i$ such that $\dim(Z_i') > 0$ and

$$\mu(\alpha, Z_i') \geq \mu(\alpha, Z_i) - \log\big(\deg(Z_i)\big) \geq \mu(\alpha, Z) - \log\left(\Delta^{(i+1)(N+1)^i}\right) \ . \tag{2.4.7}$$

Of course, there are possibly many choices for Z_i', we choose one in such a way that the subset $\alpha_i = \alpha \cap Z_i'$ is lexicographically minimal among the possible intersections. Note that α_i is uniquely determined, but Z_i' may not be. Then $\mu(\alpha_i, Z_i') = \mu(\alpha, Z_i')$ and using (2.4.5) and (2.4.7) we obtain $|\alpha_i| > \deg(Z_i)^{N+1}$. If Z_i' is the only irreducible component containing α_i then it is uniquely determined. We stop the induction and set $Z' = Z_i'$, the lemma holds in this case.

Otherwise let $T_1, T_2, \ldots$ denote those irreducible components of Z_i which contain α_i and let $Z_{i+1} = \bigcap^j T_j$ be their intersection, this is again uniquely determined. Clearly $\dim(Z_{i+1}) < \dim(Z_i)$ and we shall prove that

$$\deg(Z_{i+1}) \leq \deg(Z_i)^{N+1} \ .$$

In fact it is more convenient to prove a slightly stronger statement: for each closed subset $W \subseteq Z_i$ we have

$$\deg(W \cap Z_{i+1}) \leq \deg(W) \cdot \deg(Z_i)^{\dim(W)} . \tag{2.4.8}$$

We prove (2.4.8) by induction on $\dim(W)$, it obviously holds for $\dim(W) = 0$. Assume for a moment that W is irreducible. If it is contained in all T_j then $W \cap Z_{i+1} = W$ and (2.4.8) holds. On the other hand, if say $W \not\subseteq T_1$ then $W' = W \cap T_1$ has smaller dimension, hence satisfies the analogue of (2.4.8). But $\deg(W') \leq \deg(W)\deg(T_1) \leq \deg(W)\deg(Z_i)$, so we have

$$\deg(W \cap Z_{i+1}) = \deg(W' \cap Z_{i+1}) \leq$$

$$\leq \deg(W')\deg(Z_i)^{\dim(W)-1} \leq \deg(W)\deg(Z_i)^{\dim(W)} .$$

as we promised. In order to complete the induction step for a reducible W we simply add up the analogous inequalities for each component of W.

Then $\dim(Z_{i+1}) > 0$ by Remark 2.3.6. Now we have

$$\mu(\alpha, Z_{i+1}) = \mu(\alpha_i, Z_{i+1}) > \mu(\alpha_i, Z_i') = \mu(\alpha, Z_i') ,$$

hence Z_{i+1} satisfies (2.4.6). As we noted earlier, the induction must stop in at most N steps, which proves the lemma.

Next we show that the concentration in a closed set X does not decrease too much when we map X somewhere by a "nice" morphism.

Lemma 2.4.5. *Let $Z \subseteq \overline{\mathbb{F}}^m$ be an irreducible closed set, $\alpha \subset \overline{\mathbb{F}}^m$ an ordered nonempty finite set and $f : Z \to \overline{\mathbb{F}}^l$ a morphism such that*

$$\dim(Z) > \dim\big(\overline{f(Z)}\big) > 0$$

and

$$\dim\big(Z\big) = \dim\big(\overline{f(Z)}\big) + \dim\big(f^{-1}(t)\big)$$

for all $t \in f(\alpha \cap Z)$. Then there is a fibre $S = f^{-1}(s)$, $s \in f(\alpha \cap Z)$ such that for each value (negative, positive or 0) of the parameter ε one has

$$\begin{cases} \textit{either} & \mu\big(f(\alpha \cap Z), \overline{f(Z)}\big) & \geq & \mu(\alpha, Z) - \varepsilon \dim(S) \\ \textit{or} & \mu\big(\alpha, S\big) & \geq & \mu(\alpha, Z) + \varepsilon \dim\big(\overline{f(Z)}\big) \end{cases} \tag{2.4.9}$$

Moreover, our construction of S is uniquely determined.

Note that if all nonempty fibres of f have the same dimension, then the condition $\dim\big(Z\big) = \dim\big(\overline{f(Z)}\big) + \dim\big(f^{-1}(t)\big)$ is satisfied (see Fact 2.3.10.(e)). Note also that S is a closed set with $\deg(S) \leq \deg(f)$ by Fact 2.3.10.(c).

Proof Let us consider those fibres $f^{-1}(t)$ where the number of points $\big|\alpha \cap f^{-1}(t)\big|$ is maximal, and let $S = f^{-1}(s)$ be the one among them for which the subset $\alpha \cap S \subseteq \alpha$ is lexicographically minimal. Then by assumption we have

$$0 < \dim(S) = \dim(Z) - \dim\big(\overline{f(Z)}\big) < \dim(Z) .$$

We have

$$\big|\alpha \cap Z\big| = \sum_{t \in f(\alpha \cap Z)} \big|\alpha \cap f^{-1}(t)\big| ,$$

hence

$$\big|\alpha \cap Z\big| \leq \big|f(\alpha \cap Z)\big| \cdot \big|\alpha \cap S\big|$$

We take the logarithm of our inequality and rewrite it in terms of concentrations:

$$\mu(\alpha, Z) \cdot \dim(Z) \leq \mu\big(f(\alpha \cap Z), \overline{f(Z)}\big) \cdot \dim(\overline{f(Z)}) + \mu\big(\alpha, S\big) \cdot \dim(S)$$

We divide both sides by $\dim(Z)$ and we introduce two extra terms involving ε on the right hand side which cancel each other:

$$\mu\big(\alpha, Z\big) \ \leq \Big[\mu\big(f(\alpha\cap Z), \overline{f(Z)}\big) + \varepsilon \dim(S)\Big] \tfrac{\dim(\overline{f(Z)})}{\dim(Z)} \ + \ \Big[\mu\big(\alpha, S\big) - \varepsilon \dim\big(\overline{f(Z)}\big)\Big] \tfrac{\dim(S)}{\dim(Z)}$$

On the right hand side we see a weighted arithmetic mean of the two expressions in square brackets. We can certainly bound it it from above with the larger of them, which justifies our statement.

The following extension of Lemma 2.4.5 is our basic tool for transporting large concentration from one subset to another. The idea is that if the transport fails than we get an even larger concentration somewhere inside the first subset.

Lemma 2.4.6 (Transport)**.** *For all $\Delta > 0$ there is a real $B = B_{\text{transport}}(\Delta) \geq 0$ with the following property. Let X be an affine algebraic set, $Z \subseteq X$ a closed subset and $f : X \to \overline{\mathbb{F}}^m$ be a morphism with $\deg(Z) \leq \Delta$, $\deg(f) \leq \Delta$ and $\dim\big(\overline{f(Z)}\big) > 0$. Suppose that Z is irreducible. Then for all ordered finite subsets $\alpha \subseteq X$ and all $\varepsilon \geq 0$ either*

$$\mu\big(f(\alpha), \overline{f(Z)}\big) \geq \mu(\alpha, Z) - \log(B) - \varepsilon \cdot \dim(Z) \tag{2.4.10}$$

or there is a closed subset $S \subset Z$ such that $\deg(S) \leq B$, $0 < \dim(S) < \dim(Z)$ and

$$\mu(\alpha, S) \geq \mu(\alpha, Z) - \log(B) + \varepsilon \ . \tag{2.4.11}$$

Moreover, our construction of S is uniquely determined.

Note, that the condition $\dim\big(\overline{f(Z)}\big) > 0$ implies that $\dim(Z) > 0$, hence the concentrations appearing in the lemma are defined.

Proof To simplify notation we replace α with $\alpha \cap Z$, X with Z, Δ with Δ^2 (see Fact 2.3.10.(f)) and f with its restriction to Z, then $\alpha \subseteq Z$. If $\alpha = \emptyset$ then (2.4.10) holds automatically since the right hand side is $-\infty$. So we assume $\alpha \neq \emptyset$. This implies that $f(\alpha) \neq \emptyset$, hence the left hand side of (2.4.10) is non-negative. If $\mu(\alpha, Z) \leq \log(B)$ then inequality (2.4.10) obviously holds since the right hand side is nonpositive. So we assume $\mu(\alpha, Z) > \log(B)$ which implies $\big|\alpha\big| > B$.

First we prove a special case:

$$\begin{aligned} &\text{If} \quad \dim\big(f^{-1}(t)\big) = \dim(Z) - \dim\big(\overline{f(Z)}\big) \text{ for all } t \in f(\alpha) \text{ then} \\ &\text{the lemma is true with any} \quad B \geq 1 + \Delta. \end{aligned} \tag{2.4.12}$$

If $\dim(Z) > \dim\big(\overline{f(Z)}\big)$ then we apply Lemma 2.4.5 with parameter ε. We get a fibre $S = f^{-1}(s)$ satisfying (2.4.9). Since $\varepsilon \geq 0$, we may replace $\varepsilon \dim(\overline{f(Z)})$ with ε and $\varepsilon \dim(S)$ with $\varepsilon \dim(Z)$, hence either (2.4.10) or (2.4.11) holds for any $B \geq 1$. By Fact 2.3.10.(c) $S = f^{-1}(s)$ is closed and $\deg(S) \leq \Delta$, hence (2.4.12) is proved in this case.

On the other hand, if $\dim(Z) = \dim\big(\overline{f(Z)}\big)$ (and we are still in the special case of (2.4.12)), then all points of α are contained in finite fibres of

f, and the number of points in each finite fibre is at most $\deg(f) \leq \Delta$ (see Fact 2.3.10.(c)). Hence

$$\mu\Big(f(\alpha), \overline{f(Z)}\Big) = \frac{\log\big|f(\alpha)\big|}{\dim\big(\overline{f(Z)}\big)} \geq \frac{\log\Big(\big|\alpha\big|/\Delta\Big)}{\dim(Z)} \geq \mu(\alpha, Z) - \log(\Delta)\ ,$$

and therefore (2.4.10) holds for any $B \geq \Delta$. The special case (2.4.12) is proved.

Next we prove the lemma in full generality. We define the following subset:

$$\alpha' = \Big\{ z \in \alpha \ \Big|\ \dim\big(f^{-1}(f(z))\big) = \dim(Z) - \dim\big(\overline{f(Z)}\big) \Big\}\ .$$

First we deal with the case $|\alpha'| \geq \big|\alpha\big|/2$. We have

$$\mu(\alpha', Z) = \frac{\log\big|\alpha'\big|}{\dim(Z)} \geq \frac{\log\big|\alpha\big| - \log(2)}{\dim(Z)} \geq \mu(\alpha, Z) - \log(2)\ .$$

We apply the special case (2.4.12) of the lemma to α' and Z. We obtain that either

$$\begin{gathered}\mu\big(f(\alpha), \overline{f(Z)}\big) \geq \mu\big(f(\alpha'), \overline{f(Z)}\big) \geq \\ \geq \mu(\alpha', Z) - \log(1+\Delta) - \varepsilon \cdot \dim(Z) \geq \\ \geq \mu(\alpha, Z) - \log(2+2\Delta) - \varepsilon \cdot \dim(Z)\ ,\end{gathered}$$

or there is a closed subset $S \subset Z$ such that $\deg(S) \leq 1+\Delta$, $0 < \dim(S) < \dim(Z)$ and

$$\begin{gathered}\mu(\alpha, S) \geq \mu(\alpha', S) \geq \mu(\alpha', Z) - \log(1+\Delta) + \varepsilon \geq \\ \geq \mu(\alpha, Z) - \log(2+2\Delta) + \varepsilon\ .\end{gathered}$$

The lemma holds in this case with any $B \geq 2 + 2\Delta$.

In the remaining case we have $\big|\alpha'\big| < \big|\alpha\big|/2$. Setting

$$S = \Big\{ z \in Z \ \Big|\ \dim\big(f^{-1}(f(z))\big) > \dim(Z) - \dim\big(\overline{f(Z)}\big) \Big\}$$

we have $\big|\alpha \cap S\big| > \frac{1}{2}\big|\alpha\big|$.

The irreducibility of Z implies (see Fact 2.3.10.(e) and Fact 2.3.9.(e)) that S is a closed subset of Z and $\dim(S) < \dim(Z)$, $\deg(S) \leq \Delta'$ with a certain bound $\Delta' = \Delta'\big(\dim(Z), \Delta\big)$. We set

$$B = B_{\text{transport}}(\Delta) = \max\big(2 + 2\Delta, 2\Delta'\big)\ .$$

Then the set S has at least $|\alpha \cap S| > |\alpha|/2 \geq B/2 \geq \Delta'$ points, hence $\dim(S) > 0$ (see Remark 2.3.6). Therefore $\mu(\alpha, S)$ is defined and we can write:

$$\begin{gathered}\mu(\alpha, S) = \frac{\log|\alpha \cap S|}{\dim(S)} \geq \frac{\log|\alpha| - \log(2)}{\dim(S)} \geq \\ \geq \frac{\dim(Z)}{\dim(S)}\mu(\alpha, Z) - \log(2) \geq \mu(\alpha, Z) - \log(B) + \frac{\mu(\alpha, Z)}{\dim(S)}\ .\end{gathered}$$

We compare now the last term to ε. If $\varepsilon \leq \frac{\mu(\alpha,Z)}{\dim(S)}$ then inequality (2.4.11) holds. On the other hand, for $\varepsilon > \frac{\mu(\alpha,Z)}{\dim(S)} \geq \frac{\mu(\alpha,Z)}{\dim(Z)}$ the inequality (2.4.10) holds, since its right hand side becomes negative. We proved the lemma in all cases.

2.5. Closed sets in groups

Definition 2.5.1. A *linear algebraic group*is a closed subgroup $G \leq GL(n, \overline{\mathbb{F}})$. We use this matrix realisation of G to calculate degrees of closed subsets. We shall denote by $\mathrm{mult}(G)$ and $\mathrm{inv}(G)$ the degrees of the morphisms $(g, h) \to gh$ and $g \to g^{-1}$.

As usual, $\mathcal{Z}(G)$, $[G, G]$, and G^0 denote the centre, the commutator subgroup and the unit component of G, and for any subset $A \subseteq G$ we denote by $\langle A \rangle$, $\mathcal{N}_G(A)$ and $\mathcal{C}_G(A)$ the generated subgroup, the normaliser and the centraliser of A. The subgroup $\mathcal{C}_G(A)^0$ is usually called the *connected centraliser* of A. We shall often use products of several elements and subsets in the usual sense. In order to distinguish from this kind of product, the m-fold direct product of a subset $\alpha \subseteq G$ is denoted by $\prod^m \alpha \subseteq \prod^m G$.

Definition 2.5.2. Let $\alpha \subseteq GL(n, \overline{\mathbb{F}})$ be an ordered finite subset. This ordering extends to an ordering of the subgroup $\langle \alpha \rangle$ (hence to α^i for all i) in a natural way. We shall use this extension without further reference.

Remark 2.5.3. We measure the complexity of a closed subset $X \subseteq \overline{\mathbb{F}}^m$ with two *numerical invariants*: $\dim(X)$ and $\deg(X)$. In contrast, we measure the complexity of a closed subgroup $G \leq GL(n, \overline{\mathbb{F}})$ with four *numerical invariants*: $\dim(G)$, $\deg(G)$, $\mathrm{mult}(G)$ and $\mathrm{inv}(G)$. In order to reduce the number of variables to two, say N and Δ, we shall consider groups G with $\dim(G) \leq N$, $\deg(G) \leq \Delta$, $\mathrm{mult}(G) \leq \Delta$ and $\mathrm{inv}(G) \leq \Delta$.

It can be tiresome to bound all four numerical invariants of G. By the following proposition in most cases it is enough to bound only $\dim(G)$ and $\deg(G)$.

Proposition 2.5.4. *Let G be a linear algebraic group and $H \leq G$ a closed subgroup. Then* $\mathrm{mult}(H) \leq \deg(H)^2 \cdot \mathrm{mult}(G)$ *and* $\mathrm{inv}(H) \leq \deg(H) \cdot \mathrm{inv}(G)$. *In particular, if $G = GL(n, \overline{\mathbb{F}})$ then we have* $\mathrm{mult}(H) \leq \deg(H)^2 \cdot 2^{n^2}$ *and* $\mathrm{inv}(H) \leq \deg(H) \cdot (n+1)^{n^2}$.

Proof Follows immediately from Fact 2.3.10.(f) and Fact 2.3.9.(d).

Fact 2.5.5. *Let G be a linear algebraic group. Suppose that $f : \prod^m G \to \prod^n G$ is a morphism for some integers $m, n > 0$ whose n coordinates are all defined to be product expressions (evaluated in the group G) of length at most k of some fixed group elements, the m variables and their inverses. Then* $\deg\left(\overline{f(G)}\right) \leq \deg(f) \leq \mathrm{inv}(G)^l \,\mathrm{mult}(G)^{n(k-1)}$ *where $l \leq nk$ denotes the total number of times inverted variables occur in the n expressions (see Fact 2.3.10.*(b)*). If the product expressions do not contain the inverse of the variables then of course the bound does not depend on* $\mathrm{inv}(G)$.

Definition 2.5.6. Let G be a linear algebraic group. For all $m > 0$ and for each sequence $\underline{g} = (g_1, g_2, \ldots, g_m)$, $g_i \in G$ we define the morphism

$$\tau_{\underline{g}} : \prod^m G \to G \,,$$

$$\tau_{\underline{g}}(a_1, \ldots, a_m) = (g_1^{-1} a_1 g_1)(g_2^{-1} a_2 g_2) \ldots (g_m^{-1} a_m g_m) \,,$$

Remark 2.5.7. Let G be a linear algebraic group and $\underline{g} = (g_1, g_2, \ldots, g_m)$ any sequence. Suppose that $\dim(G) \leq N$, $\deg(G) \leq \Delta$ and $\mathrm{mult}(G) \leq \Delta$ for certain values N and Δ. According to Fact 2.5.5 there is a common upper bound on the degrees:

$$\deg\left(\tau_{\underline{g}}\right) \leq \Delta_\tau\big(m, N, \Delta\big) \ .$$

In fact, it is easy to see that conjugation by g_i is a linear transformation hence $\deg(\tau_{\underline{g}}) \leq \mathrm{mult}(G)^{m-1} \leq \Delta^{m-1}$.

Fact 2.5.8. *Let G be a connected linear algebraic group and $A, B \subseteq G$ arbitrary subsets. Then*

$$AB \subseteq \overline{A}\,\overline{B} \subseteq \overline{AB} \ .$$

We give a short proof, see also [**79**, page 56]. Let us consider the multiplication map $f : G \times G \to G$. If $AB = f(A \times B)$ satisfies a polynomial equation $p = 0$ then $p\big(f(A \times B)\big) = 0$, i.e. the polynomial $p\big(f(\_)\big)$ vanishes on $A \times B$. But then it must vanish on its closure $\overline{A \times B} = \overline{A} \times \overline{B}$, hence p vanishes on $f\big(\overline{A} \times \overline{B}\big) = \overline{A}\,\overline{B}$. □

Closed subgroups of an algebraic group can be very complicated. In contrast, centraliser subgroups are defined by linear equations, and normalisers of a closed subset X can be defined in terms of the equations of X. This proves that

Fact 2.5.9. *Let G be a linear algebraic group.*

(a) The centraliser $\mathcal{C}_G(X)$ of any subset $X \subseteq G$ is closed and its numerical invariants are bounded: $\deg\big(\mathcal{C}_G(X)\big) \leq \deg(G)$, $\mathrm{mult}\big(\mathcal{C}_G(X)\big) \leq \mathrm{mult}(G)$ *and* $\mathrm{inv}\big(\mathcal{C}_G(X)\big) \leq \mathrm{inv}(G)$. *If X is closed then its normaliser $\mathcal{N}_G(X)$ is also closed and its numerical invariants are also bounded:* $\deg\big(\mathcal{N}_G(X)\big) \leq \deg(G)\deg(X)^{\dim(G)}$, $\mathrm{mult}\big(\mathcal{N}_G(X)\big) \leq \mathrm{mult}(G)\deg(X)^{\dim(G)}$ *and* $\mathrm{inv}\big(\mathcal{N}_G(X)\big) \leq \mathrm{inv}(G)\deg(X)^{\dim(G)}$.

(b) Cosets of a closed subgroup $H \leq G$ are also closed, they all have the same degree. Therefore

$$\big|G : G^0\big| = \frac{\deg(G)}{\deg(G^0)} \leq \deg(G) \ .$$

Later we plan to apply the Transport Lemma 2.4.6 to various morphisms of the form $\tau_{\underline{g}}$. In the rest of this section we construct the appropriate sequences $\underline{g}$.

The following proposition gives a morphism which maps a direct power of a given closed subset Y onto a closed subgroup H. It should be considered folklore, see e.g. [**79**, Proposition on page 55] for a similar statement. Nevertheless, for the sake of completeness, we include a proof.

Proposition 2.5.10. *Let $Y \subseteq GL(n, \overline{\mathbb{F}})$ be an irreducible closed subset of positive dimension and $1 \in \alpha \subset GL(n, \overline{\mathbb{F}})$ an ordered finite subset. Let $H \leq GL(n, \overline{\mathbb{F}})$ denote the smallest closed subgroup which is normalised by α and contains Y. Suppose that $\dim(H) \leq m$. Then there is a sequence $\underline{g} = (g_1, g_2, \ldots, g_{2m})$ of elements $g_i \in \alpha^{m-1}$ such that*

$$H = \tau_{\underline{g}}\left(\prod\nolimits^{2m}(Y^{-1}Y)\right) = (g_1^{-1}Y^{-1}Yg_1)(g_2^{-1}Y^{-1}Yg_2)\ldots(g_{2m}^{-1}Y^{-1}Yg_{2m}) \ .$$

Moreover, our construction of $\underline{g}$ is uniquely determined, H is connected and there is a universal bound $\deg(H) \leq \delta\big(m, \deg(\overline{Y^{-1}Y})\big)$.

Remark 2.5.11. In applications the dimension of H may not be known, but if $G \leq GL(n, \overline{\mathbb{F}})$ is any closed subgroup normalised by α which contains Y then one may set $m = \dim(G)$ and one may also use the bound

$$\deg(\overline{Y^{-1}Y}) \leq \mathrm{inv}(G) \cdot \mathrm{mult}(G) \cdot \deg(Y)^2$$

(see Fact 2.3.9.(d) and Fact 2.3.10.(f)).

Proof We set $g_1 = 1$. We will define $g_i \in \alpha^{i-1}$ by induction and consider the product sets

$$Z_i = (g_1^{-1}Y^{-1}Yg_1)(g_2^{-1}Y^{-1}Yg_2)\dots(g_i^{-1}Y^{-1}Yg_i) \subseteq H \ .$$

Suppose that $g_1, g_2, \dots, g_i$ are already defined. We set $g_{i+1} \in \alpha^i$ to be the first element such that

$$\dim\left(\overline{Z_i}\right) < \dim\left(\overline{Z_i \cdot (g_{i+1}^{-1}Y^{-1}Yg_{i+1})}\right) ,$$

if there is any. Since the dimension of $\overline{Z_i}$ is strictly increasing, eventually we must arrive to an index $i \leq m$ so that g_{i+1} does not exist. But then for all $g \in \alpha^i$ the closed subsets

$$\overline{Z_i} \subseteq \overline{Z_i \cdot (g^{-1}Y^{-1}Yg)}$$

are irreducible (see Fact 2.3.10.(b)) of the same dimension, hence they are equal. This implies that $\overline{Z_i}^2 \subseteq \overline{Z_i}$ and $g^{-1}\overline{Z_i}g \subseteq \overline{Z_i}$ for all $g \in \alpha$, hence $\overline{Z_i}$ is a closed connected subgroup normalised by α i.e. $\overline{Z_i} = H$. By Fact 2.3.10.(b) the product Z_i contains a dense open subset of H, hence $H = Z_i^2$ by [**79**, Lemma on page 54]. Setting $g_{i+j} = g_j$ for $1 \leq j \leq i$ and $g_{2i+1} = \dots g_{2m} = 1$ we obtain our statement.

Lemma 2.5.12. *Let $G \leq GL(n, \overline{\mathbb{F}})$ be a closed subgroup, $Z \subseteq G \times G$ an irreducible closed set and $(a, b) \in Z$. Suppose that $\overline{\tau_{(1,1)}(Z)}$ has dimension 0 i.e. it is a finite set. Then there is an irreducible closed subset $A \subseteq G$ such that*

$$Z = \left\{(ah, h^{-1}b) \;\middle|\; h \in A\right\} \tag{2.5.1}$$

and

$$\left\{c \in GL(n, \overline{\mathbb{F}}) \;\middle|\; \dim\left(\overline{\tau_{(c,1)}(Z)}\right) = 0\right\} = \mathcal{C}_{GL(n,\overline{\mathbb{F}})}(A) \ .$$

Note that in the proof we define A explicitly (hence uniquely), but we do not use this fact later.

Remark 2.5.13. Equation (2.5.1) implies immediately that $\dim(A) = \dim(Z)$ and $1 \in A$.

Proof By assumption $\tau_{(1,1)}(Z)$ is finite and its closure is irreducible (see Fact 2.3.10.(b)), hence it is the single point $ab \in G$. Let $\mathrm{pr}_1 : G \times G \to G$ denote the projection on the first factor. We set

$$A = a^{-1}\,\mathrm{pr}_1(Z) \ .$$

We shall prove later, that it is in fact closed. Anyway, $\overline{A}$ is irreducible (see Fact 2.3.10.(b)) and by definition $1 = a^{-1}a \in A$. Then each point of Z

has the form (ah, β) with some $h \in A$ and $\beta \in G$, and for all $h \in A$ there must exist at least one such point. But then $ab = \tau_{(1,1)}(ah, \beta) = ah\beta$ hence $\beta = h^{-1}b$. This proves equation (2.5.1). The set Z is closed, hence A is closed by equation (2.5.1). Now

$$\tau_{(c,1)}(Z) = \Big\{ c^{-1}(ah)c(h^{-1}b) \;\Big|\; h \in A \Big\} = c^{-1}a \Big\{ hch^{-1} \;\Big|\; h \in A \Big\} b$$

for all $c \in GL(n, \overline{\mathbb{F}})$. This has dimension 0 iff the set $\big\{ hch^{-1} \big| h \in A \big\}$ is finite. But A is irreducible, hence its closed image $\overline{\big\{ hch^{-1} \big| h \in A \big\}}$ is also irreducible (see Fact 2.3.10.(b)), so it is finite iff it is a single point (see Fact 2.3.6) i.e. iff hch^{-1} is independent of $h \in A$. But $1 \in A$, hence this last condition is equivalent to $hch^{-1} = c$ for all $h \in A$, which simply means that c commutes with all $h \in A$. This proves the lemma.

The following corollary constructs a morphism $\tau_{\underline{g}}$ which maps a given closed subset Z of some direct power of G onto a subset of G of positive dimension.

Corollary 2.5.14. *Let $G \leq GL(n, \overline{\mathbb{F}})$ be a linear algebraic group and let $1 \in \alpha \subset G$ be an ordered finite subset whose centraliser $\mathcal{C}_G(\alpha)$ is finite. Then for each integer $m \geq 0$ and each irreducible closed subset $Z \subset \prod^m G$ of dimension $\dim(Z) > 0$ there is a sequence $\underline{g} = (g_1, g_2, \ldots, g_m) \in \prod^m \alpha$ such that the closed image $\overline{\tau_{\underline{g}}(Z)}$ has positive dimension. Moreover, our construction of $\underline{g}$ is uniquely determined.*

Proof We shall prove the theorem by induction on m. For $m = 1$ the statement is obvious. So let $m \geq 2$ and we assume that the corollary holds whenever the number of factors is smaller than m. We define several morphisms. For all $g \in G$ let

$$\sigma_g : \prod^m G \to \prod^{m-1} G \,, \qquad \sigma_g(a_1, \ldots, a_m) = \big(g^{-1}a_1 g a_2, a_3, \ldots, a_m \big)$$

and let

$$\pi : \prod^m G \to \prod^{m-2} G \,, \qquad \pi(a_1, \ldots, a_m) = (a_3, a_4, \ldots, a_m) \,,$$
$$\rho : \prod^{m-1} G \to \prod^{m-2} G \,, \qquad \rho(a_2, \ldots, a_m) = (a_3, a_4, \ldots, a_m) \,.$$

For $m = 2$ we use the convention that $\prod^0 G$ is a single point. Note, that these morphisms manipulate only the first two coordinates. In particular

$$\rho\big(\sigma_g(x)\big) = \pi(x) \qquad \text{for all } x \in \prod^m G \,.$$

Our goal is to find an element $g \in \alpha$ such that

$$\dim\big(\overline{\sigma_g(Z)}\big) > 0 \,. \tag{2.5.2}$$

Then we choose the smallest such g (in the order of α) and use the induction hypotheses for $\overline{\sigma_g(Z)} \subseteq \prod^{m-1} G$. This proves the corollary for Z as well.

We distinguish two cases. Suppose first that for all $z \in \prod^{m-2} G$ the subset $Z \cap \pi^{-1}(z)$ is finite (i.e. 0 dimensional). Then $\dim(Z) = \dim\big(\overline{\pi(Z)}\big)$ is positive (see Fact 2.3.10.(e)). But

$$\dim(Z) \geq \dim\big(\overline{\sigma_g(Z)}\big) \geq \dim\Big(\overline{\rho\big(\sigma_g(Z)\big)}\Big) = \dim\big(\overline{\pi(Z)}\big)$$

hence all these dimensions are equal. Hence (2.5.2) is achieved, the corollary holds in this case.

Suppose next that there is a point $z \in \prod^{m-2} G$ such that $Z \cap \pi^{-1}(z)$ has an irreducible component Z' with positive dimension. For simplicity we shall identify the subset $\pi^{-1}(z) = \prod^2 G \times \{z\} \subset \prod^m G$ with $\prod^2 G$ and also $\rho^{-1}(z) = G \times \{z\} \subset \prod^{m-1} G$ with G. With these identifications we have

$$\sigma_g(x) = \tau_{(g,1)}(x) \qquad \text{for all } x \in \prod^2 G \text{ and all } g \in \alpha\ .$$

If $\overline{\sigma_1(Z')} = \overline{\tau_{(1,1)}(Z')}$ has positive dimension then (2.5.2) holds with $g = 1$ since $\dim\big(\overline{\sigma_1(Z)}\big) \geq \dim\big(\overline{\sigma_1(Z')}\big)$. Otherwise we apply Lemma 2.5.12 to our Z' and get an infinite subset $A \leq G$. By assumption α does not centralise A, hence there is an element $g \in \alpha$ which does not commute with A, i.e. $g \notin \mathcal{C}_G(A) \cdot 1$. Now $\overline{\tau_{(g,1)}(Z')} = \overline{\sigma_g(Z')}$ has positive dimension. But then the potentially larger set $\overline{\sigma_g(Z)} \supseteq \overline{\sigma_g(Z')}$ has positive dimension as well. In all cases we proved (2.5.2), hence the corollary holds.

2.6. Spreading large concentration in a group

In this section we establish our main technical tool, the Spreading Theorem. Roughly speaking it says the following. Let α be a finite subset in a connected linear algebraic group G such that $\mathcal{C}_G(\alpha)$ is finite. If G has a closed subset X in which α has much larger concentration than in G then we can find a connected closed subgroup $H \leq G$ normalised by α in which a small power of α has similarly large concentration. (When G is the simple algebraic group used to define a finite group of Lie type L and α generates L then H turns out to be G itself.)

Definition 2.6.1. A finite set $\alpha \subset GL(n, \overline{\mathbb{F}})$ is called *symmetric* if $\alpha = \alpha^{-1}$.

We need the following basic facts.

Proposition 2.6.2. *Let $\alpha \subset GL(n, \overline{\mathbb{F}})$ be a symmetric subset and hH a coset of a closed subgroup $H \leq GL(n, \overline{\mathbb{F}})$. If $hH \cap \alpha \neq \emptyset$ then*

$$\mu(\alpha^2, hH) \geq \mu(\alpha, H)\ , \quad \mu(\alpha^2, H) \geq \mu(\alpha, hH)\ .$$

□

In the rest of Chapter 2 we restrict our attention to connected linear algebraic groups. It is not a serious restriction in the light of the following:

Corollary 2.6.3. *Let $G \leq GL(n, \overline{\mathbb{F}})$ be a closed subgroup and $1 \in \alpha \subset GL(n, \overline{\mathbb{F}})$ a finite symmetric subset. Then*

$$\mu(\alpha, G^0) \leq \mu(\alpha, G) \leq \mu(\alpha^2, G^0) + \log\big(\deg(G)\big)\ .$$

Proof It follows from Fact 2.5.9.(b) and Proposition 2.6.2.

Definition 2.6.4. A *spreading system* $\alpha | G$ consists of a connected closed subgroup $G \leq GL(n, \overline{\mathbb{F}})$, an ordered finite symmetric subset $1 \in \alpha \subset GL(n, \overline{\mathbb{F}})$ normalising G such that $\mu(\alpha, G) \geq 0$ and $\mathcal{C}_G(\alpha)$ is finite.
We say that $\alpha | G$ is (N, Δ, K)-*bounded* for some integer $N > 0$ and reals $\Delta > 0$, $K > 0$ if

$$\dim(G) \leq N, \quad \deg(G) \leq \Delta, \quad \mathrm{mult}(G) \leq \Delta, \quad \mathrm{inv}(G) \leq \Delta, \quad \big|\alpha \cap G\big| \geq K.$$

We say that $\alpha|G$ is (ε, M, δ)-*spreading* for some reals $\varepsilon > 0$, $\delta > 0$ and integer $M > 0$, if there is a connected closed subgroup $H \leq G$ normalised by α such that $\dim(H) > 0$ and

$$\deg(H) \leq \delta \ , \quad \mu\big(\alpha^M, H\big) \geq (1+\varepsilon) \cdot \mu(\alpha, G) \ .$$

Note, that $\mathrm{mult}(H)$ and $\mathrm{inv}(H)$ are also bounded in terms of δ and Δ by Proposition 2.5.4. We call such an H a *subgroup of spreading*, or sometimes *subgroup of* (ε, M, δ)*-spreading*.

Remark 2.6.5. Note that the assumption $\mu(\alpha, G) \geq 0$ is equivalent to $\dim(G) > 0$ and $\alpha \cap G \neq \emptyset$.

Suppose that for some $m \geq 0$ we find a closed subset $Z \subseteq \prod^m G$ in which $\prod^m \alpha$ has large concentration. We use the following lemma to find a closed subset of G in which the concentration of a small power of α is almost as large.

Lemma 2.6.6 (Back to G)**.** *For all parameters $N > 0$ and $\Delta > 0$ there are reals $B = B_b(N, \Delta) > 0$ and $K = K_b(N, \Delta) \geq 0$ with the following property. Let $\alpha|G$ be a spreading system with $\dim(G) \leq N$, $\deg(G) \leq \Delta$ and $\mathrm{mult}(G) \leq \Delta$. Then for all closed subsets $Z \subset \prod^m G$ with $0 < m \leq N$, $\dim(Z) > 0$, $\deg(Z) \leq \Delta$ and $\big|\prod^m \alpha \cap Z\big| \geq K$ there is a closed subset $Y \subseteq G$ such that $\dim(Y) > 0$, $\deg(Y) \leq B$ and*

$$\mu\big(\alpha^{3N}, Y\big) \geq \mu\big(\textstyle\prod^m \alpha, Z\big) - \log(B) \ .$$

Moreover, our construction of Y is uniquely determined.

Proof There is nothing to prove for $m = 1$, so we assume $m \geq 2$. We prove the lemma by induction on $\dim(Z)$. This is possible, since $\dim(Z) \leq N^2$, so the induction has at most N^2 steps. We assume that the lemma holds in dimensions smaller than $\dim(Z)$ with some bounds $B'\big(N, \Delta, \dim(Z)\big)$ and $K'\big(N, \Delta, \dim(Z)\big)$. By Lemma 2.4.4 if K is large enough then there is a (uniquely determined) positive dimensional irreducible closed set $Z' \subseteq Z$ of degree $\deg(Z') \leq B_{\mathrm{irr}}(N^2, \Delta)$ with large concentration:

$$\mu\big(\textstyle\prod^m \alpha, Z'\big) \geq \mu\big(\prod^m \alpha, Z\big) - \log\big(B_{\mathrm{irr}}(N^2, \Delta)\big) \ .$$

This implies immediately that

$$\Big|\textstyle\prod^m \alpha \cap Z'\Big| \geq \dfrac{|\prod^m \alpha \cap Z|^{\dim(Z')/\dim(Z)}}{B_{\mathrm{irr}}(N^2, \Delta)^{\dim(Z')}} \geq \dfrac{K^{1/N^2}}{B_{\mathrm{irr}}(N^2, \Delta)^{N^2}} \ .$$

By the above it is enough to complete the induction step for Z', so from now on we assume that Z is irreducible. Corollary 2.5.14 gives us a (uniquely determined) sequence $\underline{g} = (g_1, g_2, \ldots, g_m) \in \prod^m \alpha$ such that $\overline{\tau_{\underline{g}}(Z)}$ has positive dimension. Recall from Remark 2.5.7 the bound $\Delta_\tau(N, N, \Delta) \geq \deg\big(\tau_{\underline{g}}\big)$. Let

$$\tilde{\Delta} = \max\big(\Delta, \Delta_\tau(N, N, \Delta)\big) \ .$$

We use Lemma 2.4.6 for the two closed sets $Z \subseteq X = \prod^m G$, the morphism $f = \tau_{\underline{g}}$, the finite set $\prod^m \alpha$ (denoted by α in Lemma 2.4.6) and $\varepsilon = 0$.

We note that $\tau_{\underline{g}}\big(\prod^m \alpha\big) \subseteq \alpha^{3N}$. There are two possible outcomes. In case of Lemma 2.4.6.(2.4.10) the closed subset $T = \overline{\tau_{\underline{g}}(Z)} \subseteq G$ satisfies $\dim(T) > 0$,

$$\mu\big(\prod^m \alpha, Z\big) - \log\big(B_{\mathrm{transport}}(\tilde{\Delta})\big) \leq \mu\big(\tau_{\underline{g}}\big(\prod^m \alpha\big), T\big) \leq \mu\big(\alpha^{3N}, T\big)$$

and by Fact 2.3.10.(b) there is an upper bound $\deg(T) \leq D$ depending only on N and Δ. Hence the lemma holds now with $Y = T$ and any $B \geq \max\big(B_{\mathrm{transport}}(\tilde{\Delta}), D\big)$. In case of Lemma 2.4.6.(2.4.11) we have a closed subset $S \subseteq Z \subseteq \prod^m G$ with $0 < \dim(S) < \dim(Z)$, $\deg(S) \leq B_{\mathrm{transport}}(\tilde{\Delta})$ and

$$\mu\big(\prod^m \alpha, S\big) \geq \mu\big(\prod^m \alpha, Z\big) - \log\big(B_{\mathrm{transport}}(\tilde{\Delta})\big)\ .$$

This implies immediately that

$$\Big|\prod^m \alpha \cap S\Big| \geq \frac{|\prod^m \alpha \cap Z|^{\dim(S)/\dim(Z)}}{B_{\mathrm{transport}}(\tilde{\Delta})^{\dim(S)}} \geq \frac{K^{1/N^2}}{B_{\mathrm{transport}}(\tilde{\Delta})^{N^2}}$$

that is, we can make $\prod^m \alpha \cap S$ sufficiently large by choosing K large enough. We set $B'' = B'\big(N, B_{\mathrm{transport}}(\tilde{\Delta}), \dim(Z)\big)$ and apply the induction hypothesis to this S. This gives us a closed set $Y \subseteq G$ such that $\dim(Y) > 0$, $\deg(Y) \leq B''$ and

$$\mu\big(\alpha^{3N}, Y\big) \geq \mu\big(\prod^m \alpha, S\big) - \log(B'') \geq$$

$$\geq \mu\big(\prod^m \alpha, Z\big) - \log\big(B_{\mathrm{transport}}(\tilde{\Delta})B''\big)\ ,$$

the lemma holds again with the bound $B = B_{\mathrm{transport}}(\tilde{\Delta})B''$. The induction step is complete now, the lemma holds in dimension $\dim(Z)$.

We are now ready to prove the Spreading Theorem. Let us first give an outline of the proof which avoids technicalities. Suppose that α has "large" concentration in a subset $X \subseteq G$. We would like to "spread" this large concentration as much as possible, i.e. we are looking for a small power α^M having large concentration in a subgroup H (more precisely, we need a subgroup of spreading H).

We start with $T_0 = X$ and proceed with a simple induction. Proposition 2.5.10 gives us a surjective morphism $\tau_{\underline{g}}$ which maps $Z = \prod^{2\dim(G)}(X^{-1} \times X)$ (the direct product of $2\dim(G)$ copies of the direct product $(X^{-1} \times X)$) onto a subgroup $H \leq G$. The concentration of the product set $\prod^{4\dim(G)}\alpha$ is large in Z, and we try to transport it via $\tau_{\underline{g}}$ into H. Note, that our $\tau_{\underline{g}}$ maps $\prod^{4\dim(G)}\alpha$ into a small power α^m. According to the Transport Lemma 2.4.6 we either succeed and therefore H is a subgroup of spreading, or find a subset $S \subseteq Z$ with significantly larger concentration. This S lives in the direct product $\prod^{4\dim(G)}G$, but Lemma 2.6.6 brings it back to G, i.e. we find a subset $T_1 \subseteq G$ such that a small power α^{m_1} has significantly larger concentration in T_1 than α had in T_0 (see Lemma 2.6.7).

We repeat this process several times. Either at some point we quit the induction with a subgroup of spreading H or we obtain a sequence of subsets $T_0, T_1, \ldots$ with a quickly growing sequence of concentrations $\mu(\alpha^{m_i}, T_i)$. If we let the concentration grow sufficiently large i.e. $\mu(\alpha^m, T_i) \geq \dim(G)\mu(\alpha, X)$ for some i then already in T_i there are enough elements to force large concentration in G. Therefore we either quit the induction with

a subgroup of spreading, or in a bounded number of steps we conclude that $\mu(\alpha^{m_i}, G)$ is large i.e. G itself is a subgroup of spreading.

Lemma 2.6.7 (Try to Spread)**.** *For all parameters $N > 0$ and $\Delta > 0$ there is an integer $M_t = M_t(N)$, and there are reals $B_t = B_t(N, \Delta) > 0$ and $K = K_t(N, \Delta) \geq 0$ with the following property.*
Let $\alpha | G$ be a spreading system with $\dim(G) \leq N$, $\deg(G) \leq \Delta$, $\mathrm{mult}(G) \leq \Delta$ and $\mathrm{inv}(G) \leq \Delta$. Then for all closed subsets $Y \subset G$ with $\dim(Y) > 0$, $\deg(Y) \leq \Delta$ and $|\alpha \cap Y| \geq K$ and all values

$$\kappa \geq \log(B_t)$$

at least one of the following holds:
Either there is a connected closed subgroup $H \leq G$ normalised by α such that $\dim(H) > 0$, $\deg(H) \leq B_t$ and

$$\mu\big(\alpha^{M_t}, H\big) \geq \mu(\alpha, Y) - \kappa \,, \tag{2.6.1}$$

or there is a closed set $T \subseteq G$ such that $\deg(T) \leq B_t$, $\dim(T) > 0$ and

$$\mu\big(\alpha^{M_t}, T\big) \geq \mu(\alpha, Y) + \frac{\kappa}{8N^2} \,. \tag{2.6.2}$$

Moreover, our constructions of H and T are uniquely determined.

Proof Using Lemma 2.4.4 as in the proof of Lemma 2.6.6, we may assume that Y is irreducible. Let us recall from Lemma 2.4.6, Lemma 2.6.6, Remark 2.5.7 and Proposition 2.5.10 the functions $B_{\mathrm{transport}}$, B_{b}, Δ_τ and δ. We define the following parameters:

$$\begin{array}{lcl}
m & = & N \\
\Delta_1 & = & \max\left(\Delta^{6m}, \Delta_\tau(4m, N, \Delta)\right) \\
B_{\mathrm{transport}} & = & B_{\mathrm{transport}}(\Delta_1) \\
\Delta_2 & = & \max(\Delta, B_{\mathrm{transport}}) \\
B_{\mathrm{b}} & = & B_{\mathrm{b}}(4m, \Delta_2) \\
\varepsilon & = & \frac{\kappa}{8mN} + \log(B_{\mathrm{transport}}) + \log(B_{\mathrm{b}}) \\
 & & \\
M_{\mathrm{t}} & = & \max(4m^2, 12N) \\
B_{\mathrm{t}} & = & \max\left(\delta(N, \Delta),\, B_{\mathrm{transport}}^{8m(N+1)} \cdot B_{\mathrm{b}}^{8m(N+1)}\right)
\end{array}$$

We apply Proposition 2.5.10. to the subset Y, this gives us a sequence $\underline{g} = (g_1, g_2, \ldots, g_{2m}) \in \prod^{2m} \alpha^{m-1}$ and a connected closed subgroup $H \leq G$ normalised by α such that $\dim(H) > 0$, $\deg(H) \leq \delta(N, \Delta) \leq B_{\mathrm{t}}$ and

$$\tau_{\underline{g}}\big(\textstyle\prod^{2m} Y^{-1} Y\big) = H \,.$$

We apply Lemma 2.4.6 with parameters Δ_1 and ε to the subsets $X = \prod^{4m} G$ and $Z = \prod^{2m}(Y^{-1} \times Y)$, the morphism $f = \tau_{(g_1, g_1, g_2, g_2, \ldots, g_{2m}, g_{2m})}$, the finite set $\prod^{4m} \alpha^{m-1}$ (denoted by α in Lemma 2.4.6). We need to check that all requirements are satisfied. By assumption $\dim(Y) > 0$ and hence $\dim(H) = \dim\big(f(Z)\big) > 0$. Since Y is irreducible, Z is also irreducible (see Fact 2.3.9.(f)) with $\deg(Z) = \deg(Y)^{4m} \,\mathrm{inv}(G)^{2m} \leq \Delta^{6m}$ (see Fact 2.3.9.(d) and Fact 2.3.10.(f)) and $\deg(f) \leq \Delta_\tau(4m, N, \Delta)$. Therefore the prerequisites of Lemma 2.4.6 are satisfied, hence one of the inequalities 2.4.6.(2.4.10)

or 2.4.6.(2.4.11) is valid with the logarithmic term equal to $\log(B_{\text{transport}})$. Moreover, $\mu\big(\prod^{4m}\alpha, Z\big) = \mu(\alpha, Y)$ and

$$f\big(\prod^{4m}\alpha\big) \subseteq \alpha^{4m^2} \subseteq \alpha^{M_{\text{t}}} \ .$$

In case of 2.4.6.(2.4.10) we have

$$\mu\big(\alpha^{M_{\text{t}}}, H\big) \geq \mu\big(\alpha^{4m^2}, f(Z)\big) \geq \mu\big(f\big(\prod^{4m}\alpha\big), f(Z)\big) \geq$$

$$\geq \mu\big(\prod^{4m}\alpha, Z\big) - \log(B_{\text{transport}}) - \varepsilon\cdot\dim(Z) \geq$$

$$\geq \mu(\alpha, Y) - \log(B_{\text{transport}}) -$$

$$- \Big(\frac{\kappa}{8mN} + \log(B_{\text{transport}}) + \log(B_{\text{b}})\Big)\cdot N\cdot 4m \geq$$

$$\geq \mu(\alpha, Y) - \frac{\kappa}{2} - 4m(N+1)\Big(\log(B_{\text{transport}}) + \log(B_{\text{b}})\Big) \geq$$

$$\geq \mu(\alpha, Y) - \frac{\kappa}{2} - \frac{\log(B_{\text{t}})}{2} \geq \mu(\alpha, Y) - \kappa$$

which is exactly inequality (2.6.1).

In case of 2.4.6.(2.4.11) we have a closed subset $S \subseteq \prod^{4m} G$ with $\dim(S) > 0$, $\deg(S) \leq B_{\text{transport}}$ such that

$$\mu\big(\prod^{4m}\alpha, S\big) \geq \mu\big(\prod^{4m}\alpha, Z\big) - \log(B_{\text{transport}}) + \varepsilon =$$

$$= \mu\big(\alpha, Y\big) - \log(B_{\text{transport}}) + \Big(\frac{\kappa}{8mN} + \log(B_{\text{transport}}) + \log(B_{\text{b}})\Big) =$$

$$\geq \mu\big(\alpha, Y\big) + \ \Big(\frac{\kappa}{8N^2} + \log(B_{\text{b}})\Big) \ .$$

In particular if $K = K_{\text{t}}(N, \Delta)$ is large enough then

$$\big|\prod^{4m}\alpha \cap S\big| \geq \big|\alpha \cap Y\big|^{\dim(S)/\dim(Y)} \geq K_{\text{b}}(4m, \Delta_2) \ .$$

We apply Lemma 2.6.6 with the parameters $4N$ and Δ_2 (which are denoted there by N and Δ) to the set $S \subseteq \prod^{4m} G$ (which is denoted there by Z). Then in the inequalities we have to use $B_{\text{b}} = B_{\text{b}}(4m, \Delta_2)$. Lemma 2.6.6 gives us a subset $T \subseteq G$ (denoted there by Y) with $\dim(T) > 0$, $\deg(T) \leq B_{\text{b}}$ and

$$\mu(\alpha^{12N}, T) \geq \mu\big(\prod^{4m}\alpha, S\big) - \log(B_{\text{b}}) \geq$$

$$\geq \mu\big(\alpha, Y\big) + \Big(\frac{\kappa}{8N^2} + \log(B_{\text{b}})\Big) - \log(B_{\text{b}}) = \mu\big(\alpha, Y\big) + \frac{\kappa}{8N^2}$$

which implies inequality (2.6.2). The lemma is proved in all cases.

Theorem 2.6.8 (Spreading Theorem)**.** *For all parameters $N > 0$, $\Delta > 0$ and $\frac{1}{3} \geq \varepsilon > 0$ there is an integer $M = M_{\text{spreading}}(N, \varepsilon)$ and a real $K = K_{\text{spreading}}(N, \Delta, \varepsilon)$ with the following property.*
Let $\alpha|G$ be an (N, Δ, K)-bounded spreading system and X a closed subset in $\prod^m G$ for some $0 < m \leq N$. If $\deg(X) \leq \Delta$, $\dim(X) > 0$ and

$$\mu\left(\prod^m \alpha, X\right) \geq (1 + 3\varepsilon)\cdot\mu(\alpha, G)$$

then $\alpha|G$ is (ε, M, K)-spreading. Moreover, our construction of the subgroup of spreading is uniquely determined.

Proof Using Lemma 2.6.6 we can easily reduce the theorem to the special case of $m = 1$, so we assume $X \subseteq G$. Let us recall from Lemma 2.6.7 the functions M_{t} and B_{t}. By induction on $i \geq 0$ we shall define the following numbers:

$$\Delta_0 = \Delta \ , \quad \Delta_i = \max\left(\Delta_{i-1}, B_{\mathrm{t}}(N, \Delta_{i-1})\right) \ , \quad M_i = M_{\mathrm{t}}(N)^i \ .$$

Let $I = I(N, \varepsilon)$ be the smallest positive integer such that

$$\left(1 + \frac{\varepsilon}{4N^2}\right)^I \geq N \ . \tag{2.6.3}$$

We set $M = M_I$ and

$$K = \max\left(\Delta_I^{N/\varepsilon}, K_{\mathrm{t}}(N, \Delta_0)^N, K_{\mathrm{t}}(N, \Delta_1)^N, \ldots, K_{\mathrm{t}}(N, \Delta_{i-1})^N\right) \ .$$

Let $\alpha|G$ be an (N, Δ, K)-bounded spreading system and $X \subseteq G$ a closed subset satisfying the conditions of the theorem. Then

$$\mu(\alpha, X) > \mu(\alpha, G) \geq \frac{\log(K)}{N} \ .$$

By induction on i we build a series of closed subsets $T_i \subseteq G$ such that

$$\begin{cases} \dim(T_i) > 0 \ , \quad \deg(T_i) \leq \Delta_i \ , \\ \mu\big(\alpha^{M_i}, T_i\big) \geq \left(1 + \frac{\varepsilon}{4N^2}\right)^i \cdot \mu(\alpha, X) \geq \frac{\log K}{N} \ . \end{cases} \tag{2.6.4}$$

We run the induction until we either prove Theorem 2.6.8 or build the set T_I. We start the induction with $T_0 = X$, this certainly satisfies (2.6.4) with $i = 0$. In the i-th step of the induction we assume that T_{i-1} is already constructed and $i \leq I$.

We apply the Lemma 2.6.7 with parameters N and Δ_{i-1} to the closed subset $Y = T_{i-1}$ and to the finite set $\alpha^{M_{i-1}}$ and

$$\kappa = \varepsilon \cdot \left(1 + \frac{\varepsilon}{4N^2}\right)^{i-1} \cdot \mu(\alpha, X) \ .$$

We need to check that $\kappa \geq \varepsilon \cdot \mu(\alpha, X) \geq \frac{\varepsilon}{N} \cdot \log(K) \geq \log(\Delta_I) \geq \log(\Delta_i) \geq \log\left(B_{\mathrm{t}}(N, \Delta_{i-1})\right)$ and $\left|\alpha^{M_{i-1}} \cap T_{i-1}\right| \geq \exp\left(\mu(\alpha^{M_{i-1}}, T_{i-1})\right) \geq K^{1/N} \geq K_{\mathrm{t}}(N, \Delta_{i-1})$. Note that

$$\left(\alpha^{M_{i-1}}\right)^{M_{\mathrm{t}}(N)} = \alpha^{M_i} \subseteq \alpha^M \ .$$

There are two cases. If inequality 2.6.7.(2.6.2) holds with a subset T then

$$\mu\big(\alpha^{M_i}, T\big) \geq \mu(\alpha^{M_{i-1}}, T_{i-1}) + \frac{\kappa}{4N^2} \geq$$

$$\left(1 + \frac{\varepsilon}{4N^2}\right)^{i-1} \cdot \mu(\alpha, X) + \frac{\varepsilon}{4N^2}\left(1 + \frac{\varepsilon}{4N^2}\right)^{i-1} \cdot \mu(\alpha, X) =$$

$$= \left(1 + \frac{\varepsilon}{4N^2}\right)^i \cdot \mu(\alpha, X)$$

and $\deg(T) \leq B_{\mathrm{t}}(N, \Delta_{i-1}) \leq \Delta_i$ hence $T_i = T$ satisfies the condition (2.6.4). On the other hand, if inequality 2.6.7.(2.6.1) holds with an appropriate subgroup H then we find that $\deg(H) \leq B_{\mathrm{t}}(N, \Delta_{i-1}) \leq \Delta_i \leq K$ and

$$\mu\big(\alpha^M, H\big) \geq \mu\big(\alpha^{M_i}, H\big) \geq \mu\big(\alpha^{M_{i-1}}, T_{i-1}\big) - \kappa \geq$$

$$\geq \left(1 + \frac{\varepsilon}{4N^2}\right)^{i-1} \cdot \mu(\alpha, X) - \varepsilon \cdot \left(1 + \frac{\varepsilon}{4N^2}\right)^{i-1} \cdot \mu(\alpha, X) \geq$$

$$\geq (1-\varepsilon)\cdot\mu(\alpha,X)\geq(1-\varepsilon)(1+3\varepsilon)\mu(\alpha,G)\geq(1+\varepsilon)\mu(\alpha,G)\ .$$

The theorem holds in this case and we stop the induction.

Finally we consider the case when the induction does not stop during the first I steps and we build T_I. Using the first inequality from Proposition 2.4.2 and inequalities (2.6.4) and (2.6.3) we obtain that

$$\mu(\alpha^M,G)\geq\tfrac{\dim(T_I)}{\dim(G)}\cdot\mu(\alpha^M,T_I)\geq$$

$$\geq\tfrac{1}{N}\cdot\left(1+\tfrac{\varepsilon}{4N^2}\right)^I\cdot\mu(\alpha,X)\geq\mu(\alpha,X)\geq(1+3\varepsilon)\mu(\alpha,G)\ .$$

That is, $\alpha|G$ is (ε,M,K)-spreading with $H=G$. The theorem holds in this case too.

2.7. Variations on spreading

The following useful lemma shows that growth in a subgroup of G implies growth in G itself. See [**73**] for similar results.

Lemma 2.7.1. *Let $A\leq G\leq GL(n,\overline{\mathbb{F}})$ be closed subgroups and $1\in\alpha\subset GL(n,\overline{\mathbb{F}})$ a finite subset. Then for all integers $k>0$ one has*

$$\mu\big(\alpha^{k+1},G\big)\geq\mu\big(\alpha,G\big)+\tfrac{\dim(A)}{\dim(G)}\Big[\mu\big(\alpha^k,A\big)-\mu\big(\alpha^{-1}\alpha,A\big)\Big]$$

or equivalently

$$\frac{\left|\alpha^{k+1}\cap G\right|}{\left|\alpha\cap G\right|}\geq\frac{\left|\alpha^k\cap A\right|}{\left|\alpha^{-1}\alpha\cap A\right|}\ .$$

Proof The two inequalities are clearly equivalent, we shall prove the latter form. We shall look at the multiplication map

$$(\alpha\cap G)\times\big(\alpha^k\cap A\big)\xrightarrow{\phi}(\alpha\cap G)\cdot\big(\alpha^k\cap A\big)\subseteq\big(\alpha^{k+1}\cap G\big)$$

On the left hand side we have $|\alpha\cap G|\cdot\left|\alpha^k\cap A\right|$ elements, on the right hand side there are $\left|\alpha^{k+1}\cap G\right|$ elements. Therefore it is enough to prove that

$$\left|\phi^{-1}(g)\right|\leq\left|\alpha^{-1}\alpha\cap A\right|\quad\text{for all } g\in\alpha^{k+1}\cap G$$

and this follows from the calculation below:

$$\phi^{-1}(g)\subseteq\big\{(a,a^{-1}g)\ \big|\ a\in\alpha,\ a^{-1}g\in A\big\}\subseteq\big\{(a,a^{-1}g)\ \big|\ a\in\alpha\cap gA\big\}\ ,$$

hence

$$\left|\phi^{-1}(g)\right|\leq\left|\alpha\cap gA\right|\leq\left|(\alpha\cap gA)^{-1}(\alpha\cap gA)\right|\leq\left|\alpha^{-1}\alpha\cap A\right|\ .$$

The following result is closely related to the "escape from subvarieties" type results in [**72**] and [**73**].

Lemma 2.7.2 (Escape Lemma)**.** *For all parameters $N>0$, $\Delta>0$ and $\frac{1}{7N^2}\geq\varepsilon>0$ there is an integer $M=M_{\text{escape}}(N,\varepsilon)$ and a real $K=K_{\text{escape}}(N,\Delta,\varepsilon)$ with the following property.*
Let $\alpha|G$ be an (N,Δ,K)-bounded spreading system and $X\subsetneq Y$ two closed subsets in $\prod^m G$ for some $1\leq m\leq N$. Suppose that $\dim(Y)>0$, Y is irreducible, $\deg(X)\leq\Delta$ and

$$\mu\left(\textstyle\prod^m\alpha,Y\right)\geq(1-\varepsilon)\cdot\mu(\alpha,G)\ ,$$

$$\mu\big(\textstyle\prod^m\alpha,\ Y\setminus X\big)\leq(1-2\varepsilon)\cdot\mu(\alpha,G)\ .$$

Then $\alpha|G$ is (ε, M, K)-spreading. Moreover, our construction of the subgroup of spreading is uniquely determined.

Proof We set $M = M_{\text{escape}}(N, \varepsilon) = M_{\text{spreading}}(N, \varepsilon)$ and

$$K = K_{\text{escape}}(N, \Delta, \varepsilon) = \max\left(K_{\text{spreading}}(N, \Delta, \varepsilon), 2^{N/\varepsilon}, (2\Delta+1)^{N/(1-\varepsilon)}\right).$$

Then $\mu(\alpha, G) \geq \frac{\log(K)}{N} \geq \frac{\log(2)}{\varepsilon}$ and

$$\log\left(\frac{|\prod^m \alpha \cap Y|}{|\prod^m \alpha \cap (Y \setminus X)|}\right) = \dim(Y)\Big(\mu\big(\prod^m \alpha, Y\big) - \mu\big(\prod^m \alpha, Y \setminus X\big)\Big) \geq$$

$$\geq \dim(Y) \cdot \varepsilon \cdot \mu(\alpha, G)\big) \geq \log(2)\,.$$

Therefore $|\prod^m \alpha \cap X| \geq \frac{1}{2}\,|\prod^m \alpha \cap Y| \geq \frac{1}{2}|\alpha \cap G|^{(1-\varepsilon)\dim(Y)/\dim(G)} > \Delta$, hence $\dim(X) > 0$ and

$$\mu\left(\prod^m \alpha, X\right) \geq \tfrac{\dim(Y)}{\dim(X)}\mu\left(\prod^m \alpha, Y\right) - \log(2) \geq$$

$$\geq \Big(1 + \tfrac{1}{\dim(X)}\Big)\,(1-\varepsilon)\cdot\mu(\alpha, G) - \log(2) \geq (1+7\varepsilon)(1-\varepsilon)\cdot\mu(\alpha, G) - \log(2) \geq$$

$$\geq \big(1 + 5\varepsilon\big) \cdot \mu(\alpha, G) - \varepsilon \cdot \mu(\alpha, G) > (1 + 3\varepsilon) \cdot \mu(\alpha, G)\,.$$

Then $\alpha|G$ is (ε, M, K)-spreading by the Spreading Theorem 2.6.8.

2.8. Centralisers

If G is a simple algebraic group then a maximal torus T can be obtained as the connected centraliser of a (regular semisimple) element. Using this it follows that if an appropriate subset $\alpha \subset G$ does not grow then the concentration of a small power of α in T is at least $\mu(\alpha, G)$. We first generalise this extremely useful result. Then we define CCC-subgroups and establish some of their basic properties.

Recall from Fact 2.5.9 that the degree of any centraliser subgroup is at most $\deg(G)$.

Lemma 2.8.1 (Centraliser Lemma)**.** *For all parameters $N > 0$, $\Delta > 0$ and $1 \geq \varepsilon > 0$ there is an integer $M = M_{\text{c}}(N, \varepsilon)$ and a real $K = K_{\text{c}}(N, \Delta, \varepsilon)$ with the following property.*
Let $\alpha|G$ be an (N, Δ, K)-bounded spreading system and $C = \mathcal{C}_G(b_1, b_2, \ldots, b_m)$ the centraliser of $m \leq N$ elements $b_i \in \alpha \cap G$. If $0 < \dim(C)$ then either

$$\mu\big(\alpha^M, C^0\big) \geq \Big(1 - \varepsilon \cdot 8N\Big) \cdot \mu(\alpha, G)$$

or $\alpha|G$ is (ε, M, K)-spreading. Moreover, in the latter case our construction of the subgroup of spreading is uniquely determined.

Proof We set $M = M_{\text{c}}(N, \varepsilon) = \max\big(4, 3M_{\text{spreading}}(N, \varepsilon)\big)$, $\tilde{\Delta} = \max(\Delta, \Delta^{3m})$ and

$$K = K_{\text{c}}(N, \Delta, \varepsilon) = \max\left(\Delta^{1/\varepsilon}\,,\, \Delta \cdot K_{\text{spreading}}\big(N, \tilde{\Delta}, \varepsilon\big)\right).$$

Note that $\dim(C^0) = \dim(C) > 0$ and $\big|C : C^0\big| \leq \Delta$ by Fact 2.5.9.(b). Combining this with Proposition 2.6.2 we obtain that for some $h \in C$

$$\mu\big(\alpha^M, C^0\big) \geq \mu(\alpha^{M/2}, hC^0) \geq \mu\big(\alpha^{M/2}, C\big) - \log\big(\Delta\big)\,.$$

Since $K > \big(\Delta\big)^{1/\varepsilon}$ we have

$$\mu(\alpha, G) > \tfrac{1}{\dim(G)} \log(K) \geq \tfrac{1}{\varepsilon \cdot \dim(G)} \log\big(\Delta\big) \geq \tfrac{1}{\varepsilon \cdot N} \log\big(\Delta\big) \ .$$

By the above inequalities it is enough to prove that either $\alpha|G$ is (ε, M, K)-spreading or

$$\mu\big(\alpha^{M/2}, C\big) \geq \Big(1 - \varepsilon \cdot 7N\Big) \cdot \mu(\alpha, G) \ . \tag{2.8.1}$$

If $\dim(C) = \dim(G)$ then $G = C$ and there is nothing to prove. So we assume $\dim(C) < \dim(G)$ and apply Lemma 2.4.5 to the subsets $Z = G$ and α and to the function

$$f : G \to \prod^m G \ , \quad f(g) = \big(\ g^{-1}b_1 g,\ g^{-1}b_2 g,\ \ldots\ g^{-1}b_m g\ \big) \in \prod^m G$$

with the parameter $\varepsilon' = -7\varepsilon \frac{\mu(\alpha,G)}{\dim(C)}$. The fibres of f are just the right cosets of the subgroup C, which have equal dimension, hence we obtain a coset $S = Ca$ that satisfies inequality (2.4.9): either

$$\mu(\alpha, G) \leq \mu(\alpha, Ca) + 7\varepsilon \tfrac{\mu(\alpha,G)}{\dim(C)} \big(\dim(G) - \dim(C)\big) \leq$$

$$\leq \mu(\alpha, Ca) + \varepsilon \cdot 7 \dim(G) \cdot \mu(\alpha, G) \leq \mu\big(\alpha^2, C\big) + \varepsilon \cdot 7N \cdot \mu(\alpha, G)$$

(see Proposition 2.6.2) and the inequality (2.8.1) holds in this case, or else

$$\mu(\alpha, G) \leq \mu\big(f(\alpha \cap G), \overline{f(G)}\big) - \tfrac{7\varepsilon \cdot \mu(\alpha,G)}{\dim(C)} \dim(C) =$$

$$= \mu\big(f(\alpha \cap G), \overline{f(G)}\big) - 7\varepsilon \cdot \mu(\alpha, G) \ .$$

We know $f(\alpha \cap G) \subseteq \prod^m \alpha^3$ hence in this latter case we have

$$\mu\big(\textstyle\prod^m \alpha^3, \overline{f(G)}\big) \geq (1 + 7\varepsilon) \cdot \mu(\alpha, G) \ .$$

If $\mu(\alpha^3, G) \geq (1 + \varepsilon)\mu(\alpha, G)$ then we are done. Otherwise

$$(1 + 3\varepsilon)\mu(\alpha^3, G) \leq (1 + 3\varepsilon)(1 + \varepsilon)\mu(\alpha, G) \leq$$

$$\leq (1 + 7\varepsilon)\mu(\alpha, G) \leq \mu\big(\textstyle\prod^m \alpha^3, \overline{f(G)}\big) \ .$$

Now $\deg\big(\overline{f(G)}\big) \leq \tilde{\Delta}$ (see Fact 2.5.5). We apply the Spreading Theorem 2.6.8 with parameters N, $\tilde{\Delta}$ and ε to the spreading system $\alpha^3|G$ and $X = \overline{f(G)}$. We obtain that $\alpha^3|G$ is $(\varepsilon, \frac{1}{3}M, K)$-spreading, hence $\alpha|G$ is (ε, M, K)-spreading.

Definition 2.8.2. Let G be an algebraic group and $X \subseteq G$ an irreducible closed subset. A *CC-generator*[2] for X is a $\dim(G)$-tuple $\underline{g} \in \prod^{\dim(G)} X$ such that

$$\mathcal{C}_G(\underline{g})^0 = \mathcal{C}_G(X)^0 \ .$$

Let $X^{\mathrm{gen}} \subseteq \prod^{\dim(G)} X$ denote the set of all CC-generators and let $X^{\mathrm{nongen}} = \big(\prod^{\dim(G)} X\big) \setminus X^{\mathrm{gen}}$ denote the complement.

Note that X^{gen} depends on the group G, but for simplicity we suppressed it from the notation. When we work with a spreading system $\alpha|G$ then we always define X^{gen} with respect to G.

Proposition 2.8.3. *Let G be an algebraic group and $X \subseteq G$ an irreducible closed subset. Then X has a CC-generator i.e. $X^{\mathrm{gen}} \neq \emptyset$.*

[2]CC refers to *connected centraliser*

Proof We consider sequences $\underline{a} = a_1, a_2, \ldots, a_m$, $a_i \in X$ such that

$$G > \mathcal{C}_G(a_1)^0 > \mathcal{C}_G(a_1, a_2)^0 > \mathcal{C}_G(a_1, a_2, a_3)^0 > \ldots$$

is a strictly decreasing chain of subgroups. The dimension is strictly decreasing in such a chain, hence the length of $\underline{a}$ is $m \leq \dim(G)$. Therefore one of them, say $\underline{a}_{\max}$, is maximal i.e. it cannot be extended. But then

$$\mathcal{C}_G(X)^0 = \mathcal{C}_G(\underline{a}_{\max})^0$$

and we can build a CC-generator from $\underline{a}_{\max}$ by adding to it $\dim(G) - m$ arbitrary elements of X.

Proposition 2.8.4. *Let G be a connected linear algebraic group, X an irreducible closed set and $G \times X \to X$ a morphism which is a group action. For points $x \in X$ let G_x denote the stabiliser subgroup of x. These are closed subgroups and for each integer d the subset $\{x \in X \mid \dim(G_x) > d\}$ is closed in X. In particular, for each d the points $\underline{g} \in \prod^{\dim(G)} G$ with $\dim\left(\mathcal{C}_G(\underline{g})\right) > d$ form a closed subset in $\prod^{\dim(G)} G$.*

Proof For the first half of the proposition (about stabiliser subgroups) we refer to [**80**, Proposition in 1.4]. If we apply this to the conjugation map

$$G \times \prod^{\dim(G)} G \to \prod^{\dim(G)} G\,, \quad (h, \underline{g}) \to h^{-1}\underline{g}h$$

then we obtain the second half (about centraliser subgroups).

Lemma 2.8.5. *Let G be a connected linear algebraic group and $\emptyset \neq X \subseteq G$ an irreducible closed subset. Then X^{gen} is a dense open subset of $\prod^{\dim(G)} X$. Moreover, the degree of its complement X^{nongen} is bounded in terms of $\dim(G)$, $\deg(G)$, $\mathrm{mult}(G)$, $\mathrm{inv}(G)$ and $\deg(X)$.*

Proof First of all $X^{\mathrm{nongen}} = \left\{\underline{g} \,\middle|\, \dim(\mathcal{C}_G(\underline{g})) > \dim(A)\right\}$ is closed by Proposition 2.8.4. Its complement X^{gen} is naturally open, it is nonempty by Proposition 2.8.3, hence it is dense (see Fact 2.3.9.(e)).

Let us consider the conjugation map

$$f : G \times \prod^{\dim(G)} X \to \prod^{\dim(G)} G \times \prod^{\dim(G)} X\,, \quad f(h, \underline{g}) = \left(h^{-1}\underline{g}h, \underline{g}\right).$$

Let Y denote the diagonal subset

$$Y = \left\{(\underline{g}, \underline{g}) \,\middle|\, \underline{g} \in \prod^{\dim(G)} X\right\} \subset \prod^{\dim(G)} G \times \prod^{\dim(G)} X$$

and let $\tilde{f}$ denote the restriction of f to $f^{-1}(Y)$ composed with the second projection $Y \to \prod^{\dim(G)} X$.

The nonempty fibres of f can be easily identified with cosets of appropriate centraliser subgroups. Namely, if $f^{-1}(\underline{g}', \underline{g}) \neq \emptyset$ then $\underline{g}' = h^{-1}\underline{g}h$ for some element $h \in G$ and

$$f^{-1}(\underline{g}', \underline{g}) = \mathcal{C}_G(\underline{g})h \times \{\underline{g}\}\,.$$

All of the involved centralisers contain the subgroup

$$A = \mathcal{C}_G(X)^0$$

and by Proposition 2.8.3 at least one of them has dimension $\dim(A)$. For $\underline{g} \in \prod^{\dim(G)} X$ we have $\underline{g} \in X^{\text{nongen}}$ iff $\dim\left(f^{-1}(\underline{g},\underline{g})\right) > \dim(A)$. By Fact 2.3.10.(e) the subset

$$Z = \left\{ t \;\middle|\; \dim\left(f^{-1}\big(f(t)\big)\right) > \dim(A) \right\} \subseteq G \times \prod^{\dim(G)} X$$

is a closed subset and $\deg(Z)$ is bounded in terms of $\dim(G)$, $\deg(G)$, $\text{mult}(G)$, $\text{inv}(G)$ and $\deg(X)$. By the above $\tilde{f}\big(Z \cap f^{-1}(Y)\big) = X^{\text{nongen}} = \overline{X^{\text{nongen}}}$. By Fact 2.3.10.(f) and Fact 2.3.9.(d) we see that $\deg(X^{\text{nongen}}) = \deg\left(\overline{f(Z)} \cap Y\right) \leq \deg(f) \cdot \deg(Z) \cdot \deg(Y)$ which is bounded in terms of $\dim(G)$, $\deg(G)$, $\text{mult}(G)$, $\text{inv}(G)$ and $\deg(X)$.

Definition 2.8.6. Let G be an algebraic group. A closed subgroup $A < G$ is a *CCC-subgroup*[3] if $A = \mathcal{C}_G(X)^0$ for some irreducible closed subset $X \ni 1$ and A is different from $\{1\}$ and G^0.

Lemma 2.8.7. *Let G be an algebraic group and $A < G$ a CCC-subgroup. Then*

$$\mathcal{C}_G\big(\mathcal{C}_G(A)^0\big)^0 = A\,, \quad \deg(A) \leq \deg(G)$$

and $\deg\left(A^{\text{nongen}}\right)$ is bounded in terms of $\dim(G)$, $\deg(G)$, $\text{mult}(G)$ and $\text{inv}(G)$. If $B < G$ is another CCC-subgroup with $A \neq B$ then $A^{\text{gen}} \cap B^{\text{gen}} = \emptyset$.

Proof Let $1 \in X \subseteq G$ be an irreducible closed subset such that $A = \mathcal{C}_G(X)^0$. Then $X \subseteq \mathcal{C}_G(A)^0$, A is connected and commutes with $\mathcal{C}_G(A)^0$, hence

$$A = \mathcal{C}_G(X)^0 \supseteq \mathcal{C}_G\big(\mathcal{C}_G(A)^0\big)^0 \supseteq A\,.$$

Now $\deg(A) \leq \deg(G)$ by Fact 2.5.9 and then Lemma 2.8.5 implies that $\deg\left(A^{\text{nongen}}\right)$ is bounded in terms of $\dim(G)$, $\deg(G)$, $\text{mult}(G)$ and $\text{inv}(G)$. Finally if $\underline{g} \in A^{\text{gen}}$ then $\mathcal{C}_G\big(\mathcal{C}_G(\underline{g})^0\big)^0 = A \neq B$ hence $\underline{g} \notin B^{\text{gen}}$. This proves that $A^{\text{gen}} \cap B^{\text{gen}} = \emptyset$.

2.9. Dichotomy lemmas

A central idea of the proof of the Product theorem (2.1.4) for $L = SL(n,q)$ (as outlined in the introduction) is the following. If a generating set α of L does not grow then the intersection of α with any maximal torus of L is either relatively large or relatively small. This follows from a similar property of appropriate maximal tori in $SL(n, \overline{\mathbb{F}_q})$. Here we show that CCC-subgroups also satisfy a similar dichotomy. In fact they were designed to do so.

We first prove that if a set α does not grow (or spread), then for any closed set Z either the intersection of α with Z is relatively small or a small power of α has relatively large intersection with the centraliser of Z.

Lemma 2.9.1 (Asymmetric Dichotomy Lemma)**.** *For all parameters $N > 0$, $\Delta > 0$ and $\frac{1}{56N^3} > \varepsilon > 0$ there is an integer $M = M_{\text{a}}(N,\varepsilon)$ and a real $K = K_{\text{a}}(N,\Delta,\varepsilon)$ with the following property.*
Let $\alpha|G$ be an (N,Δ,K)-bounded spreading system. Then either $\alpha|G$ is

[3] CCC refers to "connected centraliser of a connected subgroup"

(ε, M, K)-spreading or for all irreducible closed subsets $Z \subseteq G$ such that $\dim(Z) > 0$, $\deg(Z) < \Delta$ and $\dim\big(\mathcal{C}_G(Z)\big) > 0$ one of the following holds:

$$\mu\,(\alpha, Z) < \big(1 - \tfrac{1}{7N^2}\big) \cdot \mu(\alpha, G)$$

or

$$\mu\big(\alpha^M, \mathcal{C}_G(Z)^0\big) \geq \mu\left(\prod^{\dim(G)} \alpha^M,\ \big(\mathcal{C}_G(Z)^0\big)^{\mathrm{gen}}\right) \geq \Big(1 - \varepsilon \cdot 16N\Big) \cdot \mu(\alpha, G)\ .$$

Moreover, our construction of the subgroup of spreading is uniquely determined.

Proof We define the parameters

$$\varepsilon' = \tfrac{1}{7N^2}\ , \quad \varepsilon''' = \varepsilon \cdot 8N \leq \tfrac{1}{7N^2}$$

and the closed subsets

$$\begin{aligned} Y' = \textstyle\prod^{\dim(G)} Z \quad &\supseteq \quad X' = Z^{\mathrm{nongen}} \\ Y''' = \textstyle\prod^{\dim(G)} \mathcal{C}_G(Z)^0 \quad &\supseteq \quad X''' = \big(\mathcal{C}_G(Z)^0\big)^{\mathrm{nongen}}\ . \end{aligned}$$

We know from Fact 2.5.9 that $\deg\big(\mathcal{C}_G(Z)^0\big) \leq \Delta$. By Lemma 2.8.7 there is an upper bound $\tilde{\Delta} \geq \Delta$ for $\deg(X')$ and $\deg(X''')$ which depends only on N and Δ. We set $M'' = M_{\mathrm{c}}(N, \varepsilon)$,

$$M = M_{\mathrm{a}}(N, \varepsilon) = \max\Big(M_{\mathrm{escape}}(N, \varepsilon'),\ M'',\ M'' \cdot\ M_{\mathrm{escape}}(N, \varepsilon''')\Big)$$

and

$$\begin{aligned} K = K_{\mathrm{a}}(N, \Delta, \varepsilon) = \\ = \max\Big(K_{\mathrm{escape}}(N, \tilde{\Delta}, \varepsilon'),\ K_{\mathrm{c}}(N, \Delta, \varepsilon),\ K_{\mathrm{escape}}(N, \tilde{\Delta}, \varepsilon''')\Big)\ . \end{aligned}$$

We apply the Escape Lemma 2.7.2 with the parameters N, $\tilde{\Delta}$ and ε' to the subsets X' and Y'. If the Escape Lemma 2.7.2 gives us a subgroup of $\big(\varepsilon', M_{\mathrm{escape}}(N, \varepsilon'), K_{\mathrm{escape}}(N, \tilde{\Delta}, \varepsilon')\big)$-spreading then the lemma holds since $\varepsilon \leq \varepsilon'$. Otherwise there are two possibilities. Either

$$\mu(\alpha, Z) = \mu\left(\textstyle\prod^{\dim(G)} \alpha, Y'\right) < (1 - \varepsilon') \cdot \mu(\alpha, G) = \big(1 - \tfrac{1}{7N^2}\big) \cdot \mu(\alpha, G)$$

in which case the lemma holds, or else there is at least one $\dim(G)$-tuple $\underline{g} \in \prod^{\dim(G)} \alpha \cap Z^{\mathrm{gen}}$ (in fact the Escape Lemma gives us many such tuples). We select the lexicographically minimal $\underline{g}$ among them. Note that $\mathcal{C}_G(\underline{g})^0 = \mathcal{C}_G(Z)^0 \neq \{1\}$, in particular $\dim\big(\mathcal{C}_G(\underline{g})\big) > 0$. In this latter case we apply the Centraliser Lemma 2.8.1 with parameters N, Δ and ε to the spreading system $\alpha|G$ and the subgroup $C = \mathcal{C}_G(\underline{g})$. In case we obtain a subgroup of spreading, the lemma holds. Otherwise we have

$$\mu\big(\alpha^{M''}, \mathcal{C}_G(Z)^0\big) \geq \big(1 - \varepsilon \cdot 8N\big) \cdot \mu(\alpha, G) = \big(1 - \varepsilon'''\big) \cdot \mu(\alpha, G)\ .$$

Finally we apply the Escape Lemma 2.7.2 with parameters N, $\tilde{\Delta}$ and ε''' to the spreading system $\alpha^{M''}|G$ and the subsets X''' and Y'''. Again, the lemma holds if we obtain a subgroup of spreading. Otherwise we have

$$\mu\left(\textstyle\prod^{\dim(G)} \alpha^{M''},\ \big(\mathcal{C}_G(Z)^0\big)^{\mathrm{gen}}\right) > (1 - 2\varepsilon''') \cdot \mu(\alpha, G) = \Big(1 - \varepsilon \cdot 16N\Big)\mu(\alpha, G)\ .$$

Then the lemma follows from Proposition 2.4.2 via the following calculation:

$$\mu\big(\alpha^M, \mathcal{C}_G(Z)^0\big) = \mu\big(\textstyle\prod^{\dim(G)} \alpha^M, Y'''\big) \geq$$

$$\geq \mu\big(\prod^{\dim(G)}\alpha^M, (Y''' \setminus X''')\big) = \mu\left(\prod^{\dim(G)}\alpha^M,\ \big(\mathcal{C}_G(Z)^0\big)^{\mathrm{gen}}\right) .$$

The connected centraliser of the connected centraliser of a CCC-subgroup A is A itself, hence applying the previous lemma twice we obtain the following.

Lemma 2.9.2 (Dichotomy Lemma)**.** *For all parameters $N > 0$, $\Delta > 0$ and $\frac{1}{112N^3} > \varepsilon > 0$ there is an integer $M = M_{\mathrm{dichotomy}}(N, \varepsilon)$ and a real $K = K_{\mathrm{dichotomy}}(N, \Delta, \varepsilon)$ with the following property.*
Let $\alpha|G$ be an (N, Δ, K)-bounded spreading system. Then either $\alpha|G$ is (ε, M, K)-spreading or for all CCC-subgroups $A < G$ one of the following holds:

$$\mu\left(\alpha, A\right) < \left(1 - \tfrac{1}{7N^2}\right) \cdot \mu(\alpha, G)$$

or else

$$\mu\big(\alpha^M, A\big) \geq \mu\left(\prod^{\dim(G)}\alpha^M,\ A^{\mathrm{gen}}\right) \geq \Big(1 - \varepsilon \cdot 16N\Big) \cdot \mu(\alpha, G) .$$

Moreover, our construction of the subgroup of spreading is uniquely determined.

Proof We set $M' = M_{\mathrm{a}}(N, \varepsilon)$, $M = M_{\mathrm{dichotomy}}(N, \varepsilon) = (M')^2$ and

$$K = K_{\mathrm{dichotomy}}(N, \Delta, \varepsilon) = K_{\mathrm{a}}(N, \Delta, \varepsilon) .$$

We apply the Asymmetric Dichotomy Lemma 2.9.1 with parameters N, Δ and ε to $\alpha|G$ and the irreducible subset $Z' = A$. Note that $\dim(A) > 0$ and $\dim\big(\mathcal{C}_G(A)\big) > 0$ follows from Definition 2.8.6. If we obtain a subgroup of (ε, M', K)-spreading or if

$$\mu\left(\alpha, A\right) < \left(1 - \tfrac{1}{7N^2}\right) \cdot \mu(\alpha, G)$$

then the lemma holds. Otherwise we have

$$\mu\big(\alpha^{M'}, \mathcal{C}_G(A)^0\big) \geq \Big(1 - \varepsilon \cdot 16N\Big) \cdot \mu(\alpha, G) .$$

We apply again the Asymmetric Dichotomy Lemma 2.9.1 with parameters N, Δ and ε to $\alpha^{M'}|G$ and $Z'' = \mathcal{C}_G(A)^0$. If we obtain a subgroup of (ε, M', K)-spreading then it is a subgroup of (ε, M, K)-spreading for $\alpha|G$ and the lemma holds. Otherwise $\alpha^{M'}|G$ and Z'' must satisfy one of the two inequalities of that lemma. The first one is

$$\mu\left(\alpha^{M'}, \mathcal{C}_G(A)^0\right) < \left(1 - \tfrac{1}{7N^2}\right) \cdot \mu(\alpha, G) \leq \Big(1 - \varepsilon \cdot 16N\Big) \cdot \mu(\alpha, G) ,$$

but this has already been ruled out. Therefore the other inequality holds:

$$\mu\left(\big(\alpha^{M'}\big)^{M'}, \mathcal{C}_G\big(\mathcal{C}_G(A)^0\big)^0\right) \geq$$

$$\geq \mu\left(\prod^{\dim(G)}\alpha^{M'\cdot M'},\ \Big(\mathcal{C}_G\big(\mathcal{C}_G(A)^0\big)^0\Big)^{\mathrm{gen}}\right) \geq \Big(1 - \varepsilon \cdot 16N\Big)\mu(\alpha, G) .$$

But $\mathcal{C}_G\big(\mathcal{C}_G(A)^0\big)^0 = A$ and the Dichotomy Lemma 2.9.2 follows.

2.10. Finding and using CCC-subgroups

Let G be a simple algebraic group and T a maximal torus of G. Combining the previously developed techniques we can show that if an appropriate finite subset $\alpha \subset G$ does not grow then either $\mu(\alpha, T)$ is relatively small or α itself must be very large compared to $\langle\alpha\rangle$ (which must be finite in this case). We actually prove a similar result for non-normal CCC-subgroups of arbitrary connected linear algebraic groups G. For G non-nilpotent we then construct CCC-subgroups which can be used as an input for the above result.

It is crucial in the proofs of our main theorems to find sufficiently many $\langle\alpha\rangle$-conjugates of a CCC-subgroup $A \leq G$. We define a quantity $\hat{\mu}$ measuring their number in a sense analogous to the concentration μ. To simplify the notation we restrict this definition to the case $\alpha \subset G$, in the more general situation we use a much cruder estimate.

Definition 2.10.1. Let G be a connected linear algebraic group, $A \leq G$ a closed subgroup and $\alpha \subset G$ a finite subset. Suppose that G does not normalise A. We define

$$\hat{\mu}\big(\langle\alpha\rangle, G, A\big) = \frac{\log\Big|\big\{t^{-1}At \mid t \in \langle\alpha\rangle\big\}\Big|}{\dim(G) - \dim\big(\mathcal{N}_G(A)\big)} = \frac{\log\Big|\langle\alpha\rangle : \mathcal{N}_{\langle\alpha\rangle}(A)\Big|}{\dim(G) - \dim\big(\mathcal{N}_G(A)\big)} \ .$$

Remark 2.10.2. The G-conjugates of A are parametrised by the quotient variety $X = G/\mathcal{N}_G(A)$. Let $\hat{\alpha} \subset X$ denote the image of $\langle\alpha\rangle$, these are the parameter values that correspond to the $\langle\alpha\rangle$-conjugates of A. Then $\hat{\mu}\big(\langle\alpha\rangle, G, A\big) = \mu(\hat{\alpha}, X)$.

Lemma 2.10.3 (Spreading via CCC-subgroups)**.** *For all parameters $N > 0$, $\Delta > 0$ and $\frac{1}{119N^3} > \varepsilon > 0$ there is an integer $M = M_{\mathrm{s}}(N, \varepsilon)$ and a real $K = K_{\mathrm{s}}(N, \Delta, \varepsilon)$ with the following property.*
Let $\alpha|G$ be an (N, Δ, K)-bounded spreading system and $A < G$ a CCC-subgroup such that

$$\mu(\alpha, A) > \big(1 - \tfrac{1}{7N^2}\big) \cdot \mu(\alpha, G) \ .$$

Suppose that at least one of the following holds:

(a)

$$\Big|\langle\alpha\rangle : \mathcal{N}_{\langle\alpha\rangle}(A)\Big| \geq \big|\alpha\big|^{2N} \ ,$$

(b) $\alpha \subset G$, A is not normal in G and

$$\mu(\alpha, G) \leq \Big(1 - \varepsilon \cdot 64N^3\Big) \cdot \hat{\mu}\big(\langle\alpha\rangle, G, A\big) \ .$$

Then $\alpha|G$ is (ε, M, K)-spreading. Moreover, our construction of the subgroup of spreading is uniquely determined.

Proof By Lemma 2.8.7 the conjugate subsets $h^{-1}A^{\mathrm{gen}}h$ for various h normalising G are pairwise disjoint or coincide. They are all contained in $\prod^{\dim(G)} G$ which has dimension $\dim(G)^2 \leq N^2$.

In case (b) we consider the following set:

$$X = \bigcup \Big\{h^{-1}A^{\mathrm{gen}}h \mid h \in G\Big\} \subseteq \prod^{\dim(G)} G \ .$$

Then $\dim\big(\overline{X}\big) \leq N^2$. The virtue of this estimate is that it depends only on N, but we also need a precise calculation in terms of A and G. We consider the conjugation map $\phi : G \times \overline{A^{\rm gen}} \to \prod^{\dim(G)} G$ defined as $\phi(h, \underline{a}) = h^{-1}\underline{a}h$ (note that $\overline{A^{\rm gen}} = \prod^{\dim(G)} A$). By definition $X = \phi\big(G \times A^{\rm gen}\big)$ hence $\overline{X} = \overline{{\rm im}(\phi)}$ and $\deg(\overline{X})$ is bounded in terms of N and Δ (see Fact 2.3.10.(f)). Consider any pair $(h_0, \underline{a_0}) \in G \times A^{\rm gen}$ and its image $x = h_0^{-1}\underline{a_0}h_0 \in X$. The corresponding fibre is

$$\phi^{-1}(x) = \Big\{ (nh_0, n\underline{a_0}n^{-1}) \;\Big|\; n \in \mathcal{N}_G(A) \Big\} \ ,$$

which is isomorphic (as an algebraic set, see Remark 2.3.8) to $\mathcal{N}_G(A)$. In particular, $G \times A^{\rm gen}$ (which is open and dense in the domain of ϕ) is the union of fibres of dimension $\dim\big(\mathcal{N}_G(A)\big)$. Therefore

$$\dim(\overline{X}) = \dim\big(A^{\rm gen}\big) + \Big[\, \dim(G) - \dim\big(\mathcal{N}_G(A)\big)\Big] > \dim\big(A^{\rm gen}\big)$$

(apply Fact 2.3.10.(e) to the irreducible set $G \times \overline{A^{\rm gen}}$).

In case (a) we define the parameters $\varepsilon'' = \varepsilon \cdot 16N > \varepsilon$ and $\Delta'' = \Delta^N$, in case (b) we use the same ε'' and we set $\Delta'' = \max\big(\Delta, \deg(\overline{X})\big)$. We define

$$M' = M_{\rm dichotomy}(N, \varepsilon) \ , \quad M'' = M_{\rm spreading}(N, \varepsilon'') \ ,$$
$$M = \max\big(4M' + 1 \,,\ 2M' \cdot M''\big) \ ,$$
$$K = \max\Big(K_{\rm dichotomy}(N, \Delta, \varepsilon) \,,\ K_{\rm spreading}(N, \Delta'', \varepsilon'')\Big) \ .$$

We consider all the conjugate subgroups

$$\mathcal{A} = \Big\{ t^{-1}At \;\Big|\; t \in \langle\alpha\rangle \Big\} \ ,$$

they are all CCC-subgroups of G since α normalises G.

In case (a) we have $\log\big|\mathcal{A}\big| \geq 2N \log|\alpha|$ by assumption. In case (b) we obtain instead the following estimate

$$\log\big|\mathcal{A}\big| = \hat{\mu}\big(\langle\alpha\rangle, G, A\big) \cdot \Big[\, \dim(G) - \dim\big(\mathcal{N}_G(A)\big)\Big] =$$
$$= \big[\, \dim(\overline{X}) - \dim(A^{\rm gen})\big] \cdot \hat{\mu}\big(\langle\alpha\rangle, G, A\big) \geq$$
$$\geq \big[\, \dim(\overline{X}) - \dim(A^{\rm gen})\big] \cdot \tfrac{1}{1 - \varepsilon \cdot 64N^3} \cdot \mu(\alpha, G) >$$
$$> \big[\, \dim(\overline{X}) - \dim(A^{\rm gen})\big] \cdot \big(1 + \varepsilon'' \cdot 4\dim(\overline{X})\big) \cdot \mu(\alpha, G) \ .$$

Suppose first that

$$\mu\big(\alpha^2, B\big) \geq \big(1 - \tfrac{1}{7N^2}\big)\, \mu(\alpha, G) \tag{2.10.1}$$

for all $B \in \mathcal{A}$. We apply the Dichotomy Lemma 2.9.2 with parameters N, Δ and ε to $\alpha^2|G$ and each $B \in \mathcal{A}$. We get that either $\alpha^2|G$ is (ε, M', K)-spreading i.e. $\alpha|G$ is $(\varepsilon, 2M', K)$-spreading, and the lemma holds, or

$$\mu\Big(\prod^{\dim(G)} \alpha^{2M'}, B^{\rm gen}\Big) \geq \Big(1 - \varepsilon \cdot 16N\Big)\mu(\alpha^2, G) \geq (1 - \varepsilon'')\mu(\alpha, G)$$

for all $B \in \mathcal{A}$ (in particular, $\prod^{\dim(G)} \alpha^{2M'}$ has at least one element in each $B^{\rm gen}$). Let us consider this latter possibility. By Lemma 2.8.7 the subsets $B^{\rm gen}$ are pairwise disjoint. In case (b) we obtain

$$\mu\left(\prod^{\dim(G)} \alpha^{2M'}, \overline{X}\right) = \tfrac{1}{\dim(\overline{X})} \log\left|\prod^{\dim(G)} \alpha^{2M'} \cap \overline{X}\right| \geq$$

$$\geq \tfrac{1}{\dim(\overline{X})} \log\left(\sum_{B\in\mathcal{A}} \left|\prod^{\dim(G)}\alpha^{2M'} \cap B^{\mathrm{gen}}\right|\right) \geq$$

$$\geq \tfrac{1}{\dim(\overline{X})} \left[\log|\mathcal{A}| + \log\left(\min_{B\in\mathcal{A}} \left|\prod^{\dim(G)}\alpha^{2M'} \cap B^{\mathrm{gen}}\right|\right)\right] =$$

$$\geq \tfrac{1}{\dim(\overline{X})} \left[\log|\mathcal{A}| + \dim(A^{\mathrm{gen}}) \cdot \min_{B\in\mathcal{A}} \left(\mu\left(\prod^{\dim(G)}\alpha^{2M'}, B^{\mathrm{gen}}\right)\right)\right] \geq$$

$$\geq \tfrac{1}{\dim(\overline{X})} \left[\log|\mathcal{A}| + \dim(A^{\mathrm{gen}}) \cdot \big(1-\varepsilon''\big)\,\mu(\alpha, G)\right] \geq$$

$$\geq \tfrac{1}{\dim(\overline{X})} \Big[\big[\dim(\overline{X}) - \dim(A^{\mathrm{gen}})\big] \cdot \big(1+\varepsilon''\cdot 4\dim(\overline{X})\big)\mu(\alpha, G) +$$

$$+ \dim(A^{\mathrm{gen}}) \cdot \big(1-\varepsilon''\big)\,\mu(\alpha, G)\Big] =$$

$$= \Big[1 + 4\varepsilon''\big(\dim(\overline{X}) - \dim(A^{\mathrm{gen}})\big) - \varepsilon''\tfrac{\dim(A^{\mathrm{gen}})}{\dim(\overline{X})}\Big]\,\mu(\alpha, G) >$$

$$> \big(1+3\varepsilon''\big)\cdot\mu(\alpha, G)\ .$$

In case (a) a similar, but much shorter calculation shows that

$$\mu\left(\prod^{\dim(G)}\alpha^{2M'}, \prod^{\dim(G)} G\right) \geq \tfrac{\log|\mathcal{A}|}{\dim(G)^2} \geq \tfrac{2\log|\alpha|}{\dim(G)} \geq (1+3\varepsilon'')\mu(\alpha, G)\ .$$

In both cases we apply the Spreading Theorem 2.6.8 with parameters N, Δ'' and ε'' to $\alpha^{2M'}|G$, and in case (a) to the set $\prod^{\dim(G)}G$, in case (b) to the set $\overline{X}$. We obtain that $\alpha^{2M'}|G$ is (ε'', M'', K)-spreading, hence $\alpha|G$ is $(\varepsilon, 2M'M'', K)$-spreading, the lemma holds.

Finally we assume that condition (2.10.1) does not hold for all members of $\mathcal{A}$. As the subgroup A itself satisfies it, there must be at least one subgroup $B_0 \in \mathcal{A}$ and an element $b \in \alpha$ such that B_0 satisfies (2.10.1) but $b^{-1}B_0b$ doesn't:

$$\mu\big(\alpha^2, b^{-1}B_0b\big) < \big(1 - \tfrac{1}{7N^2}\big)\,\mu(\alpha, G)\ . \tag{2.10.2}$$

Conjugating by b we transform (2.10.1) into

$$\mu\big(\alpha^4, b^{-1}B_0b\big) \geq \mu\big(b^{-1}\alpha^2 b, b^{-1}B_0b\big) =$$

$$= \mu\big(\alpha^2, B_0\big) > \big(1 - \tfrac{1}{7N^2}\big)\ \mu(\alpha, G)\ .$$

Again we apply the Dichotomy Lemma 2.9.2 with parameters N, Δ and ε to $\alpha^4|G$ and the CCC-subgroup $b^{-1}B_0b$. We obtain that either $\alpha^4|G$ is (ε, M', K)-spreading, and the lemma holds in this case, or

$$\mu\big(\alpha^{4M'}, b^{-1}B_0b\big) \geq \big(1 - \varepsilon\cdot 16N\big)\mu(\alpha, G)\ .$$

Now we compare this to inequality (2.10.2) and apply Lemma 2.7.1 to the subgroup $b^{-1}B_0b$ with $k = 4M'$. We obtain that

$$\mu\big(\alpha^{4M'+1}, G\big) \geq \mu(\alpha, G) + \tfrac{\dim(b^{-1}B_0b)}{\dim(G)}\left[\tfrac{1}{7N^2} - \varepsilon\cdot 16N\right]\mu(\alpha, G) \geq$$

$$\geq \mu(\alpha, G) + \tfrac{1}{N}\left[\tfrac{1}{7N^2} - \varepsilon\cdot 16N\right]\mu(\alpha, G) =$$

$$= \left[1 + \tfrac{1}{7N^3} - 16\varepsilon\right]\mu(\alpha, G) \geq (1+\varepsilon)\,\mu(\alpha, G)\ ,$$

hence G itself is a subgroup of $(\varepsilon, 4M'+1, K)$-spreading for $\alpha|G$.

Lemma 2.10.4. *Let G be a non-abelian connected linear algebraic group and $\mathcal{S} \subseteq G$ the closure of the set of those elements $g \in G$ whose centraliser is either the whole of G or does not contain any maximal torus. Then $\dim(\mathcal{S}) < \dim(G)$ and the degree of $\mathcal{S}$ is bounded:*

$$\deg\big(\mathcal{S}\big) \leq \Delta_{\mathrm{bad}}\big(\dim(G), \deg(G)\big) .$$

Proof Let $A \leq G$ be a Cartan subgroup. Then $A = \mathcal{C}_G(T)$ for some maximal torus $T \leq G$. Hence for each $g \in A$ we have $T \leq \mathcal{C}_G(g)$. All Cartan subgroups are conjugates of A, hence their union, denoted by $\mathcal{R}$, is the image of the conjugation map $f : A \times G \to G$, $f(a, g) = g^{-1}ag$. It is well-known that $\mathcal{R}$ contains an open subset U of G and by definition $\overline{G \setminus \mathcal{R}} \subseteq G \setminus U$, so $\dim\big(\overline{G \setminus \mathcal{R}}\big) < \dim(G)$ (see Fact 2.3.9.(e)). Moreover, $\deg\big(\overline{G \setminus \mathcal{R}}\big)$ is bounded in terms of $\dim(G)$ and $\deg(G)$ (see Fact 2.3.10.(d)). We also know that $\deg\big(\mathcal{Z}(G)\big) \leq \deg(G)$ (see Fact 2.5.9). Hence $\mathcal{S} = \big(\overline{G \setminus \mathcal{R}}\big) \cup \mathcal{Z}(G)$ also has bounded degree.

Lemma 2.10.5 (Finding CCC-subgroups)**.** *For all parameters $N > 0$, $\Delta > 0$ and $\frac{1}{56N^3} > \varepsilon > 0$ there is an integer $M = M_{\mathrm{CCC}}(N, \varepsilon)$ and a real $K = K_{\mathrm{CCC}}(N, \Delta, \varepsilon)$ with the following property.*
Let $\alpha|G$ be an (N, Δ, K)-bounded spreading system such that G is non-nilpotent. Then either it is (ε, M, K)-spreading, or there is a CCC-subgroup $A \leq G$ which contains exactly one maximal torus of G and satisfies

$$\mu\big(\alpha^M, A\big) > (1 - \varepsilon \cdot 16N)\mu(\alpha, G) .$$

In particular, A is not normal in G. Moreover, our construction of A and of the subgroup of spreading is uniquely determined.

Proof Recall the functions M_{escape}, K_{escape}, M_{c}, K_{c}, M_{a}, K_{a} and Δ_{bad} from the lemmas 2.7.2, 2.8.1, 2.9.1 and 2.10.4. We define the following constants:

$$M_{\mathrm{c}} = M_{\mathrm{c}}(N, \varepsilon) , \quad M_{\mathrm{escape}} = M_{\mathrm{escape}}(N, \varepsilon) , \quad M_{\mathrm{a}} = M_{\mathrm{a}}(N, \varepsilon) ,$$
$$\tilde{\Delta} = \max\big(\Delta, \Delta_{\mathrm{bad}}(N, \Delta)\big) , \quad M = M_{\mathrm{c}}^N \max\big(M_{\mathrm{escape}}, M_{\mathrm{a}}\big) ,$$
$$K = \max\Big(K_{\mathrm{c}}(N, \Delta, \varepsilon) , K_{\mathrm{escape}}(N, \tilde{\Delta}, \varepsilon) , K_{\mathrm{a}}(N, \Delta, \varepsilon)\Big) .$$

Set $g_0 = 1 \in G$, $G_0 = G$. We define by induction on i the elements $g_i \in \alpha^{(M_{\mathrm{c}})^{i-1}} \cap G$ in such a way that the subgroups

$$G_i = \mathcal{C}_G(g_0, g_1, g_2, \ldots, g_i)^0 = \mathcal{C}_{G_{i-1}}(g_i)^0$$

satisfy

$$\mu\left(\alpha^{(M_{\mathrm{c}})^i}, G_i\right) \geq \Big(1 - \varepsilon \cdot 8N\Big)\mu(\alpha, G) , \tag{2.10.3}$$

all G_i contain some maximal torus of G and they form a strictly decreasing series of subgroups. Then their dimension is strictly decreasing as well, hence the sequence has length smaller than N.

Suppose that such a G_i is already defined for some $N > i \geq 0$. If it is abelian then we stop the induction, otherwise continue. Let $\mathcal{S}_i \subsetneq G_i$ be the subset defined in Lemma 2.10.4. Note, that $\deg(G_i) \leq \Delta$, $\mathrm{mult}(G_i) \leq \Delta$ and $\mathrm{inv}(G_i) \leq \Delta$ (see Fact 2.5.9), hence $\deg(\mathcal{S}_i) \leq \tilde{\Delta}$. We apply the Escape Lemma 2.7.2 with parameters N, $\tilde{\Delta}$ and ε to $\alpha^{(M_{\mathrm{c}})^i}|G$ and the subsets

$X = \mathcal{S}_i$ and $Y = G_i$ of G. If we obtain a subgroup of $(\varepsilon, M_{\text{escape}}, K)$-spreading then the lemma holds. Otherwise, since (2.10.3) holds, we find at least one element

$$g_{i+1} \in \alpha^{(M_c)^i} \cap \big(G_i \setminus \mathcal{S}_i\big) \ .$$

(In fact the Escape Lemma gives us many such elements). We select the g_{i+1} which is minimal in the order of $\langle\alpha\rangle$. According to the definition of $\mathcal{S}_i$,

$$G_{i+1} = \big(G_i \cap \mathcal{C}_G(g_{i+1})\big)^0$$

contains a maximal torus of G_i, which is also a maximal torus in G, and G_{i+1} is strictly smaller than G_i. We apply the Centraliser Lemma 2.8.1 with parameters N, Δ and ε to $\alpha^{(M_c)^i}|G$ and the centraliser subgroup $\mathcal{C}_G(g_0, \ldots, g_{i+1})$. In case we obtain a subgroup of (ε, M_c, K)-spreading, the lemma holds. Otherwise we have

$$\mu\left(\alpha^{(M_c)^i M_c}, G_{i+1}\right) \geq \Big(1 - \varepsilon \cdot 8N\Big) \cdot \mu(\alpha^{(M_c)^i}, G) \geq \Big(1 - \varepsilon \cdot 8N\Big)\mu(\alpha, G)$$

i.e. G_{i+1} satisfies (2.10.3).

As we explained before, this process must stop in at most N steps. But the only way it can stop is to arrive at a connected abelian subgroup G_I which contains a maximal torus T and satisfies inequality (2.10.3).

We set $A = \mathcal{C}_G(G_I)^0$. On the one hand, T commutes with G_I, hence $T \leq A$. On the other hand, $A = \mathcal{C}_G(G_I)^0 \leq \mathcal{C}_G(T)$, and the latter one is a Cartan subgroup, which has a unique maximal torus. Therefore T is the only maximal torus in A. But G is non-nilpotent, hence G has several maximal tori. This implies that A is a CCC-subgroup which is not normal. We apply the Asymmetric Dichotomy Lemma 2.9.1 with parameters N, Δ and ε to $\alpha^{(M_c)^N}|G$ and the subset $Z = G_I$. In case we obtain a subgroup of $\big(\varepsilon, M_a, K\big)$-spreading, the lemma holds. Otherwise, since G_I satisfies (2.10.3), we obtain that

$$\mu(\alpha^M, A) \geq (1 - \varepsilon \cdot 16N)\, \mu(\alpha, G)$$

as required.

Suppose we want to prove that a certain spreading system $\alpha|G$ is (ε, M, K)-spreading. Our strategy is to obtain a CCC-subgroup $A < G$ via Lemma 2.10.5, and use Lemma 2.10.3 to establish the (ε, M, K)-spreading. In order to do this, we need to estimate the number of $\langle\alpha\rangle$-conjugates of A. In Section 2.11 we develop a powerful method for finite $\langle\alpha\rangle$. Later in Section 2.13 we deal with the much simpler case when A has infinitely many conjugates.

2.11. Finite groups of Lie type

In this section we use the general results established earlier to prove Theorem 2.1.6, our main technical result concerning fixpoint groups of Frobenius maps of linear algebraic groups.

Definition 2.11.1. Let G be a linear algebraic group over the field $\overline{\mathbb{F}_p}$.

(a) For each p-power q the usual q-th power map $\overline{\mathbb{F}_p} \to \overline{\mathbb{F}_p}$ is a field automorphism. Applying this to the entries of the $n \times n$ matrices we obtain the group automorphisms

$$\mathit{Frob}_q : GL(n, \overline{\mathbb{F}_p}) \to GL(n, \overline{\mathbb{F}_p}) \ .$$

(Note, that these are not morphisms of varieties.)

(b) More generally, $Frob_q$ can be defined the same way on any algebraically closed field of characteristic p, hence we can talk about $Frob_q$-invariant algebraic sets and $Frob_q$-equivariant morphisms (i.e. morphisms compatible with the $Frob_q$-actions on the domain and the range). (These are precisely the algebraic sets and morphisms *defined over* $\mathbb{F}_q$.)

(c) A *Frobenius map* of G is a group automorphism $\sigma : G \to G$ such that there is a p-power q, an exponent k and a faithful representation $G \hookrightarrow GL(n, \overline{\mathbb{F}_p})$ such that G is $Frob_q$-invariant, and σ^k is the restriction of the automorphism $Frob_q$ to G. The fixpoint subgroup of σ is denoted by G^σ. We define $q_\sigma = \sqrt[k]{q}$.

Remark 2.11.2. The fixpoint set of $Frob_q$ is clearly $GL(n, \overline{\mathbb{F}_p})^{Frob_q} = GL(n, \mathbb{F}_q)$. More generally, if the closed subgroup $G \leq GL(n, \overline{\mathbb{F}_p})$ is $Frob_q$-invariant then $G^{Frob_q} = G(\mathbb{F}_q)$, the set of those elements whose matrix belongs to $GL(n, \mathbb{F}_q)$.

We will combine our previous results with the following powerful extension of the Lang-Weil estimates [**76**], sometimes called the *twisted Lang–Weil estiomate*.

Proposition 2.11.3 (Hrushovski)**.** *Let G be a connected linear algebraic group and $\sigma : G \to G$ a Frobenius map. Then there is a constant $C = C\big(\dim(G), \deg(G)\big)$ such that $|G^\sigma|$ is approximately $q_\sigma^{\dim(G)}$ with error*

$$\Big||G^\sigma| - q_\sigma^{\dim(G)}\Big| \leq C \cdot q_\sigma^{\dim(G)-\frac{1}{2}} \ .$$

In the following corollary, besides various technical estimates, we establish that the finite group G^σ (if it is large enough) reflects the group-theoretic properties of G. E.g. there is a correspondence between subgroups of G and G^σ, and we have $\mathcal{C}_G(G^\sigma) = \mathcal{Z}(G)$.

Corollary 2.11.4. *For all parameters $N > 0$, $\Delta > 0$, $I > 0$ and $1 > \varepsilon > 0$ there is an integer $K = K_{\mathrm{L}}(N, \Delta, I, \varepsilon)$ with the following property.*

(a) Let G be a connected linear algebraic group, $\sigma : G \to G$ a Frobenius map and $\alpha \subseteq G^\sigma$ a finite subset. Suppose that $\dim(G) \leq N$, $\deg(G) \leq \Delta$, $\big|G^\sigma : \langle\alpha\rangle\big| \leq I$ and $|\alpha| \geq K$. Then

$$\dim(G) > 0 \ , \quad \mathcal{C}_G(\alpha) = \mathcal{Z}(G) \ , \quad \log(q_\sigma) \geq 1/\varepsilon \ .$$

(b) Let in addition $A \leq G$ be a σ-invariant closed subgroup of degree $\deg(A) \leq \Delta$. Then $A^\sigma = A \cap G^\sigma$,

$$\Big|\langle\alpha\rangle : \langle\alpha\rangle \cap A\Big| \geq \tfrac{1-\varepsilon}{I\Delta}\Big|G^\sigma\Big|^{1-\dim(A)/\dim(G)} \geq \tfrac{1-\varepsilon}{I\Delta}\Big|\langle\alpha\rangle\Big|^{1-\dim(A)/\dim(G)}$$

and if $A \neq G$ then $\langle\alpha\rangle \cap A \neq \langle\alpha\rangle$.

(c) Suppose furthermore that A is not normal in G. Then α does not normalise A and

$$(1-\varepsilon)\log(q_\sigma) \leq \hat{\mu}\big(\langle\alpha\rangle, G, A\big) \leq (1+\varepsilon)\log(q_\sigma) \ .$$

Proof Recall from Proposition 2.11.3 the constant $C = C\big(N, \Delta\big)$. By Proposition 2.11.3 we have

$$K \leq |\alpha| \leq |G^\sigma| \leq (1+C) q_\sigma^N \ ,$$

hence for large enough K

$$\log(q_\sigma) \geq \log\left(\sqrt[N]{\tfrac{K}{1+C}} \right) > 1/\varepsilon$$

and $\dim(G) > 0$ (see Remark 2.3.6). This proves the two inequalities of (a). In the rest of this proof we often use, that by choosing K large enough one can force q_σ to be arbitrary large.

It is obvious that $A^\sigma = A \cap G^\sigma$. By Proposition 2.11.3 for large enough q_σ (i.e. for large enough K) we have

$$(1 - \tfrac{\varepsilon}{3}) q_\sigma^{\dim(G)} \leq |G^\sigma| \leq (1 + \tfrac{\varepsilon}{3}) q_\sigma^{\dim(G)}$$

and

$$(1 - \tfrac{\varepsilon}{3}) q_\sigma^{\dim(A)} \leq \big|(A^0)^\sigma\big| \leq |A^\sigma| \leq \Delta \big|(A^0)^\sigma\big| \leq \Delta (1 + \tfrac{\varepsilon}{3}) q_\sigma^{\dim(A)} \ .$$

Therefore

$$\Big|G^\sigma : A^\sigma\Big| \geq \frac{(1 - \frac{\varepsilon}{3}) q_\sigma^{\dim(G)}}{(1 + \frac{\varepsilon}{3}) \Delta \, q_\sigma^{\dim(A)}} > \frac{1 - 2\frac{\varepsilon}{3}}{\Delta} q_\sigma^{\dim(G) - \dim(A)} >$$

$$> \frac{1 - 2\frac{\varepsilon}{3}}{(1 + \frac{\varepsilon}{3})\Delta} \Big|G^\sigma\Big|^{1 - \dim(A)/\dim(G)} > \frac{1-\varepsilon}{\Delta} \Big|G^\sigma\Big|^{1 - \dim(A)/\dim(G)} \ .$$

This implies the inequality in (b). If $A \neq G$ then $\dim(A) < \dim(G)$. Since $\big|G^\sigma\big| \geq K$, for large enough K we have $\big|\langle\alpha\rangle : \langle\alpha\rangle \cap A\big| > 1$, so $\langle\alpha\rangle \neq \langle\alpha\rangle \cap A$. This completes the proof of (b).

Let $g \in \mathcal{C}_G(\alpha)$ be such that $g \notin \mathcal{Z}(G)$. Clearly all elements of the $\langle\sigma\rangle$-orbit $g^{\langle\sigma\rangle}$ commute with the elements of α. On the other hand we know from (b) (say with parameter $\varepsilon' = \frac{1}{2}$) that $\langle\alpha\rangle \cap \mathcal{C}_G\big(g^{\langle\sigma\rangle}\big) \neq \langle\alpha\rangle$, which is a contradiction. Therefore $\mathcal{C}_G(\alpha) = \mathcal{Z}(G)$ which completes the proof of (a).

Suppose now that A is not normal in G. We apply (b) (say with parameter $\varepsilon'' = \frac{1}{2}$) to the proper subgroup $\mathcal{N}_G(A) < G$. We obtain that

$$\langle\alpha\rangle \neq \langle\alpha\rangle \cap \mathcal{N}_G(A) = \mathcal{N}_{\langle\alpha\rangle}(A)$$

i.e. α does not normalise A.

By Fact 2.5.9 there is an upper bound $\Delta' = \Delta'(N, \Delta) \geq \deg\big(\mathcal{N}_G(A)\big)$. We set $\varepsilon''' = \frac{\varepsilon}{2N}$. We apply (a) with a sufficiently small parameter ε' to obtain that $\log(q_\sigma) > \frac{1}{\varepsilon'''}\Big(1 + \log\big(\max(\Delta, \Delta', I)\big)\Big)$. Let $B \leq G$ be any σ-invariant closed subgroup with $\dim(B) > 0$ and $\deg(B) \leq \max(\Delta, \Delta')$. We apply Proposition 2.11.3 to B^0 and obtain

$$\Big| \log|G^\sigma| - \dim(G) \cdot \log(q_\sigma) \Big| < 1 \ .$$

This gives us upper and lower bounds on $\log\big|\langle\alpha\rangle \cap B\big|$:

$$(1 - \varepsilon''') \dim(B) \log(q_\sigma) \leq$$

$$\leq \dim(B) \cdot \log(q_\sigma) - 1 - \log(I) \leq \log\big|(B^0)^\sigma\big| - \log(I) \leq$$

$$\leq \log\big|\langle\alpha\rangle \cap B^0\big| \leq \log\big|\langle\alpha\rangle \cap B\big| \leq \log\big|B^\sigma\big| \leq$$

$$\leq \log\left|(B^0)^\sigma\right| + \log\left(\max(\Delta,\Delta')\right) \leq$$
$$\leq \dim(B)\cdot\log(q_\sigma) + 1 + \log\left(\max(\Delta,\Delta')\right) \leq$$
$$\leq (1+\varepsilon''')\dim(B)\log(q_\sigma)$$

We apply these inequalities to $B = G$ and to $B = \mathcal{N}_G(A)$:

$$(1-\varepsilon''')\dim\left(G\right)\log(q_\sigma) \leq \log\left|\langle\alpha\rangle\right| \leq (1+\varepsilon''')\dim\left(G\right)\log(q_\sigma)$$

and

$$(1-\varepsilon''')\dim\left(\mathcal{N}_G(A)\right)\log(q_\sigma) \leq \log\left|\mathcal{N}_{\langle\alpha\rangle}(A)\right| \leq$$
$$\leq (1+\varepsilon''')\dim\left(\mathcal{N}_G(A)\right)\log(q_\sigma)\ .$$

Subtracting the two estimates and dividing the result with $\dim(G) - \dim\left(\mathcal{N}_G(A)\right) > 0$ we obtain

$$(1-\varepsilon)\log(q_\sigma) \leq \frac{\log|\langle\alpha\rangle| - \log|\mathcal{N}_{\langle\alpha\rangle}(A)|}{\dim(G)-\dim(\mathcal{N}_G(A))} \leq (1+\varepsilon)\log(q_\sigma)$$

and this completes the proof of (c).

We arrived at a slightly more general version of Theorem 2.1.6 of the introduction:

Theorem 2.11.5. *For all parameters $N > 0$, $\Delta > 0$, $I > 0$ and $1 > \varepsilon > 0$ there is an integer $M = M_{\mathrm{main}}(N,\varepsilon)$ and a real $K = K_{\mathrm{main}}(N,\Delta,I,\varepsilon)$ with the following property.*
Let G be a connected linear algebraic group over $\overline{\mathbb{F}_p}$. Let $\sigma : G \to G$ a Frobenius map and $1 \in \alpha \subseteq G^\sigma$ an ordered finite symmetric subset. Suppose that $\mathcal{Z}(G)$ is finite, $\dim(G) \leq N$, $\deg(G) \leq \Delta$, $\mathrm{mult}(G) \leq \Delta$, $\mathrm{inv}(G) \leq \Delta$, $\left|G^\sigma : \langle\alpha\rangle\right| \leq I$ and

$$K \leq |\alpha| \leq q_\sigma^{(1-\varepsilon)\dim(G)}\ .$$

Then there is a σ-invariant connected closed normal subgroup $H \lhd G$ such that $\deg H \leq K$, $\dim(H) > 0$ and

$$|\alpha^M \cap H| \geq |\alpha|^{(1+\delta)\dim(H)/\dim(G)}$$

where $\delta = \frac{\varepsilon}{128N^3}$. Moreover, our construction of the subgroup H is uniquely determined.

Proof We set

$$M_{\mathrm{CCC}} = M_{\mathrm{CCC}}\left(N,\Delta,\tfrac{\varepsilon}{119N^3}\right)\ ,\quad M_{\mathrm{s}} = M_{\mathrm{s}}\left(N,\Delta,\tfrac{\varepsilon}{128N^3}\right)\ ,$$
$$M = M_{\mathrm{CCC}}\cdot M_{\mathrm{s}}\ ,$$
$$K = \max\left(\Delta+1, K_{\mathrm{CCC}}\left(N,\Delta,\tfrac{\varepsilon}{119N^3}\right), K_{\mathrm{L}}\left(N,\Delta,I,\tfrac{\varepsilon}{3}\right), K_{\mathrm{s}}\left(N,\Delta,\tfrac{\varepsilon}{128N^3}\right)\right)\ .$$

By Corollary 2.11.4.(a) $\dim(G) > 0$ and $\mathcal{C}_G(\alpha) = \mathcal{Z}(G)$, which is finite, hence $\alpha|G$ is an (N,Δ,K)-bounded spreading system. By assumption

$$\mu(\alpha,G) \leq (1-\varepsilon)\log(q_\sigma)\ .$$

Our construction of H will be uniquely determined, therefore it will be σ-invariant. By Corollary 2.11.4.(c) the rest of the conclusion of the theorem can be rewritten as follows. H is normalised by α, $\deg(H) \leq K$, $\dim(H) > 0$ and

$$\mu(\alpha^M, H) \geq (1+\delta)\mu(\alpha,G)$$

i.e. we need to prove that $\alpha|G$ is (δ,M,K)-spreading and construct a subgroup of spreading that is uniquely determined.

If G were nilpotent then $\mathcal{Z}(G)$ would have positive dimension. By assumption $\mathcal{Z}(G)$ is finite, hence G is not nilpotent. We apply Lemma 2.10.5 with parameters N, Δ and $\frac{\varepsilon}{119N^3}$ to $\alpha|G$. In case we obtain a subgroup of spreading, the theorem holds. Otherwise we find a CCC-subgroup $A \leq G$ which is not normal in G and satisfies

$$\mu\big(\alpha^{M_{\mathrm{CCC}}}, A\big) > \big(1 - \tfrac{\varepsilon}{119N^3} \cdot 16N\big)\, \mu(\alpha, G) > \big(1 - \tfrac{1}{7N^2}\big)\, \big(1 + \tfrac{\varepsilon}{119N^2}\big)\, \mu(\alpha, G)\ .$$

If $\alpha|G$ is $(\frac{\varepsilon}{119N^2}, M_{\mathrm{CCC}}, K)$-spreading, then it is (δ, M, K)-spreading, the theorem holds in this case. So from now on we assume that

$$\mu(\alpha^{M_{\mathrm{CCC}}}, G) < \big(1 + \tfrac{\varepsilon}{119N^2}\big)\mu(\alpha, G)$$

hence

$$\mu\big(\alpha^{M_{\mathrm{CCC}}}, A\big) > \big(1 - \tfrac{1}{7N^2}\big)\, \mu(\alpha^{M_{\mathrm{CCC}}}, G)\ .$$

We know from Lemma 2.8.7 that $\deg(A) \leq \deg(G)$ and Corollary 2.11.4.(c) with parameters N, Δ, I and $\frac{\varepsilon}{3}$ implies that

$$\hat{\mu}\Big(\langle\alpha\rangle, G, A\Big) \geq \big(1 - \tfrac{\varepsilon}{3}\big) \log(q_\sigma) > \big(1 - \tfrac{\varepsilon}{2}\big)\, \big(1 + \tfrac{\varepsilon}{6}\big)\, \frac{\mu(\alpha, G)}{1 - \varepsilon} \geq$$

$$\geq \frac{\big(1 - \frac{\varepsilon}{2}\big)}{1 - \varepsilon}\, \big(1 + \tfrac{\varepsilon}{119N^2}\big)\, \mu(\alpha, G) > \frac{\mu(\alpha^{M_{\mathrm{CCC}}}, G)}{1 - \frac{\varepsilon}{2}}\ .$$

We apply Lemma 2.10.3 with parameters N, Δ and $\frac{\varepsilon}{128N^3} = \delta$ to the spreading system $\alpha^{M_{\mathrm{CCC}}}|G$ and the subgroups $W = G$ and $A \leq G$. If we obtain a subgroup of $(\delta, M_{\mathrm{s}}, K)$-spreading then it is a subgroup of (δ, M, K)-spreading for $\alpha|G$, the theorem holds. Otherwise

$$\mu(\alpha^{M_{\mathrm{CCC}}}, G) > \Big(1 - \delta \cdot 64N^3\Big)\, \hat{\mu}\big(\langle\alpha^{M_{\mathrm{CCC}}}\rangle, G, A\big) = \big(1 - \tfrac{\varepsilon}{2}\big)\, \hat{\mu}\big(\langle\alpha\rangle, G, A\big)\ ,$$

a contradiction.

Remark 2.11.6. In the proof of the Product theorem (2.1.4) one can avoid using Proposition 2.11.3. We know explicitly the number of elements in all finite simple groups of Lie type and also in their maximal tori (see e.g. [**29**]). When G is a connected adjoint simple algebraic group, one can show directly that $(G^\sigma)'$ does not normalise any closed subgroup of positive dimension and small degree. This also implies that $\mathcal{C}_G\big((G^\sigma)'\big)$ is finite which is all we need for the proofs of Theorem 2.1.2 and the Product theorem (2.1.4).

The following result, communicated to us by Martin Liebeck, can be used to complete the above sketch. Let G be a connected adjoint simple algebraic group over an algebraically closed field $\mathbb{F}$ of characteristic p, and σ a Frobenius morphism of G. Let $G(q) = (G^\sigma)'$ and assume $G(q)$ is simple.

Proposition 2.11.7. *There is no proper connected subgroup of G which contains $G(q)$.*

Proof Suppose for a contradiction that $G(q) < H < G$, where H is connected.

First we consider the action of $G(q)$ on the adjoint module $L(G)$. The G-composition factors of $L(G)$ are well-known, and can be found in [**103**, 1.10]. With the exception of $G = B_n, C_n, D_n, F_4$ with $p = 2$ and G_2 with $p = 3$, G is either irreducible on $L(G)$, or has two composition factors, one of which is trivial. In any case, each composition factor is either a restricted

$\mathbb{F}G$-module, or a field twist of one. It follows that $G(q)$ is irreducible on every G-composition factor of $L(G)$. Therefore H is also irreducible on every G-composition factor of $L(G)$, and hence H must be a semisimple group.

For the moment exclude the exceptions $B_n, \dots G_2$ in the above paragraph. Clearly $G(q)$ fixes $L(H) \subset L(G)$, so it follows that $L(H)$ must be a composition factor of co-dimension 1 in $L(G)$. If U_H is a maximal connected unipotent subgroup of H, then a standard result tells us that $\dim H = 2 \dim U_H + \mathrm{rank}(H)$. Since $\dim H = \dim G - 1$, it follows that U_H is also a maximal unipotent subgroup of G, and $\mathrm{rank}(H) = \mathrm{rank}(G) - 1$. So the root system of H has the same number of roots as that of G, and H has rank 1 less than G. An easy check of root systems shows that this is impossible.

It remains to handle the exceptional cases $G = B_n, C_n, D_n, F_4$ $(p = 2)$ and G_2 $(p = 3)$. Consider G_2 and F_4, and let H_0 be a simple factor of H which contains an isomorphic copy of $G(q)$. Then H_0 is of rank at most 2 (resp. 4), and the smallest projective representation of H_0 has dimension at least that of $G(q)$, which is 7 (resp. 26). This is clearly impossible.

Next let $G = D_n$. Here the G-composition factors of $L(G)$ are of high weights $\lambda_2, 0$ (n odd) or $\lambda_2, 0^2$ (n even). We have already dealt with the case where $\dim H = \dim G - 1$, so we may assume n is even and $\dim H = \dim G - 2$. Then either $\dim U_H = \dim U_G$, $\mathrm{rank}(H) = \mathrm{rank}(G) - 2$, or $\dim U_H = \dim U_G - 1$, $\mathrm{rank}(H) = \mathrm{rank}(G)$. An inspection of root systems shows that neither of these is possible.

Now let $G = C_n$, and let V be the natural $2n$-dimensional G-module. As $G(q)$ cannot act nontrivially on a module of dimension less than $2n$, it must act tensor indecomposably on V, and hence so does H. Therefore H is simple. The possibilities for $G(q)$ are $C_n(q)$ and $Sz(q)$ (the latter just for $n = 2$). In the former case $G(q)$ has an elementary abelian subgroup $R = r^n$, where r is a prime dividing $q+1$. Note that r is odd as $p = 2$. Also $\mathrm{rank}(H) \leq \mathrm{rank}(G) = n$. An elementary argument (see [**32**, Section 2]) shows that the abelian r-rank of H is equal to $\mathrm{rank}(H)$, and hence $\mathrm{rank}(H) = n$. The only possibility is that $H = D_n$. But $G(q) = C_n(q)$ does not lie in D_n as it does not fix a quadratic form on V. If $G(q) = Sz(q)$ then H cannot have rank 2 (as C_2 has no connected simple proper subgroup of rank 2), so $H = A_1$; but $Sz(q) \not\leq A_1$, a contradiction.

Finally, if $G = B_n$ then there is a morphism from G to C_n which is an isomorphism of abstract groups, and applying this morphism to $G(q)$ and H, we reduce to the C_n case. This completes the proof.

2.12. Linear groups over finite fields

In this section we first prove our main theorem concerning simple groups of Lie type and various results for p-generated subgroups of $GL(n, \mathbb{F}_p)$ i.e. subgroups generated by elements of order p. These finite groups can be obtained roughly as fixpoint groups of Frobenius maps of linear algebraic groups. The Product theorem (2.1.4) is essentially a special case of Theorem 2.11.5. For perfect p-generated groups Theorem 2.1.7 follows by an

inductive argument based on Theorem 2.11.5. To prove Theorem 2.1.7 in the general case we need a number of finite group-theoretic results.

For the following useful results see [**116**] and [**72**, proof of Lemma 2.2].

Proposition 2.12.1 (Olson)**.** *Let $1 \in \alpha$ be a generating set of a finite group G and β a nonempty subset of G. Then $|\alpha\beta| \geq \min\big(|\beta| + |\alpha|/2, |G|\big)$. In particular, if $\alpha^3 \neq G$ then $|\alpha^3| \geq 2|\alpha|$.*

□

As noted in [**73**] the following proposition is essentially due to Ruzsa (see [**137**] and [**136**]).

Proposition 2.12.2. *Let α be a finite subset of a group. Then*

a)

$$\frac{\left|\left(\alpha \cup \alpha^{-1} \cup \{1\}\right)^3\right|}{|\alpha|} \leq \left(3\frac{|\alpha^3|}{|\alpha|}\right)^3$$

b) If $\alpha = \alpha^{-1}$ is a symmetric set with $1 \in \alpha$ and $m \geq 2$ an integer then

$$\frac{|\alpha^m|}{|\alpha|} \leq \left(\frac{|\alpha^3|}{|\alpha|}\right)^{m-2}$$

□

As mentioned in the introduction, a result of Gowers [**63**] implies the following.

Proposition 2.12.3 (Nikolov, Pyber [**113**])**.** *Let G be a finite group and let k denote the minimal degree of a complex representation. Suppose that α, β and γ are subsets of G such that*

$$|\alpha||\beta||\gamma| > \frac{|G|^3}{k} .$$

Then $\alpha\beta\gamma = G$. In particular, if $|\alpha| > |G|/\sqrt[3]{k}$ then $\alpha^3 = G$.

Proposition 2.12.4. *Let G be a simple algebraic group and $\sigma : G \to G$ a Frobenius map. If L is the simple group of Lie type obtained as a composition factor of G^σ then the minimal degree of a complex representation of L is at least $\frac{q_\sigma - 1}{2}$. If $q_\sigma \geq 20$ and $\alpha \subseteq L$ is a subset of size at least $q_\sigma^{\dim(G) - \frac{1}{4}}$ then $\alpha^3 = L$.*

Proof The first statement is an obvious consequence of the Landazuri-Seitz lower bounds ([**92**] cf. [**89**, Table 5.3A]). If $q_\sigma \geq 4$ then $|L| \leq q_\sigma^{\dim(G)}$ (see [**30**]). Now the second statement follows from Proposition 2.12.3.

We are now ready to prove our main result, the Product theorem (2.1.4).

Theorem 2.12.5. *For all parameters $r > 0$ there is a real $\varepsilon = \varepsilon(r) > 0$ with the following property.*
Let L be a finite simple group of Lie type of Lie rank at most r and $\alpha \subset L$ a generating set. Then either $\alpha^3 = L$ or

$$|\alpha^3| \geq |\alpha|^{1+\varepsilon} .$$

Proof There is a simple adjoint algebraic group G and a Frobenius map $\sigma : G \to G$ such that $L \leq G^\sigma$, and there are universal bounds $I(r)$, $N(r)$ and $\Delta(r)$ such that

$$\left|G^\sigma : L\right| \leq I(r) \ , \quad \dim(G) \leq N(r) \ ,$$
$$\deg(G) \leq \Delta(r) \ , \quad \mathrm{mult}(G) \leq \Delta(r) \ , \quad \mathrm{inv}(G) \leq \Delta(r) \ .$$

If $|\alpha| \geq q_\sigma^{\dim(G)-\frac{1}{4}}$ and $q_\sigma \geq 20$ then $\alpha^3 = L$ by Proposition 2.12.4. Assume otherwise.

Suppose first that $\alpha = \alpha^{-1}$ is symmetric with $1 \in \alpha$. We apply Theorem 2.11.5 with parameters $N(r)$, $\Delta(r)$, $I(r)$ and $\varepsilon' = \frac{1}{4\dim(G)}$ and obtain an integer $M = M(r)$ and a real $K = K(r)$. We may assume that $M \geq 3$, and by Corollary 2.11.4.(a) we may increase K so that $|\alpha| \geq K$ implies $q_\sigma \geq 20$. Since G is simple, we have $G = H$ now. If $K \leq |\alpha| \leq q_\sigma^{\dim(G)-\frac{1}{4}}$ then by Theorem 2.11.5 we have

$$|\alpha^M| \geq |\alpha|^{1+\frac{1}{512N^4}} \ .$$

Finally we assume $|\alpha| \leq K$ and $\alpha^3 \neq L$. By Proposition 2.12.1 we have

$$|\alpha^3| \geq 2|\alpha| \geq |\alpha|^{1+\varepsilon''}$$

where $\varepsilon'' = \min\left(\frac{\log(2)}{\log(K)}, \frac{1}{512N^4}\right)$ (which depends only on r). We obtain that in any case

$$|\alpha^M| \geq |\alpha|^{1+\varepsilon''} \ .$$

The theorem follows in the symmetric case from Proposition 2.12.2.(b).

The general case then follows using Proposition 2.12.2.(a).

In Theorem 2.11.5 it is essential to assume that the centre of the algebraic group G is finite. Without this assumption the statement fails. However, we can complement it for finite groups with possibly large centre using the following special case of a deep result of Nikolov and Segal ([**111**, Theorem 1.7]).

Proposition 2.12.6. *Let P be a finite perfect group generated by d elements. Then every element of G is the product of $g(d)$ commutators where $g(d) = 12d^3 + \mathcal{O}(d^2)$ depends only on d.*

Next we will describe more precisely the Nori correspondence between p-generated subgroups of $GL(n, \mathbb{F}_p)$ and certain closed subgroups of $GL(n, \overline{\mathbb{F}_p})$ and some other useful facts about perfect p-generated subgroups.

Proposition 2.12.7. *Let $P \leq GL(n, \mathbb{F}_p)$ be a p-generated subgroup. Then there are bounds $I = I_{\exp}(n)$, $\Delta = \Delta_{\exp}(n)$ and $K = K_{\exp}(n)$ with the following properties.*

- *(a) There is a $Frob_p$-invariant connected closed subgroup $G \leq GL(n, \overline{\mathbb{F}_p})$ such that $\dim(G) \leq n^2$, $\deg(G) \leq \Delta$, $\mathrm{mult}(G) \leq \Delta$, $\mathrm{inv}(G) \leq \Delta$ and P is a subgroup of $G(\mathbb{F}_p)$ of index at most I.*
- *(b) If P is perfect then the degree of any complex representation is at least $(p-1)/2$.*
- *(c) If moreover $|P| \geq K$ and $\alpha \subseteq P$ is a subset of size $|\alpha| \geq p^{\dim(G)-\frac{1}{4}}$ then $\alpha^3 = P$.*

Proof We first prove (a). By a result of Nori [**115**] there is a constant $I = I_{\exp}(n)$ such that there is a $Frob_p$-invariant connected closed subgroup $G \leq GL(n, \overline{\mathbb{F}_p})$ with $P \leq G(\mathbb{F}_p)$ of index $\big|G(\mathbb{F}_p) : P\big| \leq I$. Clearly $\dim(G) \leq n^2$. By [**94**, Proposition 3] there is an upper bound $\Delta_{\exp}(n) \geq \deg(G)$ (which can also be proved easily from [**115**] using the degree of the exponential map) and by Proposition 2.5.4 we can also assume that it is also an upper bound on the other numerical invariants $\mathrm{mult}(G)$ and $\mathrm{inv}(G)$. Let $\sigma : G \to G$ denote the restriction to G of the automorphism $Frob_p : G \to G$ of Definition 2.11.1, then $G(\mathbb{F}_p) = G^\sigma$ by Remark 2.11.2.

Assume now that P is perfect. Let $\phi : P \to GL(k, \mathbb{C})$ be a nontrivial complex representation. If $k < \frac{p-1}{2}$ then by well-known results of Brauer and Feit-Thompson (see e.g. [**81**, Theorem 14.11] and the remark after its proof) $\phi(P)$ has a normal Sylow-p subgroup. This is impossible since $\phi(P)$ is also a perfect p-generated group. This proves (b).

If K is large enough then $p \geq K^{1/n^2}$ is large as well, hence by Proposition 2.11.3 we have $|P| \leq 2p^{\dim(G)}$ and $\alpha^3 = P$ by Proposition 2.12.3.

Proposition 2.12.8. *Let $H \leq GL(n, \overline{\mathbb{F}})$ be a closed subgroup. Then for some $n' = n'\big(n, \deg(H)\big)$ there is a homomorphism $\phi_H : \mathcal{N}_{GL(n,\overline{\mathbb{F}})}(H) \to GL(n', \overline{\mathbb{F}})$ of degree bounded by n and $\deg(H)$ whose kernel is H. Moreover, if $\overline{\mathbb{F}}$ has characteristic p and H is $Frob_q$-invariant for some p-power q then the homomorphism ϕ_H we construct is $Frob_q$-equivariant (see Definition 2.11.1.*(b)*).*

This proposition is a mild strengthening of [**79**, Theorem 11.5], and it is rather clear that the proof can easily be modified to yield this version. Since we did not find a good reference, we reproduce here the argument. The modified proof is based on the notion of *families of subgroups*, we recall the definition and prove some of their basic properties.

Throughout the proof the adjectives ($Frob_q$-invariant) and ($Frob_q$-equivariant) appearing in parenthesis apply only in the case when $\mathcal{H}$ is $Frob_q$-invariant.

Definition 2.12.9. To simplify the notation let $G = GL(n, \overline{\mathbb{F}})$. Suppose that T is an affine algebraic set and $\mathcal{H} \subseteq T \times G$ is a closed subset. As in [**79**], let $K[G]$ and $K[T \times G]$ denote the coordinate rings of G and $T \times G$ respectively. For each point $t \in T$ we consider the closed subset $\mathcal{H}_t \subseteq G$ defined via the equation $\{t\} \times \mathcal{H}_t = \mathcal{H} \cap \big(\{t\} \times G\big)$. We call $\mathcal{H}$ a *family of subgroups* if $\mathcal{H}_t$ is a subgroup of G for each $t \in T$. In this case we call T the *parameter space* and $\mathcal{H}_t$ are the *members* of the family. Similarly, for vectorspaces V and W, a closed subset $\mathcal{M} \subseteq T \times W$ is a *family of subspaces* if each $\mathcal{M}_t \subseteq W$ is a subspace of W, and a closed subset $\mathcal{L} \subseteq T \times V$ is called a *family of lines* if each $\mathcal{L}_t \subseteq V$ is a line through the origin. A morphism from a family of subgroups $\mathcal{H}$ of $GL(n, \overline{\mathbb{F}})$ to another group $GL(m, \overline{\mathbb{F}})$ is a *family of homomorphisms* if the induced morphisms $\mathcal{H}_t \to GL(m, \overline{\mathbb{F}})$ are all homomorphisms.

Claim 2.12.10. *Let T be an affine algebraic set and $F < K[T \times G]$ a finite dimensional subspace. Then the smallest G-invariant subspace $W < K[T \times G]$ containing F is finite dimensional. Moreover, if T and F are $Frob_q$-invariant then W is also $Frob_q$-invariant.*

Proof G acts on $T \times G$ via the right multiplication in the second factor. Then W is finite dimensional by [**79**, Proposition 8.6], and the $Frob_q$-invariance is obvious.

Claim 2.12.11. *Let $\mathcal{H} \subseteq T \times G$ be a family of subgroups. Then there is a rational representation $\psi : G \to GL(V)$, a dense open subset $U \subseteq T$ and a family of lines $\mathcal{L} \subset U \times V$ such that*

$$\mathcal{H}_t = \left\{ g \in G \,\middle|\, \psi(g)\mathcal{L}_t = \mathcal{L}_t \right\}$$

for all $t \in U$. Moreover, if $\mathcal{H}$ is $Frob_q$-invariant then our construction yields $Frob_q$-invariant ψ, U and $\mathcal{L}$.

Proof We shall imitate [**79**, proof of 11.2]. Let $I \lhd K[T \times G]$ denote the ideal of $\mathcal{H}$ (i.e. the set of those functions vanishing on $\mathcal{H}$) and $I_t \lhd K[G]$ for $t \in T$ the ideal of $\mathcal{H}_t$. Then I is generated by a ($Frob_q$-invariant) finite dimensional subspace $F \leq K[T \times G]$. By Claim 2.12.10 there is a finite dimensional G-invariant subspace $W < K[T \times G]$ containing F (which is also $Frob_q$-invariant). For each $t \in T$ the restriction of functions to $\{t\} \times G$ is a ring homomorphism $r_t : K[T \times G] \to K[G]$.

The closed subset of G corresponding to the ideal $r_t(I)$ is precisely $\mathcal{H}_t$, but the ideal $r_t(I)$ may not be a radical ideal, hence it is not necessarily equal to I_t. It is folklore that there is a ($Frob_p$-invariant) dense open subset $T^* \subseteq T$ such that $r_t(I) = I_t$ for all $t \in T^*$. Here is a quick sketch. We consider the projection morphism $\pi : \mathcal{H} \to T$. By [**91**, Theorem I.1.6] there is a canonical open dense subset T' such that the restriction $\pi^{-1}(T') \to T'$ is flat. The fibre of π at the generic points of T' are smooth varieties (i.e. closed subgroups), hence by [**71**, Exercise III/10.2] there is a canonical open dense subset $T^* \subseteq T'$ such that the restriction $\pi^{-1}(T*) \to T*$ is smooth. By [**71**, Theorem III/10.2] the rings $K[G]/r_t(I)$ are regular for all $t \in T^*$. In particular, $r_t(I)$ are radical ideals, hence $r_t(I) = I_t$ for all $t \in T^*$.

We set $\mathcal{M}_t = W \cap r_t^{-1}(I_t)$. Then $\mathcal{M} = \bigcup_t \{t\} \times \mathcal{M}_t \subseteq T^* \times W$ is a family of subspaces, hence the function $t \to \dim(\mathcal{M}_t)$ is an upper semi-continuous function on T^*. Let $T^* = \bigcup_i T_i^*$ be the irreducible decomposition of T^* and $d_i = \max_{t \in T_i^*} \dim(\mathcal{M}_t)$. The set of points $t \in T_i^*$ which satisfy $\dim(\mathcal{M}_t) = d_i$ form an open dense subset $U_i \subseteq T_i^*$. Then $U = \bigcup_i U_i$ is a ($Frob_q$-invariant) open dense subset of T. We set $V = \bigoplus_{j=0}^{\dim(W)} \bigwedge^j W$ and the representation $\psi : G \to GL(V)$ is just the natural G-action on V. For $t \in U_i$ we set $\mathcal{L}_t = \bigwedge^{d_i} \mathcal{M}_t \leq \bigwedge^{d_i} W \leq V$ and let $\psi_t : G \to GL\big(r_t(W)\big)$ be the natural G-action on $r_t(W)$.

Then $\mathcal{L} = \bigcup_{t \in U} \{t\} \times \mathcal{L}_t \subseteq U \times V$ is a ($Frob_q$-invariant) family of lines and for each $t \in U$ the stabiliser of $\mathcal{L}_t$ in $\psi(G)$ is equal to the stabiliser of $\mathcal{M}_t$ in the image of G in $GL(W)$, which is in turn equal the stabiliser of $r_t(M_t) = I_t \cap r_t(W)$ in $\psi_t(G)$. On the other hand this last stabiliser is just $\mathcal{H}_t$ by [**79**, proof of 11.2].

Claim 2.12.12. *Let $\mathcal{H} \subseteq T \times G$ be a family of subgroups. Then there is a family of homomorphisms $\phi : \mathcal{N}_G(\mathcal{H}_t) \to GL(n', \overline{\mathbb{F}})$ for a common value of n'. In particular, there is a common upper bound on $\deg(\phi_t)$. Moreover, if $\mathcal{H}$ is $Frob_q$-invariant then our construction yields a $Frob_q$-equivariant ϕ (see Definition 2.11.1.*(b)*).*

Proof We prove the claim by induction on $\dim(T)$. We apply Claim 2.12.11 (and use its notation) to this family of subgroups. We obtain an open dense subset $U \subseteq T$. Then $\dim(T \setminus U) < \dim(T)$ so by the induction hypothesis for each ($Frob_q$-invariant) $t \in T \setminus U$ there is a ($Frob_q$-equivariant) embedding $\mathcal{N}_G(\mathcal{H}_t)/\mathcal{H}_t \to GL(n'', \overline{\mathbb{F}})$ with a common n'' and a common bound on their degrees.

Consider any ($Frob_q$-invariant) point $t \in U$ and apply [**79**, proof of Theorem 11.5] to the subgroup $N = \mathcal{H}_t$ of $\mathcal{N}_G(\mathcal{H}_t)$ (which is denoted there by G). For the representation and the line at the beginning of that proof we may choose our $G \to GL(V)$ and $\mathcal{L}_t \leq V$. The proof then constructs a representation $\phi_{\mathcal{H}_t} : \mathcal{N}_G(\mathcal{H}_t) \to GL(W)$ whose kernel is just $\mathcal{H}_t$. Moreover, the homomorphisms $\phi_{\mathcal{H}_t}$ together form a family of homomorphisms $G \times T \to GL(W)$, hence there is a common upper bound on their degrees. The construction is uniquely determined, so it must be $Frob_q$-equivariant whenever $\mathcal{H}$ and t are so. Moreover, by construction $\dim(W) \leq \dim(V)^2$, hence the Claim is valid with $n' = \max\left(n'', \dim(V)^2\right)$.

Proof [Proof of Proposition 2.12.8] By [**91**, Section I.3] there is a canonical open subset of the Chow variety of the projectivisation of G which parametrises all the closed subgroups of G of degree $\deg(H)$. This open subset is not neccessarily affine, but it is defined over $\mathbb{F}_q$, hence it is the union of finitely many $Frob_q$-invariant affine subvarieties. Hence there is a $Frob_q$-invariant family of subgroups which contains (as members) all the closed subgroups of G of degree $\deg(H)$. The proposition follows from Claim 2.12.12 applied to this family.

The proofs of all the results obtained in this section concerning not necessarily simple subgroups of $GL(n, \mathbb{F}_p)$ rest on the following somewhat technical consequence of Theorem 2.11.5. This theorem complements the results about growth of generating sets of simple groups. It would be most interesting to establish an appropriate analogue for subgroups of $GL(n, \mathbb{F}_q)$.

Theorem 2.12.13. *For all parameters $n > 0$ there is a real $\varepsilon = \varepsilon(n) > 0$ with the following property.*
Let $P \leq GL(n, \mathbb{F}_p)$ be a perfect p-generated subgroup. Let $1 \in \alpha \subseteq P$ be a symmetric generating set which projects onto each simple quotient of P. Then either $\alpha^3 = P$ or

$$|\alpha^3| \geq |\alpha|^{1+\varepsilon} .$$

Moreover, the diameter of the Cayley graph of P with respect to α is at most $d(n)$ where $d(n)$ depends on n.

Proof Let l be the smallest integer such that $|P| \leq p^{l/2}$, note that $l \leq 2n^2$. We prove the first statement (concerning α^3) by induction on l. For $l = 0$ it is clear. We assume that $l > 0$ and the statement holds for all groups of order at most $p^{(l-1)/2}$ and for all matrix sizes n with an ε-value $\varepsilon'(n, l) \leq 1$.

We apply Proposition 2.12.7 to P and obtain the bounds $I_{\exp}$, $\Delta_{\exp}$, $K_{\exp}$ (which depend only on n) and the $Frob_p$-invariant connected closed subgroup $G \leq GL(n, \overline{\mathbb{F}_p})$ for which $\left|G(\mathbb{F}_p) : P\right| \leq I_{\exp}$ and $\dim(G) \leq n^2$. We shall apply Theorem 2.11.5 with parameter $\varepsilon'' = \frac{1}{4\dim(G)}$ and obtain the

constants

$$\delta = \tfrac{\varepsilon''}{128\dim(G)^3}\ , \quad M_{\rm main} = M_{\rm main}\big(\dim(G), \varepsilon''\big)\ ,$$
$$K_{\rm main} = K_{\rm main}\big(\dim(G), \Delta_{\rm exp}, I_{\rm exp}, \varepsilon''\big)\ .$$

We shall choose later a real $K \geq \max\big(K_{\rm main}, K_{\rm exp}\big)$. If $|\alpha| \leq K$ and $\alpha^3 \neq P$ then $|\alpha^3| \geq 2|\alpha|$ by Proposition 2.12.1 and the induction step is complete in this case with any $\varepsilon \geq \log(2)/\log(K)$. So we may assume that $|\alpha| > K$. If $|\alpha| > p^{\dim(G)-1/4}$ then $\alpha^3 = P$ by Proposition 2.12.7.(c). So we assume

$$K < |\alpha| \leq p^{\dim(G)-\frac{1}{4}}\ .$$

Consider all *Frob*$_p$-invariant connected closed normal subgroups $1 \neq H \lhd G$ of degree $\deg(H) \leq K_{\rm main}$. Then by Proposition 2.11.4.(b), for sufficiently large K either $H = G$ or $\alpha \not\subseteq H$. By Proposition 2.12.8 there is a *Frob*$_p$-equivariant homomorphism $G \to GL(n', \overline{\mathbb{F}_p})$ for some common $n' = n'\big(\dim(G), K_{\rm main}\big)$ whose kernel is H. The elements of α are fixpoints of *Frob*$_p$, so by the equivariance their images are also fixpoints of *Frob*$_p$ (see Definition 2.11.1.(b)), i.e. the image set α_H of α generates a subgroup of $GL(n', \mathbb{F}_p)$ isomorphic to $P/(H \cap P)$. This subgroup is again perfect, p-generated and α_H projects onto each of its simple quotients. In particular, if $H \neq G$ i.e. $\alpha_H \neq \{1\}$ then $\big|\alpha_H\big| \geq p \geq |\alpha|^{1/n^2}$.We know from Proposition 2.11.3 that if K is large enough then $|H \cap P| \geq \big|H(\mathbb{F}_p)\big|/I_{\rm exp} > \sqrt{p}$ so $\big|P/(H \cap P)\big| < |P|/\sqrt{p} \leq p^{(l-1)/2}$ and the induction hypothesis holds for α_H and $P/(H \cap P)$ with the ε-value $\varepsilon' = \varepsilon'\big(n', l\big) \leq 1$.

Suppose that we find such an H different from G and $\big|\alpha_H^3\big| \geq \big|\alpha_H\big|^{1+\varepsilon'}$. Then using Proposition 2.6.2 we obtain

$$\big|\alpha^5\big| \geq \big|\alpha_H^3\big| \cdot \big|\alpha^2 \cap H\big| \geq \big|\alpha_H\big|^{1+\varepsilon'} \cdot \big|\alpha^2 \cap H\big| \geq \big|\alpha\big| \cdot \big|\alpha_H\big|^{\varepsilon'} \geq \big|\alpha\big|^{1+\varepsilon'/n^2}$$

and by Proposition 2.12.2.(b) the induction step is complete. So we may assume that for all such H we have $\alpha_H^3 = P/(H \cap P)$. It follows from Corollary 2.11.4.(b) that if K is sufficiently large then

$$\big|\alpha_H^3\big| = \big|P/(H \cap P)\big| \geq \big|P\big|^{1-\dim(H)/\dim(G)-\delta/(2n^2)} \geq \big|\alpha\big|^{1-\dim(H)/\dim(G)-\delta/(2n^2)}\ .$$

Suppose next that $\mathcal{Z}(G)$ is finite. We apply Theorem 2.11.5 with parameters $\dim(G)$, $\Delta_{\rm exp}$, $I_{\rm exp}$ and $\varepsilon'' = \frac{1}{4\dim(G)}$ to the subset $\alpha \subset G^{Frob_p}$. We obtain a *Frob*$_p$-invariant connected closed normal subgroup $H \lhd G$ such that $\deg(H) \leq K_{\rm main}$, $\dim(H) > 0$ and

$$\big|\alpha^{M_{\rm main}} \cap H\big| \geq \big|\alpha\big|^{(1+\delta)\dim(H)/\dim(G)}\ .$$

If $H = G$ then $\big|\alpha^{M_{\rm main}}\big| \geq \big|\alpha\big|^{(1+\delta)}$, otherwise

$$\big|\alpha^{3+M_{\rm main}}\big| \geq \big|\alpha_H^3\big| \cdot \big|\alpha^{M_{\rm main}} \cap H\big| \geq \big|\alpha\big|^{1-\frac{\dim(H)}{\dim(G)}-\frac{\delta}{2n^2}} \cdot \big|\alpha\big|^{(1+\delta)\frac{\dim(H)}{\dim(G)}} \geq \big|\alpha\big|^{1+\frac{\delta}{2n^2}}\ .$$

By Proposition 2.12.2.(b) the induction step is complete in this case as well.

Finally we suppose that $\mathcal{Z}(G)$ is infinite. In this case we consider the normal subgroup $H = \mathcal{Z}(G)^0$. By assumption $\alpha_H^3 = P/(H \cap P)$ hence α^3 intersects every $(H \cap P)$-coset in P. Hence every commutator element of P is in fact the commutator of two elements in α^3. It is well-known that P is

generated by at most n^2 elements (see [**130**]) hence by Proposition 2.12.6 each element of P is the product of Cn^6 commutators for some constant C. By assumption $|\alpha| \leq p^{\dim(G)-1/4}$. Since $|P| \geq |\alpha| > K$, if we choose K sufficiently large then $|P| \geq p^{\dim(G)-1/8}$ by Proposition 2.11.3. Therefore

$$\left|\alpha^{3\cdot 4\cdot Cn^6}\right| = |P| > |\alpha|^{1+1/8\dim(G)}$$

and by Proposition 2.12.2.(b) the induction step is complete in this case too. The first statement is proved.

Let us apply the (now established) first statement successively to $\alpha, \alpha^3, \alpha^9, \ldots$. We obtain by induction that either $\alpha^{3^i} = P$ or $\left|\alpha^{3^i}\right| \geq |\alpha|^{(1+\varepsilon)^i}$ for all i. By assumption $|\alpha| \geq p$ and $|P| < p^{n^2}$ hence $\alpha^{d(n)} = P$ where $d(n)$ is the smallest integer above $n^{2\log(3)/\log(1+\varepsilon)}$. That is, the diameter of the Cayley graph with respect to α is at most $d(n)$.

Now we prove Theorem 2.1.8 of the Introduction.

Theorem 2.12.14. *For all natural numbers n there is an integer $M = M(n)$ with the following property.*
Let $P \leq GL(n, \mathbb{F}_p)$ be a perfect p-generated subgroup. Then the diameter of the Cayley graph of P with respect to any symmetric generating set is at most $\left(\log|P|\right)^M$.

Proof Let α be a symmetric generating set of P containing 1. Let L be any simple quotient of P, we denote by $\tilde{\alpha}$ the image of α in L. The Lie rank of L is at most n (see [**54**] and [**89**, Proposition 5.2.12]). Let $\varepsilon = \varepsilon(n)$ be as in Theorem 2.12.5. Applying that theorem successively to $\tilde{\alpha}, \tilde{\alpha}^3, \tilde{\alpha}^9, \ldots$ we obtain by induction that either $\tilde{\alpha}^{3^i} = L$ or $\left|\tilde{\alpha}^{3^i}\right| \geq |\tilde{\alpha}|^{(1+\varepsilon)^i} \geq 3^{(1+\varepsilon)^i}$ for all i. With $m = \frac{\log\log|P| - \log\log(3)}{\log(1+\varepsilon)}$ we obtain that $\left|\tilde{\alpha}^{3^m}\right| \geq |P| \geq |L|$ hence α^{3^m} projects onto L. This holds for each simple quotient with the same exponent m.

By Theorem 2.12.13 the diameter of the Cayley graph corresponding to α^{3^m} is at most $d(n)$, hence the diameter of the Cayley graph corresponding to α is at most $3^m d(n) \leq \left(\log|P|\right)^{M(n)}$ where $M(n)$ is the smallest integer above $\frac{\log(3)}{\log(1+\varepsilon)} + \log\left(d(n)\right)$

We will reduce the proof of Theorem 2.1.7 to the perfect p-generated case (more precisely to Theorem 2.12.13) using finite group theory.

Definition 2.12.15. As usual $Sol(G)$ denotes the *soluble radical* and $O_p(G)$ the maximal normal p-subgroup of a finite group G. A group is called *quasi-simple* if it is perfect and simple modulo its centre. We denote by $Lie^*(p)$ the set of direct products of simple groups of Lie type of characteristic p, and by $Lie^{**}(p)$ the set of central products of quasi-simple groups of Lie type of characteristic p. If $G/Sol(G)$ is in $Lie^*(p)$ then we call G a *soluble by $Lie^*(p)$ group.*

The following deep result is essentially due to Weisfeiler [**169**].

Proposition 2.12.16. *Let G be a finite subgroup of $GL(n, \mathbb{F})$ where $\mathbb{F}$ is a field of characteristic $p > 0$. Then G has a normal subgroup H of index at most $f(n)$ such that $H \geq O_p(G)$ and $H/O_p(G)$ is the central product of*

an abelian p'-group and quasi-simple groups of Lie type of characteristic p, where the bound $f(n)$ depends on n.

It was proved by Collins [**33**] that for $n \geq 71$ one can take $f(n) = (n+2)!$. Remarkably a (non-effective) version of the above result was obtained by Larsen and Pink [**95**] without relying on the classification of finite simple groups. It is clear that H is a soluble by $Lie^*(p)$ subgroup.

Remark 2.12.17. Let P be a perfect p-generated subgroup of $GL(n, \mathbb{F}_p)$. Using Proposition 2.12.16 and [**70**, Lemma 3] one can easily show that every element of P is the product of $g(n)$ commutators where $g(n)$ depends on n. This could be used to replace the (rather more difficult) Proposition 2.12.6 in the proof of Theorem 2.12.13.

The rest of this section will be devoted to proving results concerning subsets α of $GL(n, \mathbb{F}_p)$ that satisfy $\left|\alpha^3\right| \leq K|\alpha|$. We consider the group $G = \langle \alpha \rangle$ and we will establish step by step a close relationship between α (and its powers) and the structure of G described in Proposition 2.12.16. Throughout the proof we need to establish several auxiliary results.

Proposition 2.12.18. *Let G be a group and $\alpha \subseteq G$ a symmetric generating set with $1 \in \alpha$. If H is a normal subgroup of index t in G then $\alpha^{2t} \cap H$ generates H.*

Proof It is clear that α^{t-1} contains a full system of coset representatives $g_1, \ldots, g_t$ of G/H. It is well-known (see [**150**, Theorem 2.6.9]) that H is generated by elements of the form $g_i a g_j^{-1}$ where $a \in \alpha$.

Proposition 2.12.19. *Let α be a finite subset of a group G and $\tilde{G} = G/N$ a quotient of G. Set $\tilde{\alpha} = \alpha N/N$. Then $|\alpha^4|/|\alpha| \geq |\tilde{\alpha}^3|/|\tilde{\alpha}|$. Moreover, if α is symmetric and $1 \in \alpha$ then $\left(|\tilde{\alpha}^3|/|\tilde{\alpha}|\right)^2 \geq |\tilde{\alpha}^3|/|\tilde{\alpha}|$.*

Proof There is a coset gN of N such that $|\alpha \cap gN| \geq |\alpha|/|\tilde{\alpha}|$. We may assume that $g \in \alpha$. Let $\{g_i\}$ be a system of representatives of the cosets in $\tilde{\alpha}^3$ with $g_i \in \alpha^3$. Then the sets $g_i(\alpha \cap gN)$ are disjoint subsets of α^4 hence $\left|\alpha^4\right| \geq \left|\tilde{\alpha}^3\right||\alpha|/|\tilde{\alpha}|$ as required. The other inequality follows then from Proposition 2.12.2.(b).

Proposition 2.12.20. *Let H be a soluble by $Lie^*(p)$ subgroup of $GL(n, \mathbb{F}_p)$ and $\gamma \leq H$ a symmetric generating set with $1 \in \gamma$. Assume that γ satisfies $|\gamma^3| \leq K|\gamma|$ for some $K > 2$. Then there is a soluble by $Lie^*(p)$ normal subgroup S of H such that $\gamma^6 \cap S$ projects onto all Lie type simple quotients of S and γ is covered by K^c cosets of S, where $c = c(n)$ depends only on n.*

Proof Let $H/N \cong L$ be a Lie type simple quotient of H and set $\tilde{\gamma} = \gamma N/N$. The Lie rank of L is at most n (see [**54**] and [**89**, Proposition 5.2.12]). Now $|\tilde{\gamma}^3| \leq K^2|\tilde{\gamma}|$ by Proposition 2.12.19. Hence by Theorem 2.12.5 we have two possibilities; either $|\tilde{\gamma}| \geq |\tilde{\gamma}^3|/K^2 = |L|/K^2$ or $|\tilde{\gamma}| \leq K^b$ where $b = b(n)$ depends only on n. Set $c = 6n^2(2 + nb)$. If $(p-1)/2 \leq K^{3(2+nb)}$ then we have $|GL(n, \mathbb{F}_p)| < K^c$ (since $K > 2$) and our statement holds for $S = 1$.

Otherwise let $H/N_j \cong L_j$ $(j = 1, .., t)$ be all the Lie type simple quotients of H (there are at most n such quotients e.g. by [**101**, Corollary 3.3]). Let $H/N_1, H/N_2, \ldots, H/N_i$ be the quotients for which the second possibility

holds . Consider the subgroup $S = N_1 \cap \cdots \cap N_i$. It is clear that S is a soluble by $Lie^*(p)$ normal subgroup and its Lie type simple quotients are $S/(S \cap N_{i+1}), .., S/(S \cap N_t)$. Moreover γ is covered by at most K^{nb} cosets of S.

It remains to prove that $\gamma^6 \cap S$ projects onto, say, $S/(S \cap N_{i+1})$. Consider the quotient group $\overline{H} = H/(S \cap N_{i+1})$. The image $\overline{\gamma}$ of γ in $\overline{H}$ is covered by at most K^{nb} cosets of $\overline{S} = S/(S \cap N_{i+1}) \cong L_{i+1}$ and we have $|\overline{\gamma}| \geq |\overline{S}|/K^2$. This implies that some coset of $\overline{S}$ in $\overline{H}$ contains at least $|\overline{S}|/K^{2+nb}$ elements of $\overline{\gamma}$ and it follows that $\left|\overline{\gamma}^2 \cap \overline{S}\right| \geq |\overline{S}|/K^{2+nb}$. By Remark 2.12.4 the minimal degree of a complex representation of $\overline{S}$ is at least $(p-1)/2 > (K^{2+nb})^3$ hence by Proposition 2.12.3 we have $\left(\overline{\gamma}^2 \cap \overline{S}\right)^3 = \overline{S}$, which implies our statement.

Proposition 2.12.21. *Assume that a symmetric subset α of a group G is covered by x right cosets of a subgroup H and $\alpha^2 \cap H$ is covered by y right cosets of a subgroup $S \leq H$. Then α is covered by xy right cosets of S.*

Proof We have $\alpha \subseteq Hg_1 \cup \cdots \cup Hg_x$ and $\alpha^2 \cap H \subseteq Sh_1 \cup \cdots \cup Sh_y$ where the coset representatives g_i are chosen from α. If $a \in \alpha \cap Hg_i$ then by our assumptions $ag_i^{-1} \in Sh_j$ for some j, hence $a \in Sh_jg_i$. Therefore $\alpha \subseteq \bigcup_i \bigcup_j Sh_jg_i$.

Proposition 2.12.22. *Let G and H be as in Proposition 2.12.16. Let α be a symmetric set of generators of G with $1 \in \alpha$ satisfying $\left|\alpha^3\right| \leq K|\alpha|$ for some $K > 2$. Set $\gamma = \alpha^{2f(n)} \cap H$.*

a) The set γ generates H and satisfies $\left|\gamma^3\right| \leq K_0|\gamma|$ where $K_0 = K^{7f(n)}$.

b) Let S be the subgroup constructed from γ and H in the proof of Proposition 2.12.20. If $p \geq K_0^{b_0(n)}$ (where $b_0(n) = b(n) + 4$ with the same $b(n)$ as in the proof of Proposition 2.12.20) then S is normal in G.

c) α is covered by at most $K_0^{c_0(n)}$ cosets of S (where $c_0(n) = c(n) + \log\left(f(n)\right)/\log(2)$ with the same $c(n)$ as in Proposition 2.12.20).

d) The commutator subgroup S' is an extension of a p-group by a $Lie^{**}(p)$-group.

Proof Consider $\beta = \alpha^{f(n)}$. By Proposition 2.12.18 $\gamma = \beta^2 \cap H$ generates H. Using Lemma 2.7.1 and Proposition 2.12.2 we see that

$$\frac{\left|\gamma^3\right|}{|\gamma|} \leq \frac{\left|\beta^6 \cap H\right|}{\left|\beta^2 \cap H\right|} \leq \frac{\left|\beta^7\right|}{|\beta|} \leq \frac{\left|\alpha^{7f(n)}\right|}{|\alpha|} \leq K^{7f(n)}$$

which proves (a). Part (c) follows using Proposition 2.12.21. Part (d) follows from Proposition 2.12.16.

It remains to prove (b). If H/N_j are all the Lie type simple quotients of H then $N = \bigcap_j N_j$ is the soluble radical of H. Consider the quotient $\overline{G} = G/N$. The set $\overline{\gamma}$ generates the normal subgroup $\overline{H} \lhd \overline{G}$. For each $a \in \alpha$ the conjugation by $\overline{a} \in \overline{\alpha}$ is an automorphism of $\overline{H}$. Now $\overline{H}$ is the direct product of nonabelian simple groups and an automorphism of $\overline{H}$ permutes these factors (because the direct decomposition is unique).

If S is not normal in G then there is a Lie type simple quotient of H, say $H/N_1 \cong L_1$ and an element $a \in \alpha$ such that γ projects onto at most $K_0^{b(n)}$ elements of H/N_1 and $a^{-1}\gamma a$ projects onto at least $|L_1|/K_0^2$ elements

of H/N_1. Note that $a^{-1}\gamma a = a^{-1}(\beta^2 \cap H)a \subseteq \beta^4 \cap H$. By the above we have $|\beta^2 \cap H| = |\gamma| \leq \left|\gamma^2 \cap N_1\right| K_0^{b(n)}$. On the other hand,

$$\left|\beta^8 \cap H\right| \geq \left|(\beta^4 \cap H)(\beta^2 \cap H)^2\right| \geq \left|(a^{-1}\gamma a)(\gamma^2 \cap N_1)\right| \geq \frac{|L_1|}{K_0^2}\left|\gamma^2 \cap N_1\right|.$$

Therefore $\frac{|\beta^8 \cap H|}{|\beta^2 \cap H|} \geq |L_1|/K_0^{2+b(n)}$. But we have $\frac{|\beta^8 \cap H|}{|\beta^2 \cap H|} \leq \frac{|\beta^9|}{|\beta|} \leq K^{9f(n)} < K_0^2$. We obtain that $|L_1| < K_0^{4+b(n)}$, a contradiction.

As we saw above, a subset α of $GL(n, \mathbb{F}_p)$ with $\left|\alpha^3\right| \leq K|\alpha|$ is essentially contained in a normal subgroup S of $G = \langle \alpha \rangle$ such that a small power of α projects onto all Lie type simple quotients of S. We proceed to show that the latter property also holds for the last term P of the derived series of S. Later we will prove that a small power of α in fact generates P (see Proposition 2.12.28).

Proposition 2.12.23. *Let S be a soluble by $Lie^*(p)$ subgroup of $GL(n, \mathbb{F}_p)$. Let $1 \in \alpha$ be a symmetric subset of S which projects onto all Lie type simple quotients of S. Let P be the last term of the derived series of S. Then P is a perfect soluble by $Lie^*(p)$ subgroup and $\alpha^c \cap P$ projects onto all Lie type simple quotients of P where $c = c(n)$ depends only on n.*

Proof Let S/N_i be the Lie type simple quotients of S. The commutator subgroup S' is clearly also a soluble by $Lie^*(p)$ subgroup and its Lie type simple quotients are the $S'/(S' \cap N_i) \cong S/N_i$. We need the following.

Claim 2.12.24. *$S' \cap \alpha^b$ projects onto $S'/(S' \cap N_i)$ for all i where $b = b(n)$ depends only on n.*

To see this fix i and consider the quotient $\overline{S} = S/(S' \cap N_i)$. This quotient is the direct product of $S'/(S' \cap N_i)$ and $N_i/(S' \cap N_i) \cong S/S'$ (since these have no common quotients). Take two elements $a, b \in \alpha$ which project onto noncommuting elements of S/N_i. The image of the commutator $[a, b] \in \alpha^4$ in $\overline{S}$ is a nontrivial element of $S'/(S' \cap N_i)$. Each element of $S' \cap N_i$ appears as the first coordinate of some element of the image $\overline{\alpha}$ of α in $\overline{S}$. Taking conjugates of $\overline{[a, b]}$ with these elements we obtain that the whole conjugacy class of $[a, b]$ in the simple group $S'/(S' \cap N_i)$. But this group has Lie rank at most n and therefore each element of $S'/(S' \cap N_i)$ is the product of at most $a(n)$ conjugates of an arbitrary nontrivial element where $a(n)$ depends only on n (in fact $a(n)$ is a linear function of n by [**96**]). Therefore $\alpha^{6a(n)} \cap S'$ projects onto $S'/(S' \cap N_i)$ as claimed.

The length of the derived series of any subgroup of $GL(n, \mathbb{F}_p)$ is bounded in n (in fact there is a logarithmic bound). Hence our statement follows from the Claim by an obvious induction argument.

Definition 2.12.25. If $L = L_1 \times \cdots \times L_k$ is a direct product of isomorphic groups, D a subgroup of L isomorphic to L_1 which projects onto each direct factor then we call D a *diagonal subgroup*.

Proposition 2.12.26. *Let $L = L_1 \times \cdots \times L_k$ be a direct product of k nonabelian simple groups and T a subgroup which projects onto all simple quotients of L. Then any chain of subgroups between T and L has length at most k.*

Proof Let H be a subgroup of L which projects onto all simple quotients of L (i.e. a subdirect product). Then there is a partition of the set of simple groups L_i such that the groups in any partition-class are isomorphic and H is the direct product of diagonal subgroups corresponding to these partition-classes (see [**7**, Proposition 3.3]). Our statement follows.

Proposition 2.12.27. *Let L be a $Lie^{**}(p)$-group and T a subgroup which projects onto $L/\mathcal{Z}(L)$. Then $T = L$.*

Proof We have $T\mathcal{Z}(L) = L$ which implies that T is a normal subgroup of L. Moreover, L/T is abelian and since L is perfect, we have $T = L$.

Proposition 2.12.28. *Let H be a subgroup of $GL(n, \mathbb{F}_p)$, S a soluble by $Lie^*(p)$ normal subgroup of H and P the last term in the derived series of S. Assume that P is an extension of a p-group by a $Lie^{**}(p)$-group. Let $1 \in \gamma$ be a symmetric generating set of H. Assume that $\gamma^t \cap P$ projects onto all Lie type simple quotients of P for some integer t. Then $\gamma^{t+2n+2n^2} \cap P$ generates P.*

Proof Set $Q_i = \langle \gamma^i \cap P \rangle$. We first show that Q_{t+2n} projects onto $P/O_p(P)$. Since $P/O_p(P)$ is a $Lie^{**}(p)$-group, by Proposition 2.12.27 it is sufficient to prove that Q_{t+2n} projects onto the central quotient of $P/O_p(P)$, which is exactly $P/Sol(P)$. Denote $P/Sol(P)$ by $\overline{P}$ and let $\overline{Q_i}$ denote the image of Q_i in $\overline{P}$. We need the following.

Claim 2.12.29. *If $i \geq t$ and $\overline{Q_i} \neq \overline{P}$ then $|Q_{i+2}|$ is strictly greater than $|Q_i|$.*

To see this, observe that $\overline{Q_i}$ projects onto all simple quotients of $\overline{P}$ and the only normal subgroup of $\overline{P}$ with this property is $\overline{P}$ itself. By our assumptions there is an $a \in \gamma$ for which $\overline{Q_i}$ and its conjugate $\overline{Q_i}^a$ are different subgroups of $\overline{Q_{i+2}}$. This implies the claim.

As noted earlier, $\overline{P}$ is the direct product of at most n simple groups. Hence by Proposition 2.12.26 any chain of subgroups containing $\overline{Q_t}$ has length at most n. By the above claim Q_{t+2n} projects onto $P/Sol(P)$, hence onto $P/O_p(P)$ as stated. We also need the following.

Claim 2.12.30. *If Q_i is not a normal subgroup of H and $i \geq t + 2n$ then $|Q_{i+2}| \geq |Q_i| \cdot p$.*

To see this, consider as above an element $a \in \gamma$ which does not normalise Q_i. Then Q_i and Q_i^a are different subgroups of P generated by subsets of γ^{i+2}. Hence $P \geq Q_{i+2} \gneq Q_i$. By our assumptions $|P : Q_i|$ is a power of p which implies the Claim.

Repeated applications of the Claim yield an ascending chain of subgroups $Q_{t+2n} \lneq Q_{t+2n+2} \lneq Q_{t+2n+4} \lneq \cdots \lneq Q_{t+2n+2k} = Q \leq P$ which of course has length less than n^2. The last term Q of this chain is normal in H hence in S. By our assumptions all nonabelian simple composition factors of S are among the composition factors of Q (with multiplicities). Therefore S/Q must be soluble i.e. $Q = P$.

Proposition 2.12.31. *Let G be a finite group and α a generating set such that α^k contains the subgroup P. Then*

$$\frac{\max_{g\in G}|\alpha\cap gP|}{|P|}\geq\frac{|\alpha|}{\left|\alpha^{k+1}\right|}\ .$$

Proof Let t be the number of cosets of P which contain elements of α. Then we have $\max_g|\alpha\cap gP|\cdot t\geq|\alpha|$. On the other hand it is clear that $\left|\alpha^{k+1}\right|\geq t|P|$. Hence

$$\frac{\left|\alpha^{k+1}\right|}{|P|}\geq t\geq\frac{|\alpha|}{\max_{g\in G}|\alpha\cap gP|}$$

as required.

Now we are ready to prove our main results concerning subsets α of $GL(n,\mathbb{F}_p)$ with $\left|\alpha^3\right|\leq K|\alpha|$.

Theorem 2.12.32. *Let α be a symmetric subset of $GL(n,\mathbb{F}_p)$ satisfying $|\alpha^3|\leq K|\alpha|$ for some $K\geq 1$. Then $GL(n,\mathbb{F}_p)$ has two subgroups $S\geq P$, both normalised by α, such that P is perfect, S/P is soluble, a coset of P contains at least $|P|/K^{c(n)}$ elements of α and α is covered by $K^{c(n)}$ cosets of S where $c(n)$ depends on n.*

Proof If $K\leq 2$ then let S be the subgroup generated by α and P the last term of the derived series of S. By Proposition 2.12.1 we have $\alpha^3=S$ hence $|\alpha|\geq|S|/K$, which implies that some coset of P contains at least $|P|/K$ elements. If $K>2$ and $p<K^{7f(n)b_0(n)}$ (with the notation of Proposition 2.12.22) then we set $S=P=\{1\}$. Now we have $|\alpha|<K^{7f(n)b_0(n)n^2}$ which proves our statement in this case. From now on we assume that $K>2$ and $p\geq K^{7f(n)b_0(n)}$.

Let S be as in Proposition 2.12.22. Then α is covered by $K^{7f(n)c_0(n)}$ cosets of S. By Proposition 2.12.20 the set $\alpha^{12f(n)}\cap S$ projects onto all Lie type simple quotients of S.

Let P be the last term of the derived series of S. Proposition 2.12.22.(d) implies that P is an extension of a p-group by a $Lie^{**}(p)$-group, in particular P is a p-generated group. Let $c_1(n)$ be the constant of Proposition 2.12.23 (denoted there by $c(n)$), set $c_2(n)=2f(n)\big(6c(n)+2n+2n^2\big)$. $\alpha^{c_2(n)}\cap P$ generates P and projects onto all Lie type simple quotients of P by Proposition 2.12.23 and Proposition 2.12.28. By Theorem 2.12.13 if $c(n)\geq c_2(n)d(n)$ then $\alpha^{c(n)}$ contains P.

Using Proposition 2.12.31 and Proposition 2.12.2.(b) we obtain that some coset of P contains at least

$$\frac{|P||\alpha|}{\left|\alpha^{c(n)+1}\right|}\geq\frac{|P|}{K^{c(n)}}$$

elements of α. The proof is complete.

The following is a slightly stronger version of Theorem 2.1.7.

Corollary 2.12.33. *Let α be a symmetric subset of $GL(n,\mathbb{F}_p)$ satisfying $|\alpha^3|\leq K|\alpha|$ for some $K\geq 1$. Then $GL(n,\mathbb{F}_p)$ has two subgroups $S\geq P$, both normalised by α, such that P is perfect, S/P is soluble, a coset of P is contained in α^3 and α is covered by $K^{c(n)}$ cosets of S where $c(n)$ depends on n.*

Proof If $K \leq 2$ then $\alpha^3 = \langle \alpha \rangle$ by Proposition 2.12.1 and our statement follows. Let $c'(n)$ the constant in Theorem 2.12.32. If $\frac{p-1}{2} \leq K^{3c'(n)}$ and $K > 2$ then it follows that $|\alpha| \leq K^{6c'(n)n^2}$ hence our statement holds for $S = P = 1$ with $c(n) = 6c'(n)n^2$.

We assume that $K > 2$ and $K^{3c'(n)} < \frac{p-1}{2}$. Let S and P be as in Theorem 2.12.32. By that theorem there is a subset X of P of size at least $|P|/K^{c'(n)}$ such that $aX \subseteq \alpha$ for some $a \in \alpha$. Now

$$\alpha^3 \supseteq aXaXaX = a^3(a^{-2}Xa^2)(a^{-1}Xa)X \ .$$

By our assumptions and Proposition 2.12.7.(b) if k is the minimal degree of a complex representation of P then we have $\left|a^{-2}Xa^2\right|\left|a^{-1}Xa\right|\left|X\right| \geq \left|P\right|^3/k$. hence by Proposition 2.12.3 we have $\alpha^3 \supset a^3P$ as required.

To obtain a characterisation for symmetric subsets α of $GL(n, \mathbb{F}_p)$ satisfying $\left|\alpha^3\right| \leq K|\alpha|$ with polynomially bounded constants (as in Theorem 2.12.32) seems to be a very difficult task. As another step towards such a characterisation we mention the following (folklore) conjecture.

Conjecture 2.12.34. *Let $1 \in \alpha$ be a symmetric subset of $GL(n, \mathbb{F}_p)$ satisfying $\left|\alpha^3\right| \leq K|\alpha|$ for some $K \geq 1$. Then $GL(n, \mathbb{F}_p)$ has two subgroups $S \triangleright P$ such that S/P is nilpotent, P is contained in $\alpha^{c(n)}$ and α is covered by $K^{c(n)}$ cosets of S where $c(n)$ depends on n.*

The following is well-known.

Proposition 2.12.35. *Let S be a finite group and P a normal subgroup with S/P soluble. If C is a minimal subgroup such that $PC = S$ then C is soluble.*

Proof Let M be a maximal subgroup of C. If M does not contain $C \cap P$ then $(C \cap P)M = C$ which implies $PM = PC = S$, a contradiction. Hence all maximal subgroups of C, and therefore its Frattini subgroup $\Phi(C)$ contain $C \cap P$. But $\Phi(C)$ is nilpotent, hence C is soluble.

Theorem 2.12.32 and Proposition 2.12.35 can be used to show that if Conjecture 2.12.34 holds in the case when $\langle \alpha \rangle$ is soluble then it holds in general. We omit the details. [4]

2.13. Linear groups over arbitrary fields

In this section we develop another method to show that a certain spreading system $\alpha|G$ is (ε, M, K)-spreading. As in the proof of Theorem 2.11.5, we find an appropriate CCC-subgroup $A < G$, but now we study the case when A has infinitely many $\langle \alpha \rangle$-conjugates.

We use the resulting new spreading theorem (Theorem 2.13.1) inductively to show that if α is a non-growing subset of $GL(n, \mathbb{F})$, $\mathbb{F}$ an arbitrary field, then $\langle \alpha \rangle$ is essentially contained in a virtually soluble group (see Corollary 2.13.4).

Combining Corollary 2.13.4 with various results on finite groups, in particular the Product theorem (2.1.4), we obtain Theorem 2.1.10, our main result on arbitrary finitely generated linear groups.

[4] Very recently Gill and Helfgott [**59**] have proved Conjecture 2.12.34 in the soluble case.

Theorem 2.13.1. *For all parameters $N > 0$, $\Delta > 0$ and $\frac{1}{119N^3} > \varepsilon > 0$ there is an integer $M = M_\infty(N, \varepsilon) > 0$ and a real $K = K_\infty(N, \Delta, \varepsilon) > 0$ with the following property.*
Let $\alpha|G$ be an (N, Δ, K)-bounded spreading system. Then either $\langle\alpha\rangle \cap G$ is virtually nilpotent or $\alpha|G$ is (ε, M, K)-spreading. Moreover, our construction of the subgroup of spreading is uniquely determined.

Proof Using the bounds from Lemma 2.10.5 and Lemma 2.10.3 we set

$$M_{\mathrm{CCC}} = M_{\mathrm{CCC}}(N, \varepsilon) \ , \quad M_{\mathrm{s}} = M_{\mathrm{s}}(N, \varepsilon) \ ,$$
$$M = M_{\mathrm{CCC}} \cdot M_{\mathrm{s}} \ ,$$
$$K = \max\Big(\Delta, K_{\mathrm{CCC}}(N, \Delta, \varepsilon), K_{\mathrm{s}}(N, \Delta, \varepsilon)\Big) \ .$$

Suppose that $\alpha|G$ is not (ε, M, K)-spreading. In particular, it is not $(N\varepsilon, M_{\mathrm{CCC}}, K)$-spreading either, hence

$$\mu(\alpha^{M_{\mathrm{CCC}}}, G) < \big(1 + N\varepsilon\big)\mu(\alpha, G) \ .$$

If G is nilpotent then there is nothing to prove, so we assume that G is non-nilpotent. Using Lemma 2.10.5 we obtain a CCC-subgroup $A \subseteq G$ containing a single maximal torus T such that

$$\mu\big(\alpha^{M_{\mathrm{CCC}}}, A\big) > (1 - \varepsilon \cdot 16N)\, \mu(\alpha, G) >$$
$$> \big(1 - \tfrac{1}{7N^2}\big)\,(1 + N\varepsilon)\,\mu(\alpha, G) > \big(1 - \tfrac{1}{7N^2}\big)\,\mu(\alpha^{M_{\mathrm{CCC}}}, G) \ .$$

In particular A is not normal in G. If A has infinitely many $\langle\alpha\rangle$-conjugates then $\alpha^{M_{\mathrm{CCC}}}|G$ is $(\varepsilon, M_{\mathrm{s}}, K)$-spreading by Lemma 2.10.3, a contradiction. So A has finitely many $\langle\alpha\rangle$-conjugates. Then T has finitely many $\langle\alpha\rangle$-conjugates, hence $\langle\alpha\rangle \cap \mathcal{N}_G(T)$ has finite index in $\langle\alpha\rangle \cap G$.

On the other hand $\mathcal{N}_G(T) = \mathcal{N}_G\big(\mathcal{C}_G(T)\big)$, and $\mathcal{C}_G(T)$ is a Cartan subgroup, so it is nilpotent and has finite index in its normaliser. Therefore $\mathcal{N}_G(T)$ is virtually nilpotent, hence $\langle\alpha\rangle \cap G$ is also virtually nilpotent.

Our plan is to apply Theorem 2.13.1, then apply it to the subgroup of spreading, then apply it again to the new subgroup of spreading, and so on, until we eventually arrive to a subgroup whose intersection with $\langle\alpha\rangle$ is virtually nilpotent.

We need the following fact:

Proposition 2.13.2 (Freiman [**55**])**.** *Let α be a finite subset of a group G. If $|\alpha \cdot \alpha| < \frac{3}{2}|\alpha|$, then $S := \alpha \cdot \alpha^{-1}$ is a finite group of order $|\alpha \cdot \alpha|$, and $\alpha \subset S \cdot x = x \cdot S$ for some x in the normaliser of S.*

Proposition 2.13.3. *For all parameters $n > 0$, $d > 0$ there are integers $m = m_{\mathrm{nilp}}(n, d) > 0$ and $D = D_{\mathrm{nilp}}(n, d) > 0$ with the following property.*
Let $G \leq GL(n, \overline{\mathbb{F}})$ be a (possibly non-connected) closed subgroup and $\alpha \leq G$ a finite subset such that $\dim(G) \geq 1$, $\deg(G) \leq d$ and $\big|\alpha^3\big| \leq \mathcal{K}|\alpha|$ for some $\mathcal{K}$. Then either $|\alpha| \leq \mathcal{K}^m$ or one can find a connected closed subgroup $H \leq G$ normalised by α such that $\dim(H) \geq 1$, $\deg(H) \leq D$ and $\langle\alpha\rangle \cap H$ is virtually nilpotent.

Proof During the proof we encounter several lower bounds for m, we assume that our m satisfies them all. Similarly, we shall establish several alternative upper bounds on $\deg(H)$, we set D to be the maximum of these

bounds. If $\mathcal{K} < \frac{3}{2}$ then $\langle\alpha\rangle$ is virtually cyclic by Proposition 2.13.2 and the lemma holds with $H = G^0$. If $\mathcal{C}_G(\alpha)$ is infinite then we take $H = \mathcal{C}_G(\alpha)^0$ (see Fact 2.5.9). So we assume that $\mathcal{K} \geq \frac{3}{2}$, $\mathcal{C}_G(\alpha)$ is finite and $|\alpha| > \mathcal{K}^m$. By Proposition 2.12.2.(a) we can assume that α is symmetric and $1 \in \alpha$. We order the set α.

By assumption $|G : G^0| \leq d$, hence $|\alpha^2 \cap G^0| \geq \frac{|\alpha|}{d}$. We set $\varepsilon = \frac{1}{120n^6}$, $G_0 = G^0$, and construct by induction a sequence of length at most n^2 of connected closed subgroups $G_0 > G_1 > G_2 > \dots$ normalised by α and corresponding constants e_i, K_i such that

$$\dim(G_i) \geq 1\ , \quad \deg(G_i) \leq K_i\ , \quad \left|\alpha^{e_i} \cap G_i\right| \geq \left(\tfrac{|\alpha|}{d}\right)^{\dim(G_i)/n^2}\ .$$

It will be clear from the construction that all of the appearing constants (i.e. e_i, K_i, Δ_i and M, see below) depend only on n and d. We already defined G_0, our statement holds with $K_0 = d$ and $e_0 = 2$ (since closed subgroups of $GL(n, \overline{\mathbb{F}})$ have dimension at most n^2). Suppose that G_i, K_i and e_i are already constructed for some $i \geq 0$. We assume that $\langle\alpha\rangle \cap G_i$ is not virtually nilpotent, since otherwise the lemma holds with $H = G_i$ (whose degree is bounded in terms of n and d). According to Proposition 2.5.4 the numerical invariants $\deg(G_i)$, $\mathrm{mult}(G_i)$ and $\mathrm{inv}(G_i)$ are bounded from above by a certain constant $\Delta_i = \Delta_i(n^2, K_i)$. Recall from Theorem 2.13.1 the constants $M = M_\infty(n^2, \varepsilon)$ and $K_{i+1} = K_\infty(n^2, \Delta_i, \varepsilon)$. We assume that m is large enough so that $\mathcal{K}^m > \left(\frac{3}{2}\right)^m > d\,(K_{i+1})^{n^2}$. Then the $\alpha^e|G_i$ are (n^2, Δ_i, K_{i+1})-bounded spreading systems for all $e \geq e_i$, hence according to Theorem 2.13.1 they are $(\varepsilon, M, K_{i+1})$-spreading.

Let us consider the spreading systems $\alpha^{e_i M^j}|G_i$ for $j = 0, 1, 2, \dots J-1$, where $J = 2\frac{n^2}{\varepsilon} = 240n^8$. Suppose now that for each i, G_i itself is the subgroup of spreading obtained above using Theorem 2.13.1. Then $\mu\big(\alpha^{e_i M^J}, G_i\big) \geq (1+\varepsilon)^J \mu(\alpha, G_i)$ i.e.

$$\left|\alpha^{e_i M^J} \cap G_i\right| \geq \left|\alpha^{e_i} \cap G_i\right|^{(1+\varepsilon)^J} > \left(\tfrac{|\alpha|}{d}\right)^{J\varepsilon/n^2} = \left(\tfrac{|\alpha|}{d}\right)^2 \geq |\alpha|\tfrac{\mathcal{K}^m}{d^2}\ .$$

On the other hand, by Proposition 2.12.2.(b) we have $\left|\alpha^{e_i M^J}\right| \leq |\alpha|\mathcal{K}^{e_i M^J - 2}$. We rule this case out by choosing $m \geq e_i M^J + \frac{\log(d^2)}{\log(3/2)}$. Then there is a value $j_0 < J$ such that the corresponding subgroup of spreading is a proper subgroup of G_i. This subgroup will be our G_{i+1}, and we set $e_{i+1} = e_i M^J$. We obtain

$$\left|\alpha^{e_{i+1}} \cap G_{i+1}\right| \geq \left|\alpha^{e_i M^{j_0+1}} \cap G_{i+1}\right| \geq \left|\alpha^{e_i M^{j_0}} \cap G_i\right|^{\frac{\dim(G_{i+1})}{\dim(G_i)}} \geq$$

$$\geq \left|\alpha^{e_i} \cap G_i\right|^{\frac{\dim(G_{i+1})}{\dim(G_i)}} \geq \left(\tfrac{|\alpha|}{d}\right)^{\frac{\dim(G_i)}{n^2}\,\frac{\dim(G_{i+1})}{\dim(G_i)}} \geq \left(\tfrac{|\alpha|}{d}\right)^{\frac{\dim(G_{i+1})}{n^2}}\ ,$$

the induction step is complete. The dimensions $\dim(G_i)$ strictly decrease as i grows, hence the induction must stop in at most n^2 steps. But the only way it can stop is to produce the required subgroup H.

Iterating the previous lemma we obtain that a non-growing subset $\alpha \subset GL(n, \overline{\mathbb{F}})$ is covered by a few cosets of a virtually soluble group. In the

proof we need an auxiliary subgroup G in order to do induction on $\dim(G)$. For applications the only interesting case is $G = GL(n, \overline{\mathbb{F}})$, $\deg(G) = 1$.

Corollary 2.13.4. *Let $G \leq GL(n, \overline{\mathbb{F}})$ be a (possibly non-connected) closed subgroup and $\alpha \subseteq G$ a finite subset. Suppose that $|\alpha^3| \leq \mathcal{K}|\alpha|$ for some $\mathcal{K}$. Then there is a virtually soluble normal subgroup $\Delta \lhd \langle\alpha\rangle$ and a bound $m = m\big(n, \deg(G)\big)$ such that the subset α can be covered by $\mathcal{K}^m$ cosets of Δ.*

Proof During the proof we encounter several lower bounds for m, we assume that our m satisfies them all. We prove the corollary by induction on $N = \dim(G)$. If $\mathcal{K} < \frac{3}{2}$ then $\langle\alpha\rangle$ is virtually cyclic by Proposition 2.13.2 and the lemma holds with $\Delta = \langle\alpha\rangle$. If $|\alpha| \leq \mathcal{K}^m$ then our statement holds with $\Delta = \{1\}$. So we assume that $\mathcal{K} \geq \frac{3}{2}$ and $|\alpha| > \mathcal{K}^m$. If $\dim(G) = 0$ then $|\alpha| \leq \deg(G)$, we exclude this case by choosing m large enough.

Suppose that $m \geq m_{\mathrm{nilp}}\big(n, \deg(G)\big)$. Applying Proposition 2.13.3 we obtain a subgroup H normalised by α such that $\langle\alpha\rangle \cap H$ is virtually nilpotent, $\dim(H) \geq 1$, and $\deg(H)$ is bounded in terms of n and $\deg(G)$. Consider the algebraic group $\overline{G} = \mathcal{N}_G(H)/H$, let $\overline{\alpha} \subseteq \overline{G}$ denote the image of α. By Proposition 2.12.19 we have $|\overline{\alpha}^3| \leq \mathcal{K}^2|\overline{\alpha}|$. By Proposition 2.12.8 and Fact 2.3.10.(f) there is an embedding $\overline{G} \leq GL(n', \overline{\mathbb{F}})$ where n' and $\deg(\overline{G})$ are bounded in terms of n, $\deg(G)$ and $\deg(H)$. Clearly $\dim(\overline{G}) < \dim(G)$, so by the induction hypothesis we obtain a virtually soluble normal subgroup $\overline{\Delta} \lhd \langle\overline{\alpha}\rangle$ such that $\overline{\alpha}$ is covered by $\mathcal{K}^{2m(n', \deg(\overline{G}))}$ cosets of $\overline{\Delta}$. We define Δ to be the preimage of $\overline{\Delta}$ in $\langle\alpha\rangle$. Then Δ is virtually soluble since the class of virtually soluble groups is closed under extensions (see e.g. [**88**]). The induction step is complete.

The following consequence of well-known results is of independent interest.

Lemma 2.13.5. *Let Δ be a virtually soluble subgroup of $GL(n, \overline{\mathbb{F}})$ and let S be the soluble radical of Δ. Then Δ has a characteristic subgroup $\Delta_0 \geq S$ such that Δ_0/S is a direct product of simple groups of Lie type of the same characteristic as $\overline{\mathbb{F}}$ and $|\Delta/\Delta_0| \leq f(n)$ (where $f(n)$ is as in Proposition 2.12.16). Moreover the Lie rank of the simple factors appearing in Δ_0/S is bounded by n and the number of simple factors is also at most n.*

Proof If $\mathrm{char}(\overline{\mathbb{F}}) = 0$ this is a theorem of Platonov (see [**166**]). Assume $\mathrm{char}(\overline{\mathbb{F}}) = p > 0$. Let D be the Zariski closure of Δ. Then D^0 is soluble (see [**166**, Theorem 5.11]) and $(D^0)\Delta = D$ hence $\tilde{\Delta} = \Delta/\big(\Delta \cap D^0\big) \cong D/D^0$. By a result of Platonov (see [**166**, Lemma 10.10]) we have $D = (D^0)G$ where G is some finite subgroup of D, hence $G/\big(G \cap D^0\big) \cong D/D^0$. Now $\tilde{\Delta}$ is isomorphic to a quotient of the finite group $G \leq GL(n, \overline{\mathbb{F}})$ by a soluble normal subgroup. Therefore Proposition 2.12.16 implies that $\tilde{\Delta}$ has a characteristic subgroup H of index at most $f(n)$ such that $H/Sol(\tilde{\Delta})$ is in $Lie^*(p)$ (we can take $H/Sol(\tilde{\Delta})$ to be the $Lie^*(p)$ part of the socle of $\tilde{\Delta}/Sol(\tilde{\Delta})$). Using [**37**, Theorem 3.4B] it follows that $H/Sol(\tilde{\Delta})$ is isomorphic to a quotient of a finite subgroup of $GL(n, \overline{\mathbb{F}_p})$. As in the proof of Proposition 2.12.20 we see that the number of simple factors in $H/Sol(\tilde{\Delta})$ and their Lie ranks are bounded by n. Let Δ_0 be the subgroup of Δ which corresponds to

H. This is a characteristic subgroup since the kernel of the homomorphism $\Delta \to \left(\tilde{\Delta}/Sol(\tilde{\Delta})\right)$ is $Sol(\Delta)$, which is characteristic in Δ. We obtain our statement.

Combining Corollary 2.13.4 and Lemma 2.13.5 we see that a non-growing subset $\alpha \subset GL(n, \overline{\mathbb{F}})$ is covered by a few cosets of a soluble by $Lie^*(p)$ normal subgroup of $\langle \alpha \rangle$. To obtain another such subgroup Γ for which $\alpha^6 Sol(\Gamma)$ contains Γ we need a bit more work. The following two lemmas taken together describe the structure of a (possibly infinite) soluble by $Lie^*(p)$ linear group.

Lemma 2.13.6. *Let $S \leq GL(n, \overline{\mathbb{F}})$ be a soluble subgroup normalised by a subset $\alpha \subseteq GL(n, \overline{\mathbb{F}})$. Then there is a closed subgroup $D \leq GL(n, \overline{\mathbb{F}})$ containing α and S, and a homomorphism $\phi : D \to GL(n', \overline{\mathbb{F}})$ such that $\ker(\phi)$ is soluble, contains S, and n' depends only on n.*

Proof If S is abelian then we consider the centralisers $A = \mathcal{C}_{GL(n,\overline{\mathbb{F}})}(S)$ and $B = \mathcal{C}_{GL(n,\overline{\mathbb{F}})}(A)$. By [**166**, Theorem 6.2] we have homomorphisms

$$\phi_1 : \mathcal{N}_{GL(n,\overline{\mathbb{F}})}(A) \to GL(n^2, \overline{\mathbb{F}}) , \quad \phi_2 : \mathcal{N}_{GL(n,\overline{\mathbb{F}})}(B) \to GL(n^2, \overline{\mathbb{F}})$$

whose kernels are precisely A and B. Note that $A \cap B = \mathcal{Z}(A)$ contains S. Since α normalises S, it also normalises A and B. The lemma holds in this case with the following settings:

$$D = \mathcal{N}_{GL(n,\overline{\mathbb{F}})}(A) \cap \mathcal{N}_{GL(n,\overline{\mathbb{F}})}(B) ,$$

$$\phi = (\phi_1, \phi_2) : D \longrightarrow GL(n^2, \overline{\mathbb{F}}) \times GL(n^2, \overline{\mathbb{F}}) \ \leq \ GL(2n^2, \overline{\mathbb{F}}) .$$

In the general case we do induction on the derived length of S, which is bounded in terms of n [**166**, Theorem 3.7]. The commutator subgroup S^* is normalised by the subset $\alpha^* = \alpha \cup S$, we apply to them the induction hypothesis. We obtain a closed subgroup $D^* \leq GL(n, \overline{\mathbb{F}})$ containing $\alpha \cup S$ and a homomorphism $\phi^* : D^* \to GL(m^*, \overline{\mathbb{F}})$ such that $\ker(\phi^*)$ is soluble, contains S^*, and m^* depends only on n. The image $\phi^*(S)$ is abelian and it is normalised by $\phi^*(\alpha)$. By the above settled case there is a closed subgroup $D^{**} \leq GL(m^*, \overline{\mathbb{F}})$ containing $\phi^*(\alpha)$ and a homomorphism $\phi^{**} : D^{**} \to GL(m^{**}, \overline{\mathbb{F}})$ such that $\ker(\phi^{**})$ is soluble, contains $\phi^*(S)$, and m^{**} depends only on m^*, hence only on n. We set

$$D = \phi^{*-1}(D^{**}) , \quad \phi = \phi^{**} \circ \phi^* , \quad m = m^{**} ,$$

the induction step is complete.

Lemma 2.13.7. *Let Λ be a subgroup of $GL(n, \overline{\mathbb{F}})$, $\mathrm{char}(\overline{\mathbb{F}}) = p$ and L a finite normal subgroup of Λ such that L is in $Lie^*(p)$. Then $\Lambda/L\mathcal{C}_\Lambda(L) \leq f(n^2)$ where $f()$ is as in Proposition 2.12.16.*

Proof By [**166**, Theorem 6.2] $\Lambda/\mathcal{C}_\Lambda(L)$ is a subgroup of $GL(n^2, \overline{\mathbb{F}})$ hence by Proposition 2.12.16 it has a soluble by $Lie^*(p)$ normal subgroup N of index at most $f(n^2)$. On the other hand $\Lambda/\mathcal{C}_\Lambda(L)$ is isomorphic to a subgroup A of $Aut(L)$ containing $Inn(L) \cong L$. It is easy to see that the socle of A is $Inn(L)$. Therefore all soluble by $Lie^*(p)$ normal subgroups of A are actually $Lie^*(p)$ subgroups of $Inn(L)$. Our statement follows.

We need two more auxiliary results on $Lie^*(p)$ groups.

Lemma 2.13.8. *Let H be a normal subgroup of a group G and assume that H is a direct product of at most m finite simple groups of Lie type of rank at most m. Let α be a symmetric subset of G covered by x cosets of H. If $|\alpha| \geq |H|/y$ then H has a (possibly trivial) characteristic subgroup N such that N is contained in α^6 and $|H/N| \leq (xy)^{Cm^2}$ for some constant C.*

Proof If L is a simple direct factor of H and $k = k(L)$ is the minimal degree of a non-trivial complex representation of L then by Proposition 2.12.4 we have $|L| < k^{\frac{C}{3}m}$ for some absolute constant C. Let $k_0 < k_1 < ....$ be the different numbers $k(L)$. Define H_i as the product of the direct factors L for which $k(L) \geq k_i$. The H_i are characteristic subgroups of H. By our assumptions for all indices i we have $|\alpha^2 \cap H_i| \geq |\alpha|/x|H/H_i| \geq |H_i|/xy$. By Proposition 2.12.3 if $|\alpha^2 \cap H_i| > |H_i|/(k_i)^{1/3}$ then we have $H_i \subseteq \alpha^6$. Let j be the smallest index for which this holds. By the above for all $i < j$ we have $k_i \leq (xy)^3$ hence if L is a simple constituent of $|H/H_j|$ then $|L| < (xy)^{Cm}$. Setting $N = H_j$ we obtain that $|H/N| \leq (xy)^{Cm^2}$, as required.

Lemma 2.13.9. *Let $L = L_1 \times \cdots \times L_m$ be a direct product of simple groups of Lie type of rank at most r. Let α be a symmetric generating set of L which projects onto all simple quotients of L. Then $\alpha^{c(m,r)} = L$ where $c(m, r)$ depends only on m and r.*

Proof We need the following

Claim 2.13.10. *Let $x = (x_1, \ldots x_t)$ be an element of a product $L_1 \times \cdots \times L_t$ of simple groups of Lie type of rank at most r such that all x_i are non-trivial. Then each element of $L_1 \times \cdots \times L_t$ is a product of at most Cr conjugates of x for an absolute constant C.*

For $t = 1$ this is proved in [**96**] and the general case is an obvious consequence.

We prove the lemma by induction on m. It is clear that α^2 has two elements whose first projections are the same, hence α^3 contains a non-trivial element $a = (a_1, \ldots, a_m)$ such that $a_1 = 1$. Assume that $a_{i+1}, \ldots, a_m$ are the projections of a different from 1. By the induction hypothesis we know that $\beta = \alpha^{c(m-1,r)}$ projects onto the quotient L/L_1. By the claim each element of $L_{i+1} \times \cdots \times L_m$ is a product of at most Cr conjugates of a by elements of β, hence this subgroup is contained in $(\alpha^3\beta^2)^{Cr}$. Using again the induction hypothesis we see that β projects onto $L_1 \times \cdots \times L_{m-1}$ hence $L \leq \beta L_m \leq (\alpha^3\beta^3)^{Cr}$. We obtain that $L \leq \alpha^{3Cr(c(m-1,r)+1)}$ which completes the induction step.

Finally we are ready to prove Theorem 2.1.10.

Theorem 2.13.11. *Let $\alpha \subseteq GL(n, \overline{\mathbb{F}})$ be a finite symmetric subset such that $\left|\alpha^3\right| \leq \mathcal{K}|\alpha|$ for some $\mathcal{K} \geq \frac{3}{2}$. Then there are normal subgroups $S \leq \Gamma$ of $\langle\alpha\rangle$ and a bound m depending only on n such that $\Gamma \subseteq \alpha^6 S$, the subset α can be covered by $\mathcal{K}^m$ cosets of Γ, S is soluble, and the quotient group Γ/S is the product of finite simple groups of Lie type of the same characteristic as $\overline{\mathbb{F}}$. (In particular, in characteristic 0 we have $\Gamma = S$.) Moreover, the the Lie rank of the simple factors appearing in Γ/S is bounded by n, and the number of factors is also at most n.*

Proof If $\mathrm{char}(\overline{\mathbb{F}}) = 0$ then our statement follows from Corollary 2.13.4 and Lemma 2.13.5. Assume that $\mathrm{char}(\overline{\mathbb{F}}) = p > 0$. Corollary 2.13.4 and Lemma 2.13.5 imply that $\Lambda = \langle \alpha \rangle$ has a normal subgroup Δ such that $\Delta/Sol(\Delta)$ is in $Lie^*(p)$ and α is covered by $K^{a(n)}$ cosets of Δ where $a(n)$ depends on n. Moreover $\Delta/Sol(\Delta)$ is the direct product $L_1 \times \cdots \times L_t$ of at most n simple groups of Lie type of rank at most n. We set $S = Sol(\Delta)$. The proof of our theorem reduces to the following.

Claim 2.13.12. *The group Λ has a normal subgroup Γ such that $\Delta \geq \Gamma \geq S$, $S\alpha^6 \geq \Gamma$ and α is covered by K^m cosets of Γ.*

To prove the claim, by Lemma 2.13.6 and Proposition 2.12.19 we might as well assume (at the cost of enlarging n and K) that $S = \{1\}$, i.e. $\Delta = L_1 \times \cdots \times L_t$. In this case Proposition 2.13.7 implies that Λ has a normal subgroup H of index at most $f(n^2)$ such that H is the direct product of Δ and $C = \mathcal{C}_\Lambda(\Delta)$. Set $\gamma = \alpha^{2f(n^2)} \cap H$. Slightly adjusting the proof of Proposition 2.12.22.(a) we see that γ generates H and $|\gamma^3| \leq K_0|\gamma|$ where $K_0 = K^{7f(n^2)}$.

Denote by N_j the (unique) direct complement of L_j in H. Using Theorem 2.12.5 (as in the proof of Proposition 2.12.20) we see that for the quotients $H/N_j \cong L_j$ we have two possibilities; either γ^3 projects onto H/N_j (in which case $|\gamma N_j/N_j| \geq |L_j|/K_0^2$ by Proposition 2.12.19) or $|\gamma N_j/N_j| \leq K_0^{b(n)}$ where $b(n)$ depends only on n. Let $H/N_1, \ldots, H/N_i$ be the quotients for which the first possibility holds and which also satisfy $|L_j| > K_0^{b(n)+4}$.

Since H/C is a direct product of nonabelian simple groups it follows that conjugation by α permutes the simple factors, therefore it permutes the subgroups N_j. By an argument as in the proof of Proposition 2.12.22.(b) we see that the set $\{N_1, \ldots, N_i\}$ is invariant under conjugation by α. Therefore $N = N_1 \cap \cdots \cap N_i$ and $I = N_{i+1} \cap \cdots \cap N_t$ are normal subgroups of Λ. By our assumptions γ^3 projects onto all simple quotients of H/N and $(\gamma N)/N$ generates this group. By Lemma 2.13.9 we see that $\gamma^{c(n)}$ projects onto H/N where $c(n)$ depends on n. This implies $|\alpha|K^{d(n)} \geq |H/N|$ where $d(n) = 2f(n^2)c(n)$.

The subgroup $D = I \cap \Delta = L_1 \times \cdots \times L_i$ is also normal in Λ and we have $H/N \cong D$, hence $|\alpha|K^{d(n)} \geq |D|$. By our assumptions γ projects onto at most $K_0^{n(b(n)+4)} = K^{e(n)}$ elements of H/I. Since $\alpha^2 \cap H \subseteq \gamma$, the natural isomorphism between H/I and Δ/D implies that $\alpha^2 \cap \Delta$ projects onto at most $K^{e(n)}$ elements of Δ/D. Using Proposition 2.12.21 we see that α is covered by $K^{a(n)+e(n)}$ cosets of D. Since $|\alpha| \geq |D|/K^{d(n)}$, Lemma 2.13.8 implies that D has a characteristic subgroup Γ contained in α^6 such that $|D/\Gamma| \leq K^{(a(n)+d(n)+e(n))Cn^2}$. The subgroup Γ is normal in Λ and α is covered by $|D/\Gamma|K^{a(n)+e(n)}$ cosets of Γ. Our statement follows.

Theorem 2.1.10 does not hold for all $\mathcal{K} \geq 1$. For example α could be a subgroup of $GL(n, \overline{\mathbb{F}})$ isomorphic to $Alt(n)$. However the structure of subsets α with $|\alpha^3| < \frac{3}{2}|\alpha|$ is completely described in Proposition 2.13.2.

2.14. Examples

In this section we give some examples which show that the constant $\varepsilon(r)$ for which Theorem 2.12.5 holds must be less than $\frac{C}{r}$. It will be convenient to rely on [**5**, Section 3] in describing our examples.

Example 2.14.1. Consider the group $SL(n,q)$ where $n \geq 3$ (which has Lie rank $r = n-1$). Let H be the subgroup of all diagonal matrices, this has order $(q-1)^{n-1}$. If N denotes the subgroup of all monomial matrices then $N/H \simeq S_n$ Choose an element s of N projecting onto an n-cycle of N/H. If $e_1, \ldots, e_n$ is the standard basis of ${\mathbb{F}_q}^n$, consider the subgroup $L_{1,2} \simeq SL(2,q)$ which fixes $e_3, \ldots, e_n$. In [**5**, Theorem 3.1] a 3-element generating set $\{a,b,c\}$ of $L_{1,2}$ is chosen. As shown in [**5**] s, a, b and c generate $SL(n,q)$ (moreover, the diameter of the corresponding Cayley graph is logarithmic).

Now s normalises the diagonal subgroup H and it is clear that a, b and c normalise a subgroup H_0 of index $(q-1)^2$ in H (the group of diagonal matrices fixing e_1 and e_2). Our generating A set will consist of H, a, b, c and s. We claim that

$$\left|A^3\right| \leq |H|\big(3(q-1)^2 + 58\big) + 64\ .$$

It is straightforward to see that

$$\left|A^3\right| \leq \Big|H\{a,b,c,s\}H\Big| + 57|H| + 64\ .$$

Since s normalises H we have $\left|HsH\right| = |H|$. Since a (resp. b and c) normalises H_0 we have $\left|HaH\right| \leq |H|(q-1)^2$ (and analogous inequalities hold for b and c) which implies the claim.

Setting $q = 3$ we obtain the generating set with $\left|A^3\right| \leq 100|A|$ mentioned in the introduction.

Clearly, there are many ways in which the above construction can be extended. For example the full diagonal subgroup H can be replaced by its characteristic subgroups isomorphic to C_t^{n-1} where t divides $q-1$. This way e.g. we can construct large families of generating sets of constant growth whenever q is odd.

It would be most interesting to find some essentially different families of examples of large generating sets of $SL(n,q)$ with constant growth.

The above generating sets of "moderate growth" are "dense" subsets of the union of a few cosets of some subgroup. This can be avoided. Assume that $q = 2^p$ where $p \geq n$ is an odd prime. It is well-known that all divisors of $q-1$ are greater that $2p+1$. Replace H in the above construction by a subset $P \subseteq H$ of the form $\prod^{n-1}\{g, g^2, \ldots, g^n\} \subseteq \prod^{n-1} C_{q-1} \simeq H$ which is invariant under conjugation by the cyclic element s. Now $A = P \cup \{a,b,c,s\}$ is a generating set of size roughly n^{n-1} with A^3 of size roughly n^n. It is easy to see that P is far from being a subgroup of $SL(n,q)$.

2.15. Appendix to Chapter 2

In this appendix we prove rigorously the algebraic geometry facts used in Chapter 2. For reference we use [**71**, Sections I.1, I.2, I.7 and II.3], and also [**91**, Section I.3]. Besides that, we need Proposition 2.15.1, which is a version of Bézout's theorem, stated and proved in [**57**].

Let $\overline{\mathbb{F}}^m$ denote the m-dimensional affine space over the algebraically closed field $\overline{\mathbb{F}}$, and $\mathbb{P}^m$ denote its projective closure. For a locally closed subset $X \subseteq \overline{\mathbb{F}}^m$, in this appendix $\overline{X}$ denotes (as before) the closure of X in $\overline{\mathbb{F}}^m$, and $\overline{X}^{\mathbb{P}^m}$ denotes the closure of X in $\mathbb{P}^m$. Similarily, $\deg(X)$ and $\deg(\overline{X})$ denotes the degrees in the sense of Definition 2.3.5, and $\deg_{\mathbb{P}^m}(\overline{X}^{\mathbb{P}^m})$ denotes the degree of the projective variety $\overline{X}^{\mathbb{P}^m} \subseteq \mathbb{P}^m$ in the sense of [**71**, Section I.7]. Note, that both notions of degree depend not only on the isomorphism type of X, but also on the particular embedding of X into the affine (or projective) space.

Proposition 2.15.1 (Fulton, see [**57**], this is a variant of Bézout's theorem)**.** *Let P, Q be irreducible closed subsets of the projective space $\mathbb{P}^m$, and let $Z_1, \dots, Z_k$ be the irreducible components of $P \cap Q$. Then*

$$\deg_{\mathbb{P}^m}(P) \cdot \deg_{\mathbb{P}^m}(Q) \geq \sum_{i=1}^{k} \deg_{\mathbb{P}^m}(Z_i) \ .$$

Definitions 2.3.1, 2.3.2, 2.3.3 and 2.3.4 are standard, we do not comment on them. On the other hand, the degree is usually defined for projective varieties, and in Definition 2.3.5 we deal with locally closed subsets of $\overline{\mathbb{F}}^m$. The connection with the usual notions is explained by the following:

Proposition 2.15.2. *For a locally closed subset $X \subseteq \overline{\mathbb{F}}^m$ we have*

$$\dim(X) = \dim(\overline{X}) = \dim(\overline{X}^{\mathbb{P}^m}) \ ,$$

$$\deg(X) = \deg(\overline{X}) = \deg_{\mathbb{P}^m}(\overline{X}^{\mathbb{P}^m}) \ .$$

Moreover, X is irreducible iff $\overline{X}^{\mathbb{P}^m}$ is irreducible.

Proof The last statement follows from [**71**, Ex.I.1.6]. Then it is enough to prove the two equalities for irreducible X. So we assume that X is irreducible. The equality of dimensions is [**71**, Ex.I.2.7]. Let $\mathcal{L}$ denote the collection of affine subspaces $L \subseteq \overline{\mathbb{F}}^m$ of dimension $m - \dim(X)$. For all members $L \in \mathcal{L}$, the intersection $\overline{L}^{\mathbb{P}^m} \cap \overline{X}^{\mathbb{P}^m}$ is either infinite, or it has at most $\deg_{\mathbb{P}^m}(\overline{X}^{\mathbb{P}^m})$ points. Moreover, for almost all L the intersection $\overline{L}^{\mathbb{P}^m} \cap \overline{X^{\mathbb{P}^m}}$ have exactly $\deg_{\mathbb{P}^m}(\overline{X}^{\mathbb{P}^m})$ points and $\overline{L}^{\mathbb{P}^m}$ avoids the smaller dimensional boundary $\overline{X}^{\mathbb{P}^m} \setminus X$. This proves that $\deg(X) = \deg_{\mathbb{P}^m}(\overline{X^{\mathbb{P}^m}})$. The same argument applied to $\overline{X}$ completes the proof.

Remark 2.3.6 follows immediately from our definition of $\deg(X)$, as a single point has degree 1. Definition 2.3.7 and Remark 2.3.8 are standard, we do not comment on them.

Proof [Proof of Fact 2.3.9] (a), (b) and (c) follows from Proposition 2.15.2 and the analogous statements for projective varieties. (e) follows from [**71**, Ex.I.1.6] and from the definition of the dimension.

Combining Proposition 2.15.2 with $\overline{X \cup Y}^{\mathbb{P}^m} = \overline{X}^{\mathbb{P}^m} \cup \overline{Y}^{\mathbb{P}^m}$, $\overline{X \cap Y}^{\mathbb{P}^m} \subseteq \overline{X}^{\mathbb{P}^m} \cap \overline{Y}^{\mathbb{P}^m}$, $X \setminus \overline{Y} \subseteq \overline{X}^{\mathbb{P}^m} \setminus \overline{Y}^{\mathbb{P}^m}$ we obtain most of (d), with the exception of its last equality. Next we consider the intersection

$$\left(X \times \overline{\mathbb{F}}^m\right) \cap \left(\overline{\mathbb{F}}^m \times Y\right) = X \times Y \subseteq \overline{\mathbb{F}}^{2m} \ .$$

Taking closures in $\mathbb{P}^{2m}$ and applying [**71**, Theorem I.7.7] we obtain the last equality of (d).

If X and Y are irreducible then $\overline{X} \times \overline{Y} = \overline{X \times Y}$ is irreducible by [**71**, Ex.I.3.15(d)], hence (f) follows from [**71**, Ex.I.1.6].

Next we introduce two invariants of closed subsets. If $Z \subseteq \overline{\mathbb{F}}^m$ is a closed set with irreducible decomposition $Z = \bigcup_i Z_i$ then we define

$$N(Z) = \sum_i (d+1)^{\dim(Z_i)} \deg(Z_i) \quad \text{and} \quad D(Z) = \sum_i d^{\dim(Z_i)} \deg(Z_i) \, .$$

Let F be the zero set of a polynomial of degree d which does not vanish identically on Z. By Proposition 2.15.1 we have $N(Z_i \cap F) < N(Z_i)$ and $D(Z_i \cap F) \leq D(Z_i)$ whenever $Z_i \subsetneq F$, therefore $N(Z \cap F) < N(Z)$ and $D(Z \cap F) \leq D(Z)$. To obtain X we start from $\overline{\mathbb{F}}^m$, and add the equations of X of degree d one by one, until their common zero locus becomes X. We obtain that $\deg(X) \leq D(X) \leq D(\overline{\mathbb{F}}^m) = d^m$, and the invariant N decreases in each step, i.e. we need at most $N(\overline{\mathbb{F}}^m) = (d+1)^m$ equations. One direction of (g) is proved. The other direction of (g) follows from [**91**, Section I.3] (the construction of the Chow variety).

Proof [Proof of Fact 2.3.10] Let $X \subseteq \overline{\mathbb{F}}^n$ and $Y \subseteq \overline{\mathbb{F}}^m$ denote the ambient spaces (see the note after Definition 2.3.1), and let $\pi : \overline{\mathbb{F}}^n \times \overline{\mathbb{F}}^m \to \overline{\mathbb{F}}^m$ denote the projection to the second factor. Note that Γ_f is isomorphic to X (via the first projection), and $f(X) = \pi(\Gamma_f)$.

We already proved (a) with the exception of the degree estimates which we postpone for a while.

In the proof of (b) we may (and do) assume that X is irreducible. If $\overline{f(X)} = A \cup B$ were a proper decomposition into closed subsets then $X = f^{-1}(A) \cup f^{-1}(B)$ would also be a proper decomposition, a contradiction. Hence $\overline{f(X)}$ is also irreducible. By [**71**, Ex.II.3.19(b)] the subset $f(X)$ contains a dense open subset $U \subseteq \overline{f(X)}$. It remains to estimate $\deg(f)$. Let $L \subseteq \overline{\mathbb{F}}^m$ be an affine subspace of dimension $m - \dim\big(\overline{f(X)}\big)$ which intersects U in exactly $\deg(U) = \deg\big(\overline{f(X)}\big)$ points (see Definition 2.3.5 and Fact 2.3.9.(a)). Then $\pi^{-1}(L)$ is an affine subspace, hence $\deg(f) = \deg(\Gamma_f) \geq \deg\big(\Gamma_f \cap \pi^{-1}(L)\big)$. But $\Gamma_f \cap \pi^{-1}(L)$ is isomorphic to $f^{-1}\big(L\big) = f^{-1}\big(U \cap L\big)$, hence it has at least $\deg\big(\overline{f(X)}\big)$ connected components. This implies that $\deg(f) \geq \deg\big(\overline{f(X)}\big)$, (b) is proved.

Next we prove (c). We know that $f^{-1}(T)$ is isomorphis to $\Gamma_f \cap \pi^{-1}(T)$, and $\pi^{-1}(T) = \overline{\mathbb{F}}^n \times T$ have degree $\deg(T)$ by Fact 2.3.9.(d). Then $\deg\big(f^{-1}(T)\big) \leq \deg(T)\deg(f)$ by Fact 2.3.9.(d). In the spacial case $T = \{y\}$ we obtain $\deg\big(f^{-1}(y)\big) \leq \deg(f)$, which completes the proof of (c).

The closed complement considered in (d) is the union of a number of the locally closed subsets of (a), hence its degree bound follows immediately from (a). So (d) is proved.

[**71**, Ex.II.3.22(b)] contains the inequality of (e) as well as the openness and denseness of $X_{\min}$. The difference $X \setminus X_{\min}$ is the inverse image of the union of a number of the locally closed subsets of (a), hence its degree bound follows from (a) and (c). This proves (e).

In (f), the graph of the restricted morphism $f\big|_S$ is $\Gamma_f \cap (S \times \overline{\mathbb{F}}^m)$. By Fact 2.3.9.(d) it has degree at most $\deg(\Gamma_f)\deg\left(S \times \overline{\mathbb{F}}^m\right) = \deg(f)\deg(S)$. Moreover, if S is an irreducible component of X then the graph of $f\big|_S$ is the corresponding component of Γ_f. This proves (f).

Proof [Proof of Fact 2.3.10.(a), counting the sheep] First we bound the number of the parts in the partitions of Z. In the proof we partition Z'_j in at most $d+1$ steps. In the very first step we subdivide Z'_j into $(d+1)^2+2$ parts, and the algorithm stops in two of them. Suppose that C is a partition class constructed before the $(l-1)$-th polynomial division and the algorithm did not stop in C. Before the l-th division we subdivide C into $d+2$ parts, in one of them the algorithm stops, in the other $d+1$ it continues. Altogether we cut Z'_j into at most $2+\sum_{l=1}^{d}(d+1)^{l+1} \leq (d+2)^{d+1}$ pieces, and we repeat this cutting less than $(d+1)^{k+1}$ times. Hence we obtain altogether at most $(d+2)^{(d+1)((d+1)^{k+1}-1)}$ parts Z_i. Finally we cut each Z_i again into at most $d+2$ parts.

Let $p(t, \underline{x})$ and $q(t, \underline{x})$ be polynomials of t-degree at most d and $\underline{x}$-degree at most e. We divide by the leading t-coefficients, then all t-coefficients are rational functions of degree at most (with nonstandard notation) e/e. We do polynomial division: both the quotient and the remainder have coefficients of degree at most e^2/e^2. We run Euclid's algorithm for p and q. We do at most d divisions. In each quotient and in each remainder the t-coefficients have degrees at most e^{2^d}/e^{2^d}. Then we multiply through with the denominators.

In the proof of Claim 2.3.11.(a) we run Euclid's algorithm at most $(d+1)^{k+1}-1$ times. So each polynomial we encounter (including the P_i) has t-degree at most d and $\underline{x}$-degree at most $d^{((d+1)^{k+1}-1)2^d}$, hence their total degree is at most $d^{(d+1)^{k+1}2^d}$. In the proof of Claim 2.3.11.(b) each Z_i is subdivided into at most $d+2$ locally closed subsets defined via the vanishing or non-vanishing of several k-variate polynomials of degree at most $d^{(d+1)^{k+1}2^d}$.

In the proof of Fact 2.3.10.(a) we start from $\overline{f(X)}$ (which has degree at most $\deg(f)$), and apply Claim 2.3.11 at most $\dim(X) + \deg(f) - 1$ times. Each time we subdivide each locally closed subset into at most $\Phi\big(\Phi\big(\dots\Phi\big(\deg(f)\big)\big)\dots\big)$ pieces and each piece is defined with equations of degree at most $\Phi\big(\Phi\big(\dots\Phi\big(\deg(f)\big)\big)\dots\big)$. At the end we obtain altogether at most D locally closed parts and their degrees are at most D (see Fact 2.3.9.(g)).

Finally, in Fact 2.3.10.(d) the subset in question is the union of a number of the locally closed subsets of (a), and the subset in Fact 2.3.10.(e) is the inverse image of such a union. Hence their degrees are at most D^2 and $D^2\deg(f)$ respectively.

CHAPTER 3

Helfgott's conjecture, soluble version

3.1. Introduction

For the convenience of the reader, we restate here Theorem 2.1.4 from Chapter 2. The result was proved simultaneously and independently by Breuillard–Green–Tao [**25**] and Pyber–Szabó [**133**] in 2010.

Theorem 3.1.1 (Product theorem)**.** *Let L be a finite simple group of Lie type of rank r and A a generating set of L. Then either $A^3 = L$ or*

$$|A^3| \gg |A|^{1+\varepsilon}$$

where ε and the implied constant depend only on r.

For $G = PSL(2,p)$, p prime, this is a famous result of Helfgott [**72**]. For $PSL(3,p)$ resp. $PSL(2,q)$, q a prime-power, this was proved earlier by Helfgott [**73**] resp. Dinai [**36**] and Varjú [**165**].

For the groups $G = PSL(n,q)$ (which are simple groups of Lie type of rank $n-1$) the Product theorem can be reformulated as follows: If A is a generating set of G, such that $|A^3| < K|A|$ for some number $K \geq 1$, then A is contained in $K^{c(n)}$ (i.e. polynomially many) cosets of some normal subgroup H of G contained in A^3. This (somewhat artificial) reformulation turns out to be quite useful when we seek an extension of the Product theorem which describes non-growing subsets of linear groups.

The Product theorem has quickly become a central result of finite asymptotic group theory with many applications. Let us briefly mention a few of them. The Product theorem easily implies, that the Babai conjecture[1] [**7**] holds for finite simple groups of Lie type of bounded rank. With Gill, Pyber, Short [**61**] (see Theorem 7.1.3) we proved that the Conjecture of Liebeck, Nikolov, Shalev[2] [**97**] holds for simple groups of Lie type of bounded rank. The proof is based on the Product theorem, and a deep result of Liebeck and Shalev [**104**] and an extra trick which handles small subsets. Breuillard, Green, Guralnick and Tao [**22**] proved that a large number of Cayley graphs are expanders.[3] Their proof is based on the Product theorem, and the

[1] Babai conjectures, that for every non-abelian finite simple group L and every symmetric generating set S of L the diameter of the Cayley graph of L corresponding to S is at most $C\big(\log|L|\big)^c$ where c and C are absolute constants.

[2] Conjecturally there exists an absolute constant c such that if L is a finite simple group and S is a subset of L of size at least two, then L is a product of N conjugates of S for some $N \leq c\log|L|/\log|S|$.

[3] Let G be a finite simple group of Lie type of rank r. They proved that the Cayley graph of G corresponding to two random elements is an $\varepsilon(r)$ expander with probability going to 1 as $|G| \to \infty$.

so called "Bourgain-Gamburd expansion machine" developed by Bourgain–Gamburd [**12**].

For the above applications of the Product theorem it is essential that the size of a generating set A is bounded by a polynomial of the tripling constant $K = |A^3|/|A|$ (unless A is very large). For a discussion of related issues by Tao see [**155**]. Guided by this insight, and remarks of Helfgott [**73**, page 764], and Breuillard, Green, Tao [**27**, Remark 1.8], and various discussions on Tao's blog, we started to study the structure of non-growing subsets in linear groups. Our main result in Chapter 3 is the following.

Theorem 3.1.2 (Polynomial Inverse theorem)**.** *Let S be a symmetric subset of $GL(n, \mathbb{F})$ satisfying $|S^3| \leq K|S|$ for some $K \geq 1$, where $\mathbb{F}$ is an arbitrary field. Then S is contained in the union of polynomially many (more precisely $K^{c(n)}$) cosets of a finite-by-soluble subgroup Γ normalised by S.*

Moreover, Γ has a finite subgroup P normalised by S such that Γ/P is soluble, and S^3 contains a coset of P.

The theorem extends and unifies several earlier results. Most importantly (to us) it contains the Product theorem (for symmetric sets) as a special case. For subsets of $SL(2, p)$ and $SL(3, p)$ similar results were obtained by Helfgott [**72, 73**].

In characteristic 0 the above theorem was first proved by Breuillard, Green and Tao [**25**]. The earliest result in this direction for subsets of $SL(2, \mathbb{R})$ is due to Elekes and Király [**43**]. The proof in [**25**] uses the fact that a virtually soluble subgroup of $SL(n, \mathbb{C})$ has a soluble subgroup of n-bounded index, which is no longer true in positive characteristic.

Finally the theorem implies a result of Hrushovski for linear groups over arbitrary fields obtained by model-theoretic tools [**77**]. In Hrushovski's theorem the structure of Γ is described in a less precise way and the number of covering cosets is only bounded by some large function of n and K.

It seems possible that a result similar to Theorem 3.1.2 can be proved with H nilpotent (possibly at the cost of loosing the normality of H and P in $\langle S \rangle$). Indeed in characteristic 0 such a result follows by combining [**25**] with [**21**], and for prime fields it follows from Theorem 2.1.7 (proved in [**133**]) and the results of [**59**] for soluble groups. In general this may be technically quite challenging even though soluble linear groups have a nilpotent-by-abelian subgroup of n-bounded index.

It may also be possible that a more general polynomial inverse theorem holds (see [**20, 23**]). We pose the following question which bypasses abelian and finite obstacles.

Question 3.1.3. Let S be a finite symmetric subset of a group G such that $|S^3| \leq K|S|$ for some $K \geq 1$. Is it true that S is contained in $K^{c(G)}$ cosets of some virtually soluble subgroup of G?

The above question may be viewed as a counterpart of the Polynomial Freiman-Ruzsa Conjecture which asserts that (a variant of) Freiman's famous Inverse theorem holds with polynomial constants. The existence of some huge bound $f(K)$ for the number of covering cosets, as above, follows from the very general Inverse theorem of Breuillard, Green and Tao [**27**]. See Breuillard's survey [**20**] for a detailed discussion of these issues.

Theorem 3.1.2 shows that the answer is positive for the groups $G = SL(n, \mathbb{F})$. It would interesting to investigate this question for various groups of intermediate word growth, such as the Grigorchuk groups (see e.g. [**35**]).

It would be extremely interesting if the number of cosets required would be bounded by K^c for some absolute constant c for all groups G. Obtaining such a result for all linear groups G (in which c does not depend on the dimension) already seems to require some essential new ideas.

3.2. Basic results

Notation. For any group G let $\deg_{\mathbb{C}}(G)$ denote the minimum degree of a non-trivial complex representation.

Proposition 3.2.1 (Nikolov, Pyber [**113**])**.** *Let G be a finite group with $\deg_{\mathbb{C}}(G) \geq k$. Suppose that α, β and γ are subsets of G such that*

$$|\alpha||\beta||\gamma| > \frac{|G|^3}{k} .$$

Then $\alpha\beta\gamma = G$. In particular, if $|\alpha| > |G|/\sqrt[3]{k}$ then $\alpha^3 = G$.

Proposition 3.2.2. *Let $1 \in \alpha$ be a symmetric finite subset of a group G and $\tilde{G} = G/N$ a quotient of G. Set $\tilde{\alpha} = \alpha N/N$. Then $\left(|\alpha^3|/|\alpha|\right)^2 \geq |\tilde{\alpha}^3|/|\tilde{\alpha}|$.*

Definition 3.2.3. Let $1 \in \alpha$ be a symmetric finite generating set of a group G. We call α *weakly K-tripling* if for all quotients $\tilde{G} = G/N$ the projection $\tilde{\alpha} = \alpha N/N$ satisfies $\left|\tilde{\alpha}^3\right| \leq K^2 \cdot |\tilde{\alpha}|$.

Proposition 3.2.4 (Helfgott)**.** *Let $1 \in \alpha$ be a symmetric finite subset of a group and $k \geq 2$ an integer. Then*

$$\frac{\left|\alpha^k\right|}{|\alpha|} \leq \left(\frac{\left|\alpha^3\right|}{|\alpha|}\right)^{k-2}$$

Lemma 3.2.5. *Let $1 \in \alpha$ be a finite subset of a group G, and H a subgroup of G. Then for all integers $k > 0$ one has*

$$\frac{\left|\alpha^k \cap H\right|}{\left|\alpha^{-1}\alpha \cap H\right|} \leq \frac{\left|\alpha^{k+1}\right|}{\left|\alpha\right|} .$$

In particular, if α is symmetric then we have

$$\frac{\left|\alpha^k \cap H\right|}{\left|\alpha^2 \cap H\right|} \leq \left(\frac{\left|\alpha^3\right|}{\left|\alpha\right|}\right)^{k-1} .$$

Proposition 3.2.6. *Let G be a finite group and α a subset such that α^k contains a right coset Hx of a subgroup H. Then*

$$\frac{\max_{g \in G} |\alpha \cap gH|}{|H|} \geq \frac{|\alpha|}{\left|\alpha^{k+1}\right|} .$$

Proof Let t be the number of left cosets of H which contain elements of α. Then we have $\max_g |\alpha \cap gH| \cdot t \geq |\alpha|$. On the other hand it is clear that $\left|\alpha^{k+1}\right| \geq t|H|$. Hence

$$\frac{\left|\alpha^{k+1}\right|}{|H|} \geq t \geq \frac{|\alpha|}{\max_{g \in G} |\alpha \cap gH|}$$

as required.

Lemma 3.2.7. *Let $1 \in \alpha$ be a finite symmetric subset of a group G and assume that α^k contains a coset of a normal subgroup N. If $|\alpha^3| \leq K \cdot |\alpha|$ and $\deg_{\mathbb{C}}(N) \geq K^{3k}$ then α^3 contains a coset of N. In particular, if $N = G$ then $\alpha^3 = G$.*

Proof By Proposition 3.2.4 we have $|\alpha^{k+1}| \leq |\alpha| \cdot K^{k-1}$. By Proposition 3.2.6 there is a subset X of N of size at least $|N|\frac{|\alpha|}{|\alpha^{k+1}|} \geq \frac{|N|}{K^{k-1}}$ such that $aX \subseteq \alpha$ for some $a \in \alpha$. Now $\alpha^3 \supseteq aXaXaX = a^3(a^{-2}Xa^2)(a^{-1}Xa)X$. By our assumptions we have

$$|a^{-2}Xa^2||a^{-1}Xa||X| > |N|^3/\deg_{\mathbb{C}}(N) .$$

Hence by Proposition 3.2.1 we have $\alpha^3 \supseteq a^3N$, which proves our statement.

3.3. Affine conjugating trick

Proposition 3.3.1 (Rhemtulla [**135**])**.** *Let $G = \langle A, x_1, \ldots, x_n\rangle$ where A is an abelian normal subgroup of a group G. Then the set*

$$S = \{[a_1, x_1][a_2, x_2] \cdots [a_n, x_n] | a_i \in A\}$$

is precisely the subgroup $[A, G]$.

Definition 3.3.2. Let H be a group and A a $\mathbb{Z}H$-module. We denote by $[H, A]$ the submodule of A spanned by the set

$$\{a - ga \,|\, a \in A,\, g \in H\} .$$

This is equal to the commutator subgroup $[G, A]$ where G stands for the semidirect product $A \rtimes H$.

Lemma 3.3.3. *Let H be a d-generated finite group with $\deg_{\mathbb{C}}(H) > K^{21}$, and A a $\mathbb{Z}H$-module. Let $0 \in \alpha \subseteq A$ be a symmetric H-invariant set generating A. Then (with multiplicative notation) $|\alpha^3| > K \cdot |\alpha|$ or α^{7d} contains the submodule $[H, A]$.*

Proof Assume, that $|\alpha^3| \leq K \cdot |\alpha|$. Let us consider the semidirect product $G = A \rtimes H$, and the finite subset $\beta = H \cdot \alpha \subseteq G$. It is clear, that $\beta^j = H \cdot \alpha^j$ for all $j \geq 0$, hence $|\beta^3| \leq K \cdot |\beta|$. It is also clear, that β generates G.

We claim that β^6 contains all conjugates of H. Otherwise there is a conjugate H_0 of H contained in β^6 such that for some $b \in \beta$ the conjugate $H^* = b^{-1}H_0b$ is not contained in β^6. Since H^* is contained in β^8, by Lemma 3.2.5 and Proposition 3.2.4 we have

$$|H^*| = |\beta^8 \cap H^*| \leq \frac{|\beta^9|}{|\beta|} \cdot |\beta^2 \cap H^*| \leq K^7 \cdot |\beta^2 \cap H^*| .$$

So $H^* \subseteq \beta^6$ by Proposition 3.2.1, a contradiction.

The above claim implies that

$$h^{-1}a^{-1}ha \in \beta \cdot \beta^6 \cap A = \alpha^7$$

for all $a \in A$ and all $h \in H$.

Since A is abelian, this implies that α^7 contains all commutators of the form $g^{-1}a^{-1}ga$ for all $a \in A$ and all $g \in AH = G$. By Proposition 3.3.1 we have

$$\alpha^{7d} \supseteq [G, A] = [H, A] .$$

3.4. Finite nilpotent-by-*Lie** groups

Lemma 3.4.1. *For each d there is a constant $m = m(d)$ with the following property:*
Let N be a nilpotent normal subgroup in a finite group G satisfying $[G, N] = N$. Assume that G/N is d-generated, $\deg_{\mathbb{C}}(G) > K^m$ for some $K > 1$. Then for every symmetric generating set $1 \in \alpha \subseteq G$ that projects onto G/N we have $\left|\alpha^3\right| > K \cdot |\alpha|$ or $\alpha^3 = G$.

This is an immediate consequence of Proposition 3.2.2 and the following (somewhat technical) result.

Lemma 3.4.2. *For each d there is a constant $m = m(d)$ with the following property:*
Let N be a nilpotent normal subgroup in a finite group G satisfying $[G, N] = N$. Assume that G/N is d-generated, and $\deg_{\mathbb{C}}(G) > K^m$ for some $K > 1$. Let $1 \in \alpha$ be a symmetric generating set of G which is weakly K-tripling and projects onto G/N. Then $\alpha^3 = G$.

Proof We set $m = 378d + 1092$. We prove the lemma by induction on $|N|$. It follows from the induction hypothesis that if N_0 is any normal subgroup of G contained in N then α^3 projects onto G/N_0, i.e. $\alpha^3 N_0 = G$. (Note, that the condition $[G, N] = N$ is inherited in all quotient groups.)

We claim, that if there are elements $a, b \in G$ such that $1 \neq [a, b] \in \mathcal{Z}(G) \cap N$ then $\alpha^3 = G$. To see this consider the subgroup A generated by $[a, b]$. It is normal in G, and it consists of the commutator elements $[a^k, b] = [a, b]^k \quad (k = 1, 2, \dots)$. By the induction hypothesis α^3 contains elements $x_k \in a^k A$ and $y \in bA$. Since $A \leq \mathcal{Z}(G)$ we have $[x_k, y] = [a^k, b]$ for all k. Hence $(\alpha^3)^4 = \alpha^{12}$ contains A, therefore α^{15} contains $\alpha^3 A = G$. Since $\left|\alpha^3\right| \leq K^2 \cdot |\alpha|$, Lemma 3.2.7 implies our claim.

Suppose first that $\mathcal{Z}(G) \geq \mathcal{Z}(N)$. If N is abelian, then $[G, N] = 1$, so by assumption $N = \{1\}$, and $\alpha = G$, the induction step is complete in this case. On the other hand, if N is a nonabelian nilpotent group, then there are elements $a, b \in N$ such that $1 \neq [a, b] \in \mathcal{Z}(N) \leq \mathcal{Z}(G) \cap N$. The above claim completes the induction step in this case.

Next assume that $\mathcal{Z}(N)$ is not contained in $\mathcal{Z}(G)$. Let $M \leq \mathcal{Z}(N)$ be a minimal normal subgroup of G which is not contained in $\mathcal{Z}(G)$. We distinguish two subcases.

If $[G, M] \neq M$ then $\{1\} \neq [G, M] \leq \mathcal{Z}(G)$ by the definition of M. Since $[G, M] \leq N$, the above claim completes the induction step in this subcase.

Finally assume $[G, M] = M$. Since M is in the centre of N, M is a $\mathbb{Z}G/N$-module which by assumption satisfies $\left[G/N, M\right] = M$. First we show, that in this subcase α^7 contains an element $x \in M \setminus \mathcal{Z}(G)$. If $M \cap \mathcal{Z}(G) = \{1\}$ then we use the fact that $\alpha^3 M = G$ by the induction hypothesis. Since α generates G, α^4 contains two elements of some coset of M, hence α^7 contains a nontrivial element of M, i.e. an element of $M \setminus \mathcal{Z}(G)$. On the other hand,

if $B = M \cap \mathcal{Z}(G) \neq \{1\}$ then α^3 contains elements from each B-coset in G by the induction hypothesis. Hence already α^3 contains an element x of $M \setminus B = M \setminus \mathcal{Z}(G)$.

By construction x generates M as a $\mathbb{Z}G/N$-module hence $\alpha^7 \cap M$ also generates M. Since α projects onto G/N, taking the union of all α-conjugates of $\alpha^7 \cap M$ we obtain a symmetric G/N-invariant generating set β of M. By construction $\beta \subseteq \alpha^9$. Using $\beta^3 \subseteq \alpha^{27} \cap M$ and Lemma 3.2.5 we obtain that

$$\frac{|\beta^3|}{|\beta|} \leq \frac{|\alpha^{27} \cap M|}{|\alpha^7 \cap M|} \leq \frac{|\alpha^{27} \cap M|}{|\alpha^2 \cap M|} \leq \left(\frac{|\alpha^3|}{|\alpha|}\right)^{26} \leq \left(K^2\right)^{26} = K^{52}$$

We apply Lemma 3.3.3 to the module M and the generating set β. We obtain that $M \subseteq \beta^{7d} \subseteq \alpha^{63d}$, hence $\alpha^{63d+3} = G$ by the induction hypothesis. Since by assumption $\left|\alpha^3\right| \leq K^2 \cdot |\alpha|$, Lemma 3.2.7 implies that $\alpha^3 = G$. The induction step is complete in all cases.

Proposition 3.4.3 (Nikolov, Pyber [**113**])**.** *Let G be a finite subgroup of $GL(n, \mathbb{C})$. Then $G/Sol(G)$ has an embedding into the symmetric group of degree cn^2 for some absolute constant c. In particular $G/Sol(G)$ embeds into $GL(cn^2, \mathbb{C})$.*

Corollary 3.4.4. *Let N be a soluble normal subgroup in a finite perfect group G. Then*

$$\deg_{\mathbb{C}}\left(G/N\right) \leq c \deg_{\mathbb{C}}(G)^2 \leq \deg_{\mathbb{C}}(G)^{2+\log_2 c}\,.$$

Proof Consider a complex representation of G of degree $k = \deg_{\mathbb{C}}(G)$. Let $K \lhd G$ denote the kernel of this representation. The image KN/K of N in this representation is a soluble normal subgroup of the perfect group G/K. Hence $(G/K)/Sol(G/K)$ is a nontrivial quotient group of G/N. By Proposition 3.4.3 this quotient group has a nontrivial complex representation of degree ck^2, which is also a representation of G/N.

3.5. Generation

Lemma 3.5.1. *Let $1 \in \alpha$ be a finite symmetric subset of a group, G a finite perfect normal subgroup of $\langle\alpha\rangle$, and S a soluble normal subgroup of G. Then each proper subgroup H of G that projects onto G/S has an α-conjugate $H^a < G$ with $\left|H : (H^a \cap H)\right| \geq \deg_{\mathbb{C}}\left(G/S\right)$.*

Proof We argue indirectly. Let $H = H_0, H_1, \ldots, H_n$ be the list of all $\langle\alpha\rangle$-conjugates of H ordered so, that for each $i \geq 1$ there is an index $\sigma(i) < i$ such that H_i is an α-conjugate of $H_{\sigma(i)}$. Let M_i denote the intersection $M_i = H_0 \cap H_1 \cap \cdots \cap H_i$. Then, setting $A_i = \bigcap_{j \neq i, j \neq \sigma(i)} H_j$, for all $i > 0$ we have

$$\left|M_{i-1} : M_i\right| = \left|A_i \cap H_{\sigma(i)} : A_i \cap H_{\sigma(i)} \cap H_i\right| \leq \left|H_{\sigma(i)} : H_{\sigma(i)} \cap H_i\right| < \deg_{\mathbb{C}}\left(G/S\right)\,.$$

We project the chain $H = M_0 \geq M_1 \geq \cdots \geq M_n$ into G/N. We obtain a chain of subgroups, starting at G/N, such that each member has index less than $\deg_{\mathbb{C}}\left(G/S\right)$ in the previous one. The minimal index of a proper subgroup in G/S is at least $\deg_{\mathbb{C}}\left(G/S\right)$. This implies, that each M_i projects onto G/S. In particular, $SM_n = G$.

By construction M_n is α-invariant, in particular $M_n \triangleleft G$. By assumption $G/M_n \cong S/(M_n \cap S)$ is soluble and G is perfect, hence $G = M_n$, a contradiction.

Definition 3.5.2. A *section* of a group G is a quotient group $\Sigma = H/N$ where $N \trianglelefteq H$ are subgroups in G. Let α be a subset of G. The *trace of* α in Σ, denoted by $\mathrm{tr}(\alpha, \Sigma) \subseteq \Sigma$, is the projection of $\alpha \cap H$. We say, that α *covers* Σ, if $\mathrm{tr}(\alpha, \Sigma) = \Sigma$.

Theorem 3.5.3. *For each d there is a constant $m = m(d)$ with the following property:*
Let $1 \in \alpha$ be a finite symmetric subset of a group, and $N \leq G$ be finite normal subgroups of $\langle \alpha \rangle$ satisfying $[G, N] = N$. Suppose that N is nilpotent, G/N is d-generated, and $\deg_{\mathbb{C}}(G) > K^m$ for some $K > 1$. Assume, that α covers G/N. Then either $\left|\alpha^3\right| > K \cdot \left|\alpha\right|$, or α^6 contains G.

Lemma 3.5.4. *For each d there is a constant $m = m(d)$ with the following property:*
Let $1 \in \alpha$ be a finite symmetric subset of a group, and $N \leq G$ be finite normal subgroups of $\langle \alpha \rangle$ satisfying $[G, N] = N$. Suppose that N is nilpotent, G/N is d-generated, and $\deg_{\mathbb{C}}(G) > K^m$ for some $K > 1$. Assume, that α is weakly K-tripling and α covers G/N. Then $\alpha^2 \cap G$ generates G and α^6 contains G.

Proof We set $m = \max\big(10m_0(d)(2 + \log_2 c), 15\big)$ where $m_0(d)$ is the bound given in Lemma 3.4.1 and c is the constant in Corollary 3.4.4. Let $H \leq G$ be the subgroup generated by $\beta = \alpha^2 \cap G$. Since α covers G/N, we have $HN = G$.

By Lemma 3.2.5

$$\frac{|\beta^3|}{|\beta|} \leq \frac{|\alpha^6 \cap G|}{|\alpha^2 \cap G|} \leq \left(\frac{|\alpha^3|}{|\alpha|}\right)^5 \leq K^{10} \,. \tag{3.5.1}$$

We argue by induction on $|G|$. It follows from the induction hypothesis, that if N_0 is any normal subgroup of $\langle \alpha \rangle$ contained in N then $HN_0 = G$ and $\alpha^6 N_0$ contains G.

If $G = H$ then we apply Lemma 3.4.1 to $\beta \subseteq G$. Using (3.5.1) we obtain that $\beta^3 = G$, hence the induction step follows in this case. So we assume that $G \neq H$.

We claim that $\mathcal{Z}(G) \cap N = \{1\}$. Indeed, otherwise the induction hypothesis (applied to the quotient by $N_0 = \mathcal{Z}(G) \cap N$) implies that $H \cdot \mathcal{Z}(G) = G$. But then $H \triangleleft G$, and $G/H \cong \mathcal{Z}(G)/(H \cap \mathcal{Z}(G))$ is a nontrivial Abelian quotient of G, a contradiction.

Let M be a minimal α-invariant subgroup of $\mathcal{Z}(N)$ and $B = H \cap M$. Then $M \triangleleft G$, hence $B = H \cap M \triangleleft H$. Since we have $B \triangleleft N$ and by assumption $HN = G$, we obtain that B is normal in G.

By the induction hypothesis $HM = G$. Consider the natural isomorphism between $G/M = HM/M$ and $H/(H \cap M) = H/B$. Then $(N \cap H)/B$ corresponds to $(N \cap H)M/M$. By the modularity law

$$(N \cap H)M/M = (HM \cap N)/M = N/M \,.$$

The assumption $[G,N]=N$ implies that $\big[G/M,N/M\big]=N/M$ and hence $\big[H/B,(N\cap H)/B\big]=(N\cap H)/B$. On the other hand, $\big[H/B,(N\cap H)/B\big]=\big[H,N\cap H\big]B/B$. We obtain that

$$N\cap H=\big[H,N\cap H\big]B\,.$$

By assumption M is a simple $\mathbb{Z}\langle\alpha\rangle$-module. Consider $\mathrm{soc}(M)$ the $\mathbb{Z}G$ submodule of M generated by its simple submodules. Since G is normal in $\langle\alpha\rangle$, it follows that $\mathrm{soc}(M)$ is a $\mathbb{Z}\langle\alpha\rangle$ submodule of M, hence $\mathrm{soc}(M)=M$. Note that N acts trivially on M, hence M is actually a $\mathbb{Z}G\!/\!N$-module, and by the above M is generated by its simple $\mathbb{Z}G\!/\!N$-submodules. It follows that the $\mathbb{Z}G\!/\!N$-submodule B is also generated by simple submodules.

If S is a non-trivial simple submodule of B then clearly we have $[G/N,S]=S$. But B has no trivial summand as $\mathcal{Z}(G)=\{1\}$. Hence $\big[G/N,B\big]=B$. Since H projects onto G/N, we have $B=[H,B]$, which is contained in the subgroup $[H,N\cap H]$.

We obtain that $\big[H,N\cap H\big]=N\cap H$. Since $H\big/(N\cap H)\cong G/N$, it follows that H is perfect. Using Corollary 3.4.4 we see that $\deg_{\mathbb{C}}(H)\geq\deg_{\mathbb{C}}\big(G/N\big)^{1/(2+\log_2 c)}\geq K^{m/(2+\log_2 c)}\geq\big(K^{10}\big)^{m_0(d)}$. Since β projects onto $H\big/(N\cap H)$, we can apply Lemma 3.4.1 to the group H, the nilpotent normal subgroup $N\cap H$ and the generating set β. Using (3.5.1) we obtain that H is contained in $\beta^3\subseteq\alpha^6$.

By Lemma 3.5.1 H has an α-conjugate $H^a\leq G$ with $\big|H:(H^a\cap H)\big|>K^m$. Then $H^a\subseteq\alpha H\alpha$, hence α^8 contains H^a. On the other hand, $\alpha^2\cap H^a=(\alpha^2\cap G)\cap H^a\subseteq H\cap H^a$. By Lemma 3.2.5 we have

$$\big(K^2\big)^7>\left(\frac{|\alpha^3|}{|\alpha|}\right)^7\geq\frac{\big|\alpha^8\cap H^a\big|}{\big|\alpha^2\cap H^a\big|}\geq\frac{|H^a|}{|H\cap H^a|}>K^m\geq K^{15}\,,$$

a contradiction.

3.6. Linear Groups

In this section $\mathbb{F}$ denotes a field of characteristic $p>0$, unless stated otherwise.

Definition 3.6.1. As usual $Sol(G)$ denotes the soluble radical and $O_p(G)$ the maximal normal p-subgroup of a finite group G.

A finite perfect group G has a unique perfect central extension H of maximal order. The Schur multiplier $M(G)$ is the centre of H. A finite group is called *quasi-simple* if it is a perfect central extension of a finite simple group. We denote by $Lie^*(p)$ the set of direct products of simple groups of Lie type of characteristic p, and by $Lie^{**}(p)$ the set of central products of quasi-simple groups of Lie type of characteristic p. If $G/Sol(G)$ is in $Lie^*(p)$ then we call G a *soluble by $Lie^*(p)$ group*.

The following deep result is essentially due to Weisfeiler [**169**].

Proposition 3.6.2. *Let G be a finite subgroup of $GL(n,\mathbb{F})$. Then G has a normal subgroup H of index at most $f(n)$ such that $H\geq O_p(G)$ and $H/O_p(G)$ is the central product of an abelian p'-group and quasi-simple groups of Lie type of characteristic p, where the bound $f(n)$ depends on n.*

It was proved by Collins [**33**] that for $n \geq 71$ one can take $f(n) = (n+2)!$. Remarkably a (non-effective) version of the above result was obtained by Larsen and Pink [**95**] without relying on the classification of finite simple groups.

Lemma 3.6.3. *For each n there is a constant $D_1(n)$ with the following property.*
Let P be a finite subgroup of $GL(n, \mathbb{F})$ such that $\deg_{\mathbb{C}}(P) > D_1(n)$. Then

(a) $P/Sol(P)$ is in $Lie^*(p)$,
(b) the normal subgroup $N = \big[P, Sol(P)\big]$ is a p-group which satisfies $[P, N] = N$, and
(c) P/N is in $Lie^{**}(p)$ and $\big|Sol(P)/N\big| \leq (2n+1)^n$.
(d) Let α be a subset of $GL(n, \mathbb{F})$ which covers $P/Sol(P)$. Then α^3 covers P/N.

Proof We use the notation of Proposition 3.6.2, setting $G = P$. Assuming $D_1(n) \geq f(n)$, it follows that $P = H$. Assuming $D_1(n) \geq 2$, i.e. that P is perfect, we see that $P/O_p(P)$ is in $Lie^{**}(p)$. This implies (a), and implies also that $N \leq O_p(P)$.

If X is any normal subgroup of the perfect group P, then we have $\big[[X, P], P\big] = [X, P]$ by a consequence of the Three-Subgroup Lemma, see [**2**, (8.9)]. This completes the proof of (b).

The quotient P/N is perfect central extension of a group in $Lie^*(p)$, hence (see [**151**, Theorem 6.4]) it is in $Lie^{**}(p)$.

Let $P/Sol(P)$ be the direct product of the simple groups $L_1, \ldots, L_t$. We have $t \leq n$ by [**101**, Corollary 3.3]. Moreover, the Lie rank l_i of L_i is at most n for all i. It is easy to see that

$$\big|Sol(P)/N\big| \leq \prod \big|M(L_i)\big| \, .$$

It follows from [**89**, Theorem 5.14] that $M(L_i) \leq 2l_i + 1$ holds, provided that $|L_i|$ is greater than some absolute constant. Assuming that $D_1(n)$ is larger than this constant it follows that

$$\big|Sol(P)/N\big| \leq (2n+1)^n \, ,$$

i.e. (c) holds.

Consider the group P/N. The projection of $\alpha \cap P$ is a subset β of size at least $\big|P/Sol(P)\big| \geq |P/N|/(2n+1)^n$. Assuming that $D_1(n) > (2n+1)^{3n}$, Proposition 3.2.1 implies that $\beta^3 = P/N$. This proves (d).

Remark 3.6.4. In the above lemma we actually have $O_p(P) = N$. This can be shown using the (somewhat delicate) fact that if L is a finite simple group of Lie type of characteristic p then $\big|M(L)\big|$ is coprime to p with finitely many exceptions (see [**89**, Theorem 5.14]).

The following classical theorem of Malcev (see e.g. in [**166**]) makes it possible to use finite group theory to study properties of finitely generated linear groups.

Proposition 3.6.5. *Let Γ be a finitely generated subgroup of $GL(n, \mathbb{F})$. For every finite set of elements $g_1, \ldots, g_t$ of Γ there exists a finite field $\mathbb{K}$ of the same characteristic and a homomorphism $\phi : \Gamma \to GL(n, \mathbb{K})$ such that $\phi(g_1), \ldots, \phi(g_t)$ are all distinct.*

Proposition 3.6.6. *For each n there is a constant $D_2(n)$ with the following property.*
Let $1 \in \alpha$ be a finite symmetric subset of $GL(n,\mathbb{F})$ which satisfies $\left|\alpha^3\right| \leq K \cdot |\alpha|$ for some $K \geq \frac{3}{2}$. Let P be a soluble by $Lie^*(p)$ normal subgroup of $\langle\alpha\rangle$ such that α covers $P/Sol(P)$ and $P/Sol(P)$ is the product of at most n simple groups.

(a) If P is finite and $\deg_{\mathbb{C}}(P) > K^{D_2(n)}$ then α^{18} contains P.
(b) If $\deg_{\mathbb{C}}(\tilde{P}) > K^{2D_2(n)}$ holds for all finite quotients $\tilde{P}$ of P then P is finite.

Proof We set $K^{D_2(n)} = \max\left(D_1(n), K^{7m(2n)}\right)$ where $m()$ is as in Theorem 3.5.3.

Assume first that P is finite. Consider $N = \left[P, Sol(P)\right]$ and the $Lie^{**}(p)$ quotient P/N. Since finite simple groups are 2-generated, there is a $2n$-generated subgroup H of P/N which projects onto $P/Sol(P)$. Therefore H and $\mathcal{Z}(P/N)$ generate P/N, which implies that $(P/N)/H$ is abelian. But P/N is perfect, hence $P/N = H$ is $2n$-generated. By Lemma 3.6.3 the set $\beta = \alpha^3$ covers P/N, and by Proposition 3.2.4 we have $\frac{|\beta^3|}{|\beta|} \leq \frac{|\alpha^9|}{|\alpha|} \leq K^7$. Applying Theorem 3.5.3 we see that $\beta^6 = \alpha^{18}$ contains P as required.

Assume now by way of contradiction that P is infinite but $\deg_{\mathbb{C}}(\tilde{P}) > K^{2D_2(n)}$ for all finite quotients of P. Choose a set of $t > |\alpha^{18}|$ elements $g_1, \ldots, g_t$ in P. By Proposition 3.6.5 there is a finite field $\mathbb{K}$ and a homomorphism $\phi : \Gamma \to GL(n,\mathbb{K})$ such that $\phi(g_1), \ldots, \phi(g_t)$ are all distinct. Hence $\left|\phi(P)\right| > |\alpha^{18}|$. By Proposition 3.2.2 we have $\left|\phi(\alpha)^3\right| \leq K^2 \cdot \left|\phi(\alpha)\right|$. Applying (a) to $\phi(\alpha)$ and $\phi(P)$ we obtain that $\phi(\alpha)^{18}$ contains $\phi(P)$, which is impossible.

Proposition 3.6.7 (Freiman [**55**])**.** *Let α be a finite subset of a group G. If $|\alpha \cdot \alpha| < \frac{3}{2}|\alpha|$, then $H := \alpha \cdot \alpha^{-1}$ is a finite group of order $|\alpha \cdot \alpha|$, and $\alpha \subset H \cdot x = x \cdot H$ for some x in the normaliser of H.*

Theorem 3.6.8. *Let $\alpha \subseteq GL(n,\mathbb{F})$ be a finite symmetric subset such that $\left|\alpha^3\right| \leq K|\alpha|$ for some $K \geq \frac{3}{2}$. Then there are normal subgroups $S \leq \Gamma$ of $\langle\alpha\rangle$ and a bound $a(n)$ depending only on n such that $\Gamma \subseteq \alpha^6 S$, the subset α is contained in the union of $K^{a(n)}$ cosets of Γ, S is soluble, and the quotient group Γ/S is the product of finite simple groups of Lie type of the same characteristic as $\mathbb{F}$. Moreover, the Lie rank of the simple factors appearing in Γ/S is bounded by n, and the number of factors is also at most n.*

We need the following consequence of the Landazuri-Seitz bounds [**92**] on the minimal degrees of representations of finite simple groups.

Proposition 3.6.9. *Let L be a finite simple group of Lie type of rank r. Then $|L| < \deg_{\mathbb{C}}(L)^{Cr}$ for some absolute constant C.*

Proposition 3.6.10. *In the above theorem we may assume that $\deg_{\mathbb{C}}(\Gamma/S) > K^{D_3(n)}$ for some given function $D_3(n)$ by appropriately changing Γ and m.*

Proof Let Γ/S be the direct product of the simple groups $L_1, \ldots, L_t$. Let G be the product of the L_i with $\deg_{\mathbb{C}}(L_i) > K^{D_3(n)}$, and Γ^* the preimage

of G in Γ. Since G is a characteristic subgroup of Γ/S, we see that Γ^* is a characteristic subgroup of Γ. In particular, it is also normal in $\langle\alpha\rangle$.

The order of each of the remaining L_i is at most $\left(K^{D_3(n)}\right)^{Cn}$ by Proposition 3.6.9. Hence the index $\left|\Gamma : \Gamma^*\right| = \left|\Gamma/S : G\right|$ is at most $K^{Cn^2D_3(n)}$. Setting $a^*(n) = a(n) + Cn^2D_3(n)$ we see that α can be covered by $K^{a^*(n)}$ cosets of Γ^*. This proves our statement.

Proposition 3.6.11. *Let Γ be a soluble by $Lie^*(p)$ subgroup of $GL(n,\mathbb{F})$. Let $1 \in \alpha$ be a symmetric subset of Γ which covers all Lie type simple quotients of Γ. Assume that the Lie rank of the simple factors appearing in $\Gamma/Sol(\Gamma)$ is bounded by n. Let P be the last term of the derived series of Γ. Then P is a perfect soluble by $Lie^*(p)$ subgroup and $\alpha^{b(n)}$ covers all simple quotients of P, where $b(n)$ depends only on n.*

Proof This is proved for subgroups of $GL(n,\mathbb{F}_p)$ as Proposition 2.12.23. The proof given there applies to subgroups of $GL(n,\mathbb{F})$ without change.

Next we restate Lemma 2.13.9.

Lemma 3.6.12. *Let $L = L_1 \times \cdots \times L_m$ be a direct product of simple groups of Lie type of rank at most r. Let α be a symmetric generating set of L which projects onto all simple quotients of L. Then $\alpha^{b(m,r)} = L$ where $b(m,r)$ depends only on m and r.*

Theorem 3.6.13. *Let α be a symmetric subset of $GL(n,\mathbb{F})$ satisfying $|\alpha^3| \leq K|\alpha|$ for some $K \geq 1$. Then there are normal subgroups $P \leq \Gamma$ of $\langle\alpha\rangle$ such that P is finite and perfect, Γ/P is soluble, a coset of P is contained in α^3, and α is contained in the union of $K^{c(n)}$ cosets of Γ, where $c(n)$ depends on n.*

Proof When $K < \frac{3}{2}$ this follows from Proposition 3.6.7 with $\Gamma = \langle\alpha\rangle = \alpha^3$ and $c(n) = 1$. Assume that $K \geq \frac{3}{2}$. We set

$$D_3(n) = (2 + \log_2 c)d(n)\max\left(6D_2(n), 54\right) \quad \text{where} \quad d(n) = 6b(n)b(n,n)$$

and c is the constant in Corollary 3.4.4. Using Theorem 3.6.8 and Proposition 3.6.10 we obtain a soluble by $Lie^*(p)$ normal subgroup Γ of $\langle\alpha\rangle$ such that α^6 covers $\Gamma/Sol(\Gamma)$, α is contained in the union of $K^{a(n)}$ cosets of Γ, and $\deg_{\mathbb{C}}\left(\Gamma/Sol(\Gamma)\right) > K^{D_3(n)}$. Moreover, the number and the Lie rank of the simple factors appearing in $\Gamma/Sol(\Gamma)$ is at most n.

Let P be the last term in the derived series of Γ. Then P is a perfect soluble by $Lie^*(p)$ normal subgroup of $\langle\alpha\rangle$ such that $P/Sol(P) \cong \Gamma/Sol(\Gamma)$. Using Corollary 3.4.4 we obtain that $\deg_{\mathbb{C}}(\tilde{P}) > K^{D_4(n)}$ for all finite quotients $\tilde{P}$ of P where $D_4(n) = d(n)\max\left(6D_2(n), 54\right)$. Using Proposition 3.6.11 and Lemma 3.6.12 we obtain that $\beta = \alpha^{d(n)}$ covers $P/Sol(P)$ where $d(n) = 6b(n)b(n,n)$ as above.

By Proposition 3.2.4 we have $\frac{|\beta^3|}{|\beta|} \leq \frac{|\alpha^{3d(n)}|}{|\alpha|} < K^{3d(n)}$. Applying Proposition 3.6.6 to P and β we obtain that P is finite and $\beta^{18} = \alpha^{18d(n)}$ contains P.

Lemma 3.2.7 implies that α^3 contains a coset of P. The proof is complete.

dc_650_12

CHAPTER 4

Triple points in three families of plane curves

4.1. Introduction

The (very) general problem. Let Γ be a family of continuous curves in $\mathbb{R}^2$. We pick a set of n curves $\mathcal{G} = \{\gamma_1, \ldots, \gamma_n\} \subset \Gamma$ and a set of m points $\mathcal{P} = \{P_1, \ldots, P_m\} \in \mathbb{R}^2$ and define a graph on $\mathcal{G} \cup \mathcal{P}$ by connecting γ_i to P_j if γ_i passes through P_j. We shall call this (bipartite) graph the *incidence graph* of $\mathcal{G}$ and $\mathcal{P}$.

Certain properties of such graphs, especially the maximum possible number of edges as a function of n and m (i.e. bounds on the number of incidences) play central role in Computational Geometry as well as in Discrete or Combinatorial Geometry.

In Chapter 4 we study a "reverse" question:

> *if we know only the incidence graph (or some of its properties), can we infer something about the properties of the family* Γ*?*

Apart from trivial observations like "if two curves share two common points then Γ cannot be the family of straight lines", very little is known. (Actually, [**48**] contains a result that points to this direction, see Theorem 4.2.1 below.)

Many triple points. In terms of incidence graphs, a point P_j is a *triple point* if it is connected to at least three of the n curves in $\mathcal{G}$. Since three general curves do not pass through a common point, triple points can be considered as interesting coincidences.

Given a family Γ and a positive integer $n \in \mathbb{N}^+$, we select n curves $\gamma_1, \ldots, \gamma_n \in \Gamma$ so that the number of triple points is maximized, and denote this maximum by $\mathcal{T}_\Gamma(n)$. More generally, for three (not necessarily distinct) families $\Gamma_1, \Gamma_2, \Gamma_3$, we select n curves from each Γ_i $(i = 1, 2, 3)$ and call a point P a triple point if, for $i = 1, 2, 3$, there exist distinct $\gamma_i \in \Gamma_i$ that pass through P. (Usual bipartite graphs cannot represent such structures; certain "four–partite" graphs can, but we do not need them.) We denote the maximum number of such triple points by $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n)$, taken over all possible selections of the $n+n+n$ curves. We must emphasize that, even in this general case, we require that a triple point be the intersection of three *distinct* curves.

If any two curves intersect in at most B points (where B is a constant while n is large) then the maxima defined above really exist; in particular

$$\mathcal{T}_\Gamma(n) \leq B\binom{n}{2} \qquad \text{and} \qquad \mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n) \leq Bn^2,$$

since already the number of *pairwise* intersections in Γ (or between, say, Γ_1 and Γ_2) cannot exceed the claimed bound.

If no such B exists then no bound can be found for the $\mathcal{T}$ (e.g., if, for $i = 1 \dots 3$, Γ_i consists of the graphs of $y = i \cdot \sin x + t$, for $t \in \mathbb{R}$).[1] That is why, in what follows, we shall always assume the existence of such a B, i.e. that

(4.1.1) *no two curves intersect in more than B points.*

On the other hand, the number of "double" points can really attain this quadratic order of magnitude if the curves we select are in "sufficiently general position", e.g., if any two share a common point and these points are all distinct. This observation indicates that the "magic multiplicity" 3 is the smallest interesting value. In some cases even the number of triple points can be of order cn^2, e.g., for straight lines like those in Figure 4.5.1(c). However, as we shall see, in many cases the number of triple points is only $O(n^{2-\eta})$ for some constant $\eta \in (0, 1)$.

Problem 4.1.1. Characterize those families Γ, or triples of families $\Gamma_1, \Gamma_2, \Gamma_3$, for which $\mathcal{T}_\Gamma(n)$ or $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n)$, respectively, attains a quadratic order of magnitude (i.e. at least cn^2, for a fixed $c > 0$ and infinitely many n).

If the function $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n)$ for certain families $\Gamma_1, \Gamma_2, \Gamma_3$ attain a quadratic order of magnitude, a simple way to prove this is to exhibit n (or $n + n + n$) curves — for all $n \in \mathbb{N}$ — that have this many triple points.

The converse is harder: if a quadratic order of magnitude is impossible, how to demonstrate this? That is why our main result Theorem 4.4.1 concerns a sufficient condition for not having many triple points.

The main result at a "philosophical" level. Roughly speaking, we show the following (all notions will be defined rigorously, including "envelopes").

> *using suitable (slightly different from usual) definitions of "parametrised families" and "envelopes", if one of three algebraically parametrised families has an envelope which is not an envelope for any of the other two families, then*
>
> $$\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n) = \mathcal{O}\big(n^{2-\eta}\big),$$
>
> *for a positive $\eta > 0$ that depends only on the degree of the families.*

Since we do not want to spoil the Introduction with a lot of technical details, we must, for the time being, postpone the exact formulation of our main result; see Theorem 4.4.1 for a precise statement.

[1] It is perhaps unfortunate but we use the word "graph" in two completely different ways: until this point it was used to represent/emphasize the incidences of geometric curves. From now on graph theory is forgotten and the graph means the graph of a function.

Earlier results for straight lines. Studying the incidence structures of points and straight lines (more generally, of points and certain curves) has been one of the fundamental tasks of Combinatorial Geometry for a long time.

About 140 years ago Sylvester [**152**] posed his famous "Orchard Problem" which, in an equivalent (dual) form, asks for an arrangement of n straight lines in the Euclidean plane so that the number of triple points be maximized. Sylvester showed that if $\mathcal{L}$ denotes the family of all straight lines, then $\mathcal{T}_{\mathcal{L}}(n) = n^2/6 + \mathcal{O}(n)$ (cf. [**58**]). Recently Green–Tao [**66**] has shown that the largest possible value of $\mathcal{T}_{\mathcal{L}}(n)$ is $\lfloor n(n-3)/6 \rfloor + 1$.

The study of general "k–orchards" for $k \geq 4$ was initiated by Erdős. [2]

One of his conjectures resulted in a beautiful and widely applicable upper bound proven by Szemerédi and Trotter [**154**]. The most interesting special case of this bound asserts that

> *the number of incidences between n points and n straight lines in the Euclidean plane is at most $Cn^{4/3}$, for some absolute constant C.*

Since then, various proof techniques have been found, some of them even extending the Szemerédi–Trotter bound to "pseudo–lines" (i.e. curves with the property that any two intersect in at most one point) and "families with two degrees of freedom" (i.e. through any two given points there pass at most a bounded number of curves), see [**146**], [**119**], [**153**], and also the excellent monographs [**108**], [**118**].

Earlier results on unit circles. Another "orchard–like" problem was posed by Erdős in [**50**]: arrange n unit circles in the Euclidean plane so that the number of triple points be maximized. Denoting the family of all unit circles by $\mathcal{U}$, an upper bound of $\mathcal{T}_{\mathcal{U}}(n) \leq n(n-1)$ is obvious (since, as before, already the number of *pairwise* intersections obeys this bound). A lower bound of $\mathcal{T}_{\mathcal{U}}(n) \geq cn^{3/2}$ was proved in [**38**]. The gap between these two estimates is still wide open.

Also from another point of view, unit circles play a special role in Combinatorial Geometry. One of the most challenging unproved conjectures of Erdős concerns the maximum possible number of unit distances between n points in $\mathbb{R}^2$, and this can be bounded from above by half the number of incidences between the n points and n unit circles around them.

Since such circles obviously form a family with two degrees of freedom, they obey the aforementioned Szemerédi–Trotter bound — and it readily implies the best currently known upper bound on the number of unit distances [**146**].

The Szemerédi–Trotter bound is known to give the best order of magnitude for point-and-straight-line configurations, which is not the case for points and unit circles (let alone more general families with two degrees of freedom). Actually, it is widely believed that for unit circles and points

[2] The "k–orchard" problem asks: Given n points in the plane, how many straight lines can contain k points of them if no r of them are on a straight line ($r > k$). See [**19**], p315.

much better upper bounds hold on the number of incidences. Thus, according to the famous Erdős conjecture on unit distances, n points and n unit circles cannot have more than $n^{1+\varepsilon}$ incidences, for any $\varepsilon > 0$ and $n > n_0(\varepsilon)$.

However, to the best of our knowledge, no such bound has been found so far, since all existing methods consider the set of unit circles just as a family with two degrees of freedom. That is why the known tools cannot distinguish them from straight lines — for which the bound cannot be improved.

As an application of our Main Theorem 4.4.1, we show a combinatorial distinction between families of straight lines and families of unit circles in Section 4.5.

An outline of Chapter 4. Assume we have an *algebraically parametrised family* $\Gamma = \{\gamma^{(t)} \ : \ t \in T\}$ of curves, i.e. there is a polynomial $p \in \mathbb{R}[x, y, t]$ or $p \in \mathbb{C}[x, y, t]$ such that $\gamma^{(t)} = \{(x, y) \ : \ p(x, y, t) = 0\}$, for all t in the parameter domain T. Here we do not care whether the points of the individual curves are parametrised somehow; rather, *curves* are assigned to each parameter $t \in T$.

If three such curves, say $\gamma^{(t_1)}$, $\gamma^{(t_2)}$, $\gamma^{(t_3)}$ pass through a common point (x, y), then three equations $p(x, y, t_i) = 0$ are satisfied. Eliminating x and y we get another polynomial equation

$$F(t_1, t_2, t_3) = 0. \tag{4.1.2}$$

It was shown in [**48**] that, if some n elements of Γ determine $> cn^2$ triple points, then the surface $S_F := \{F = 0\}$ must be very special: *there exist three independent univariate coordinate transforms on the three axes which, together, transform S_F into a plane* — unless S_F is a cylinder. The details are given in the forthcoming Surface Theorem 4.2.1.

Unfortunately, that theorem does not provide a "good characterization" in the sense that it only states the equivalence of *existence* assumptions. (A "really good" and efficient tool would be one that says: "structure A exists if and only if structure B does not"; this would allow for an easy proof of "A does not exist" by simply exhibiting a B.)

Fortunately, a good characterization was also found in [**48**]: if we express, say, parameter t_3 from equation (4.1.2) then the implicit function $t_3(t_1, t_2)$ must satisfy a partial differential equation of order three. Theoretically this allows for proving subquadratic upper bounds on $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n)$ via elementary calculations, by showing that the differential equation is not satisfied.

In practice, however, even in simple, natural cases, these calculations may be impossible to carry out, even for powerful computers (see Section 4.5).

Our Main Theorem 4.4.1 becomes useful under such circumstances: it allows for similar bounds, based upon simple geometric considerations.

In Section 4.2 we present one of the most important tools for the proof of our Main Theorem: the Theorem 4.2.1, also called "Surface Theorem", proven in [**48**].

In order to prepare for the proof of our main result, we define partial envelopes and present some of their properties in Section 4.3. The main proof itself comes in Section 4.4.

In Sections 4.5 and 4.6 we state and prove our motivating Theorem 4.5.1: a combinatorial distinction between unit circles and straight lines.

Finally, we make some concluding remarks and formulate some conjectures.

4.2. Special surfaces

The first main ingredient of our proof is Theorem 4.2.1 below, proven in [**48**].

Assume we consider a plane $\alpha x + \beta y + \gamma z = \delta$, intersecting the cube $[0, n]^3$. If the coefficients $\alpha, \beta, \gamma, \delta$ are rationals with small numerators and denominators then this plane will contain $\sim n^2$ lattice points. If we apply independent univariate transformations in the three coordinates, x, y, z, then we can easily produce 2-dimensional surfaces — described by some equation $f(x) + g(y) + h(z) = \delta$ — containing a quadratic number of points from a product set $X \times Y \times Z$, where $|X| = |Y| = |Z| = n$. The main result of [**48**] asserts that if some appropriate algebraicity conditions hold, then (apart from being a cylinder) this is the only way for a surface $F(x, y, z) = 0$ to contain a near–quadratic number of points from such a product set $X \times Y \times Z$.

As usual, we call a (real or complex) function in one or two variable(s) *analytic* at a point if it can be expressed as a convergent power series in a neighborhood. Also, it is analytic on an open set if it is analytic at each point of the open set.

A *cylinder* over a curve $f(x, y) = 0$ is the surface

$$S := \big\{(x, y, z) \in \mathbb{C}^3 \ : \ f(x, y) = 0, \ z \in \mathbb{C}\big\}.$$

The definitions of cylinders over $g(x, z) = 0$ or $h(y, z) = 0$ are similar. It is worth noting that such cylinders always contain n^2 points of suitable $(\leq n) \times (\leq n) \times (\leq n)$ Cartesian products. To see this, just pick n arbitrary points on the curve $f(x, y) = 0$ and n arbitrary values $z_1, z_2, \ldots, z_n \in \mathbb{C}$. Denote the x and y coordinates of the points by X and Y, respectively, and let $Z := \{z_1, z_2, \ldots, z_n\}$. Then $|X|, |Y| \leq |Z| = n$ and $X \times Y \times Z$ contains at least n^2 points of S.

For the convenience of the reader we restate here Theorem 1.1.3, in a form slightly adapted to our current situation.

Theorem 4.2.1 (Surface Theorem)**.** *For any positive integer d there exist positive constants $\eta = \eta(d) \in (0, 1)$ and $n_0 = n_0(d)$ with the following property.*
If $V \subset \mathbb{C}^3$ is an algebraic surface (i.e. each component is two dimensional) of degree $\leq d$ then the following are equivalent:

(a) For at least one $n > n_0(d)$ there exist $X, Y, Z \subset \mathbb{C}$ such that $|X| = |Y| = |Z| = n$ and

$$|V \cap (X \times Y \times Z)| \geq n^{2-\eta};$$

(b) Let $\mathbb{D} \subset \mathbb{C}$ denote the open unit disc. Then either V contains a cylinder over a curve $F(x, y) = 0$ or $F(x, z) = 0$ or $F(y, z) = 0$ or, otherwise, there are one-to-one analytic functions $g_1, g_2, g_3 : \mathbb{D} \to \mathbb{C}$ with analytic inverses such that V contains the $g_1 \times g_2 \times g_3$-image of a part of the

plane $x+y+z=0$ *near the origin:*

$$V \supseteq \left\{ \Big(g_1(x), g_2(y), g_3(z)\Big) \in \mathbb{C}^3 \ : \ x,y,z \in \mathbb{D},\ x+y+z=0 \right\}.$$

(c) For all positive integers n *there exist* $X, Y, Z \subset \mathbb{C}$ *such that* $|X| = |Y| = |Z| = n$ *and* $|V \cap (X \times Y \times Z)| \geq (n-2)^2/8$.

(d) Both (b) *and* (c) *can be localized in the following sense. There is a finite subset* $H \subset \mathbb{C}$ *and an irreducible component* $V_0 \subseteq V$ *such that whenever* $P \in V_0$ *is a point whose coordinates are not in* H *and* $U \subseteq \mathbb{C}^3$ *is any neighborhood of* P*, then one may require that* $\Big(g_1(0), g_2(0), g_3(0)\Big) = P$ *in (b), and the Cartesian product* $X \times Y \times Z$ *in (c) lies entirely inside* U*. Furthermore,* P *has a neighborhood* U' *such that each irreducible component* W *of the analytic set* $V_0 \cap U'$*, with appropriate* g_1*,* g_2 *and* g_3*, can be written in the form*

$$W = \left\{ \Big(g_1(x), g_2(y), g_3(z)\Big) \in \mathbb{C}^3 \ : \ x,y,z \in \mathbb{D},\ x+y+z=0 \right\}.$$

If $V \subset \mathbb{R}^3$ *then the equivalence of* (a), (b), (c) *and* (d) *still holds true with* real *analytic functions* g_1*,* g_2*,* g_3 *defined on the interval* $(-1,1)$.

Remark 4.2.2. This version of (d) is in fact stronger than the original one in [**48**], but the proof given there applies without change to the stronger statement.

This result indicates a significant "jump": either V has the special form described in (b), in which case a quadratic order of magnitude is possible, by (b)⇒(c); or else we cannot even exceed $n^{2-\eta}$, by (a)⇒(b).

4.3. Families, and their envelopes

Definition 4.3.1. Let G be an open domain in $\mathbb{R}^2$ or $\mathbb{C}^2$. A *curve* in its closure $\overline{G}$ is a level set of a continuous function $\overline{G} \to \mathbb{C}$ which is analytic inside G.

Remark 4.3.2. We note, that these kind of curves are not necessarily connected, and they may have isolated points. However, this will not cause any trouble.

We consider families Γ of curves in $\mathbb{R}^2$ or $\mathbb{C}^2$, parametrised by the elements of a "parameter space" $T \subset \mathbb{R}$ or $T \subset \mathbb{C}$, like

$$\Gamma = \{\gamma^{(t)} \ : \ t \in T\}. \tag{4.3.1}$$

The parametrisation is an "implicit analytic parametrisation" if there exists a trivariate function f, analytic on an open domain $G \subset \mathbb{R}^3$ or $G \subset \mathbb{C}^3$ and continuous on its closure $\overline{G}$, such that

$$\gamma^{(t)} = \{(u,v) \ : \ f(u,v,t) = 0\}, \quad \text{for all } \ t \in T.$$

As opposed to implicit ones, we prefer explicit parametrisations.

Definition 4.3.3. Γ in (4.3.1) is *explicitly analytically parametrised* if there exists a *bivariate* function f, analytic on an open domain $G \subset \mathbb{R}^2$ or $G \subset \mathbb{C}^2$ and continuous on its closure $\overline{G}$, such that

$$\gamma^{(t)} = \{(u,v) \in \mathrm{cl}(G) \ : \ f(u,v) = t\} \quad \text{for all} \quad t \in T.$$

Remark 4.3.4. Curves of an implicitly analytically parametrised family can usually be cut into sub–arcs that can be parametrised explicitly — though we do not need this fact.

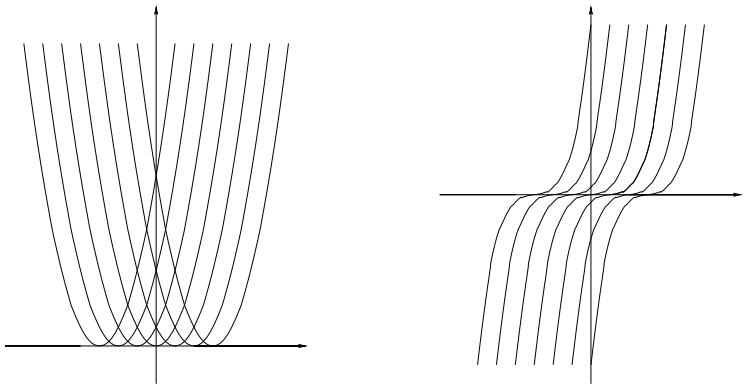

FIGURE 4.3.1. Implicitly analytically parametrised families:
(a) $y-(x-t)^2=0$ and **(b)** $y-(x-t)^3=0$.

The parabolas in Figure 4.3.1(a) cannot be parametrised explicitly since more than one curve passes through any point above the x–axis. As for the cubics in Figure 4.3.1(b), $t=x-\sqrt[3]{y}$ is a *continuous* parametrisation but it is not differentiable at any point of the x–axis (and so not analytic either). However, it is an explicit analytic parametrisation for suitable closed sub–arcs, say those in Figure 4.3.2(b).

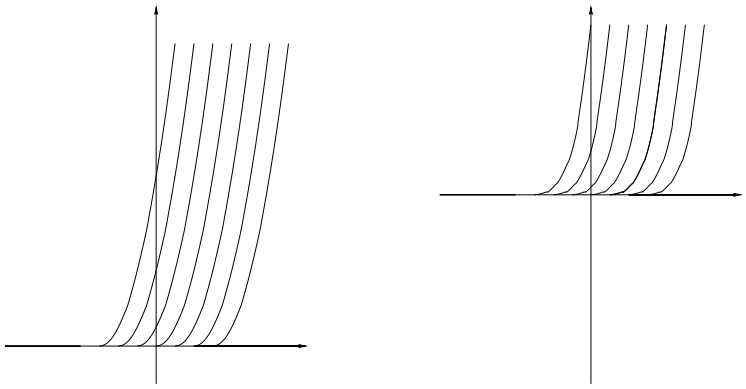

FIGURE 4.3.2. Explicitly analytically parametrised families:
(a) $t=x-\sqrt{y}$ and **(b)** $t=x-\sqrt[3]{y}$.

Envelopes of explicitly parametrised families. Usually in Differential Geometry an envelope of a family Γ of curves is a smooth curve that is tangent to each $\gamma\in\Gamma$. For *explicitly* parametrised families the situation is not that simple. E.g., in Figure 4.3.2(a)-(b), the x–axis is not a proper tangent line of the curves; rather, it only is a "half–tangent". Since this is typical in the case of sub–arcs of explicitly parametrised families, we shall use this general definition.

Definition 4.3.5. Let G be an open domain in the real or complex plane and let $\gamma \subset \overline{G}$ be a curve. A line L is the *half–tangent* of γ at a point P of the boundary $\mathrm{bd}(G)$ if $P \in \gamma \cap L$, P is not an isolated point of γ, and the following estimate holds:

$$\mathrm{dist}(Q, L) = o\left(\mathrm{dist}(Q, P)\right) \quad \text{for} \quad Q \in \gamma\ .$$

Definition 4.3.6. Two plane curves *touch* each other at a point P if there exists a straight line through P that is a tangent or half–tangent of both of the curves at P.

Definition 4.3.7. A smooth (open or closed) curve $\mathcal{E}$ is a *partial envelope* for an explicitly analytically parametrised family Γ, if

(i) $\mathcal{E}$ is the graph of an analytic real or complex function, say $y = h(x)$ or $x = h(y)$, defined on an open or closed interval or disk, respectively (i.e. $\mathcal{E} = \{(x, y)\ :\ y = h(x)\}$ or $\mathcal{E} = \{(x, y)\ :\ x = h(y)\}$);
(ii) no (non–empty open) sub–arc of $\mathcal{E}$ is contained in any $\gamma^{(t)} \in \Gamma$;
(iii) for each point $P \in \mathcal{E}$, there exists a t for which the curve $\gamma^{(t)} \in \Gamma$ touches $\mathcal{E}$ at P.

The adjective "partial" refers to the fact that we do not require that *each* $\gamma^{(t)} \in \Gamma$ touches $\mathcal{E}$.

Remark 4.3.8. (a) As we shall see in Lemma 4.3.10(ii), for explicitly analytically parametrised families, $\mathcal{E}$ must be a subset of $\mathrm{bd}(G)$. (Here $\mathcal{E} \subset \mathrm{cl}(G)$ is obvious since $\gamma^{(t)} \subset \mathrm{cl}(G)$ for all $\gamma^{(t)} \in \Gamma$.)
(b) Any non-trivial sub–arc of a partial envelope is a partial envelope;
(c) It is also worth noting that if a real $\mathcal{E}$ is a partial envelope for a family of analytically parametrised real curves then h can be extended to a complex analytic function whose graph defines a partial envelope for the family of the naturally extended, analytically parametrised complex curves.

The technical problems caused by *explicit* parametrisation may be tedious but, in general, they are not too difficult to manage.

Example 4.3.9. The unit circles through a given point, say the origin, form a family of implicitly analytically parametrised curves. Indeed, if (t, u) is the center of such a circle, then we can eliminate, say, u from the equations

$$(x - t)^2 + (y - u)^2 = 1 = t^2 + u^2, \tag{4.3.2}$$

and get a polynomial equation

$$4(x^2 + y^2)t^2 - 4x(x^2 + y^2)t + (x^2 + y^2)^2 - 4y^2 = 0.$$

Moreover, the circle $x^2 + y^2 = 4$ is obviously an envelope for them, in the usual Differential Geometric sense.
In order to get *explicitly* parametrised families, we express, say,

$$t = \frac{x}{2} \pm \frac{y}{2}\sqrt{\frac{4 - x^2 - y^2}{x^2 + y^2}}. \tag{4.3.3}$$

(Equivalently, we could express u in a symmetric manner.) Since the right hand side of (4.3.3) has no limit at the origin, we exclude a neighborhood of it, of a small radius δ, and consider the open set given by $x^2 + y^2 < 4$, $x^2 +$

$y^2 > \delta^2$, $y < \sqrt{1-(x-1)^2}$ and $x > \sqrt{1-(y+1)^2}$ as G (see the left hand side of Figure 4.6.1, where this domain is labelled as G_i^1, and the excluded neighbourhood is labelled as $B_\delta(a_i, b_i)$). Then the appropriate arcs of the unit circles are *explicitly* analytically parametrised on G by (4.3.3) with $+$ on the right hand side. We need four rotated copies of the domain G (labelled by $G_i^1, \dots G_i^4$ on Figure 4.6.1) to cover all "right-banding" semi-circles, and we need four more mirrored and rotated copies (labelled by $G_i^5, \dots G_i^8$ on Figure 4.6.1) to cover the "right-banding" semi-circles. Thus the whole family can be decomposed into eight explicitly parametrised (sub)families this way, four of them parametrised by t and four by u.
Moreover, each family has a quarter of the large circle as a partial envelope. (No portion of the small "inner circle" is an envelope since the unit circles do not touch it.)

A lemma on envelopes. In the proof of the Main Theorem 4.4.1, the following statement will play an important role.

Lemma 4.3.10. *Let Γ be a family of curves, explicitly analytically parametrised by $f : cl(G) \to \mathbb{C}$ or $\to \mathbb{R}$, as in Definition 4.3.3, and let $\mathcal{E}$ be a partial envelope. Then the following hold.*

(i) There are no points of $\mathcal{E}$ to which f can be extended analytically;
(ii) Consequently, we have $\mathcal{E} \subset bd(G)$.

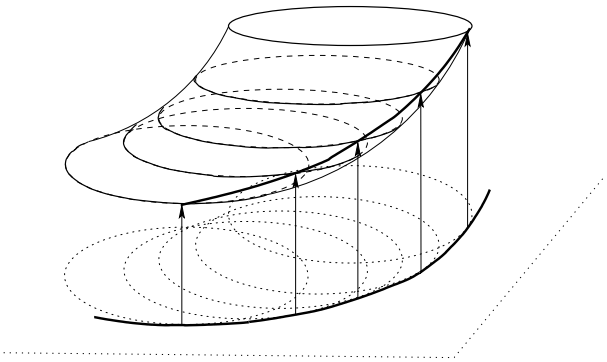

FIGURE 4.3.3. An envelope $\mathcal{E}$ (dashed) and its "lifting" by g on the cylinder over $\mathcal{E}$.

Proof To prove (i), we assume that f can be extended analytically to an open set $\tilde{G}$ which contains G and intersects $\mathcal{E}$. This means, that there is an analytic function $\tilde{f} : \tilde{G} \to \mathbb{C}$ which agrees with f on G. We replace $\mathcal{E}$ with $\tilde{G} \cap \mathcal{E}$, so from now on $\tilde{f}$ is defined and analytic at each point of $\mathcal{E}$. Also, let us define the extended curves $\tilde{\gamma}^{(t)} = \{(u, v) \,:\, t = \tilde{f}(u, v)\}$ for all t.
The function $f(x, y)$, if restricted to $\mathcal{E}$, gives, by definition, the parameter t of the curve $\gamma^{(t)} \in \Gamma$ that touches $\mathcal{E}$ at (x, y). Also by definition, $\mathcal{E}$ is the graph of an analytic function, say $y = h(x)$, on an interval or disk I (the case of $x = h(y)$ is similar). We consider the composition

$$g(x) := f(x, h(x)) : I \to \mathbb{C}.$$

This g is clearly continuous on I; moreover, since we assumed that f can be extended analytically to every $(x, h(x)) \in \mathcal{E}$, it is also differentiable, as

an univariate function, in the interior $\mathrm{int}(I)$, by the Chain Rule for the derivative of compositions of type $\mathbb{R} \to \mathbb{R}^2 \to \mathbb{R}$ or $\mathbb{C} \to \mathbb{C}^2 \to \mathbb{C}$.

Also, g cannot be a constant on $\mathcal{E}$ since $\mathcal{E}$ is not a subset of any $\gamma \in \Gamma$; thus there must exist a point $P_0(x_0, h(x_0)) \in \mathrm{int}(\mathcal{E})$ where $g'(x_0) \neq 0$. We are going to get the required contradiction by showing that the tangent plane of the graph of $\tilde{f}$ above P_0, i.e. at point $P_0^+ := \big(x_0, h(x_0), f(x_0, h(x_0))\big)$, is vertical — which is impossible.

To this end, we define two spatial curves on the graph of $\tilde{f}$ that pass through P_0^+ such that, at that point, the tangent lines of the two curves will both exist but will not coincide — hence they must span the tangent plane in question. Specifically, we consider the curves

$$\{\big(x, h(x), g(x)\big) \ : \ x \in I\}; \qquad \text{and}$$
$$\{\big(x, y, g(x_0)\big) \ : \ (x, y) \in \tilde{G}, \ \tilde{f}(x, y) = g(x_0)\};$$

the former one is the "lifting of $\mathcal{E}$ by function g" while the latter the lifting of the $\tilde{\gamma}^t$ that touches $\mathcal{E}$ at P_0 (i.e. it is $\tilde{\gamma}^{g(x_0)}$) to the fixed height $g(x_0)$. By assumption, there is a line L which is tangent to $\mathcal{E}$ and half–tangent to γ^t at P_0, hence must be tangent to the extended curve $\tilde{\gamma}^t$ at P_0. Hence both lifted curves have, indeed, tangent lines at P_0^+; that of the latter curve is obviously horizontal while that of the former one is not, by $g'(x_0) \neq 0$. Since both lines project to L in the base plane, we conclude that the tangent plane at P_0^+ must be vertical — the required contradiction to the assumption that f can be extended analytically to $\tilde{G}$.

Now (ii) follows from (i) since it implies that $\mathcal{E}$ can contain no (interior) point of the open set G.

This completes the proof of Lemma 4.3.10. ∎

4.4. The Main Theorem

The following is our main result. Though it concerns families of analytically parametrised curves, we need the technical assumption that there is an *algebraic*, i.e. polynomial relation between the families (the reason being that the Surface Theorem 4.2.1 works only for this case).

Theorem 4.4.1 (Main Theorem)**.** *Let $\Gamma_1, \Gamma_2, \Gamma_3$ be families explicitly parametrised by the functions f_1, f_2, f_3, analytic on open domains G_1, G_2, G_3 and continuous on $\overline{G_1}, \overline{G_2}, \overline{G_3}$, respectively, and with the property that $\mathbf{G} = G_1 \cap G_2 \cap G_3$ is connected. Assume that any two curves intersect in at most B points, and the concurrency of three curves $\gamma^{(t_i)} \in \Gamma_i$ $(i = 1, 2, 3)$ is described by a polynomial relation in the sense that, denoting a triple point where they intersect by (u, v), the three parameters $t_i = f_i(u, v)$ satisfy a polynomial relation $F(t_1, t_2, t_3) = 0$, or, more explicitly*

$$F\big(f_1(u, v), f_2(u, v), f_3(u, v)\big) = 0 \tag{4.4.1}$$

identically on $\overline{\mathbf{G}}$, for a polynomial $F \in \mathbb{C}[t_1, t_2, t_3]$. Assume, moreover, that

(i) Γ_3 has a partial envelope $\mathcal{E}$;
(ii) $\mathcal{E} \subseteq G_1 \cap G_2$;

(iii) No f_i ($i = 1, 2, 3$) is a constant on any non–empty open sub–arc of $\mathcal{E}$. (Intuitively: no non–empty open sub–arc of $\mathcal{E}$ is contained in any $\gamma \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$.)

Then

$$\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n) < B \cdot n^{2-\eta},$$

for a suitable $\eta = \eta(\deg(F))$ — provided that $n > n_0 = n_0(\deg(F))$.

Remark 4.4.2. The existence of an envelope $\mathcal{E}$ is sufficient but *not necessary* to make $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n)$ subquadratic. Actually, if no such envelope exists, then anything can happen. To see this, consider the three families of concentric circles about three points $P_1, P_2, P_3 \in \mathbb{R}^2$, respectively. (Obviously, none of these families possesses an envelope.) On the one hand, the method shown in [**39**] gives that, if the P_i are collinear, then $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n) \geq cn^2$. On the other hand, if they are non–collinear, then $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n)$ is subquadratic (see [**48**], Theorem 33).

Remark 4.4.3. The applicability of Theorem 4.4.1 is limited to *one–parameter* families Γ_i (the reason, again, being that the Surface Theorem 4.2.1 works only for such families).

Remark 4.4.4. It is worth noting that requirement (iii) in Theorem 4.4.1 is not just a technical assumption. E.g., the $n + n$ straight lines and n parabolas

$$\Gamma_1 := \{y = t_1^2 \ : \ t_1 = 0, 1, \ldots, n-1\};$$
$$\Gamma_2 := \{x = t_2 \ : \ t_2 = 0, 1, \ldots, n-1\};$$
$$\Gamma_3 := \{y = (x - t_3)^2 \ : \ t_3 = 0, 1, \ldots, n-1\};$$

have n^2 triple points — three curves of parameter t_1, t_2, t_3, respectively, pass through a common point if and only if $t_1 = |t_2 - t_3|$ — while the x–axis as $\mathcal{E}$ and the polynomial $F(t_1, t_2, t_3) := t_1^2 - (t_2 - t_3)^2$ satisfy all requirements but (iii).

Proof of the Main Theorem

(I) Without loss of generality we may assume that both the polynomial F and the surface $S_F = \{F = 0\}$ are irreducible. Indeed, the open domain $\mathbf{G}$ is connected, hence irreducible (as an analytic set). Therefore its image under the mapping

$$\mathbf{f} = f_1 \times f_2 \times f_3 \ : \ \overline{\mathbf{G}} \to S_F \subset \ \mathbb{R}^3 \text{ or } \mathbb{C}^3,$$

defined by

$$(u, v) \mapsto \Big(f_1(u, v), f_2(u, v), f_3(u, v)\Big)$$

is, again, irreducible. Then $\mathbf{f}(\mathbf{G})$ must be contained in a single irreducible component of the surface S_F, and one can simply throw away all other components. Moreover, the analytic functions f_i are non-constant, hence the polynomial F must depend on all three variables, and the surface S_F does not contain a cylinder over a curve (see Theorem 4.2.1(b)). Let $\eta = \eta(\deg(F)) \in (0, 1)$ and $n_0 = n_0(\deg(F))$ be the constants the existence of which is stated in Theorem 4.2.1. We want to show that $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n) < B \cdot n^{2-\eta}$, for $n > n_0$.

(II) Assume for a contradiction that $n+n+n$ curves with parameter sets T_1, T_2, T_3, respectively, determine $\geq B \cdot n^{2-\eta}$ triple points for an $n > n_0$. Any three curves, say of parameter t_1, t_2, t_3, respectively, share at most B common points. Therefore, the surface S_F passes through $\geq n^{2-\eta}$ points of the $n \times n \times n$ Cartesian product $T_1 \times T_2 \times T_3$. In other words, $V = S_F$ and the T_i as X, Y, Z satisfy Theorem 4.2.1(a).

(III) Consequently, we can use Theorem 4.2.1(d) and localize Theorem 4.2.1(b). This gives us a finite subset $H \subset \mathbb{R}$ or $\mathbb{C}$ of "exceptional" or "forbidden" values, and after picking a point P and a surface W in (IV) below, we shall also obtain three analytic functions $g_1, g_2, g_3 : \mathbb{D} \to \mathbb{C}$. Without loss of generality, we may assume that the partial envelope $\mathcal{E}$ of Γ_3 whose existence we assumed in the Main Theorem 4.4.1, has the property that

$$\forall P \in \mathcal{E} \text{ and } i = 1,2,3, \quad g_i(P) \notin H. \tag{4.4.2}$$

Indeed, this only excludes finitely many points from any closed sub–arc of $\mathcal{E}$ — since the g_i are nowhere constant by assumption (iii) — thus, if necessary, $\mathcal{E}$ can be restricted to a suitable open sub–arc.

(IV) Now we pick an arbitrary point $Q \in \mathcal{E}$. Clearly, $\mathbf{f}(Q) \in S_F$, since $\mathcal{E} \subset \overline{\mathbf{G}}$ by assumption (ii) and S_F is closed. Recall that $V_0 = V = S_F$ by the irreducibility assumption in (I), and $f_i(Q) \notin H$ for $i = 1,2,3$, by the assumption we made in (III), equation (4.4.2), so we can apply Theorem 4.2.1(d) and (b) to the point $P = \mathbf{f}(Q)$. Then we get a neighbourhood U' of $\mathbf{f}(Q)$, and the promised one-to-one analytic functions (with analytic inverses), $g_1, g_2, g_3 : (-1,1) \to \mathbb{R}$ or $\mathbb{D} \to \mathbb{C}$ with the following property: The function $\mathbf{g} = g_1 \times g_2 \times g_3$ maps the origin $(0,0,0)$ to $\mathbf{f}(Q)$, and maps an open subset of the plane $x+y+z=0$ onto the irreducible component of $W \subset S_F \cap U'$ containing $\mathbf{f}(\mathbf{G}) \cap U'$. This latter set is nonempty, since P lies inside U' and in the closure of $\mathbf{f}(\mathbf{G})$.

(V) Denote the inverses of the g_i by φ_1, φ_2, φ_3, respectively. Then the "coordinate-wise inverse" $\mathbf{g}^{-1} = \varphi_1 \times \varphi_2 \times \varphi_3$ maps W into the plane $x+y+z=0$. In other words, for $(t_1,t_2,t_3) \in W$ we have

$$\varphi_1(t_1) + \varphi_2(t_2) + \varphi_3(t_3) = 0,$$

since the three quantities on the left hand side are coordinates of a point in the plane $x+y+z=0$. But $\mathbf{f}(\mathbf{G}) \cap U' \subseteq W$, hence

$$\varphi_1\Big(f_1(u,v)\Big) + \varphi_2\Big(f_2(u,v)\Big) + \varphi_3\Big(f_3(u,v)\Big) = 0 \tag{4.4.3}$$

identically, in a neighborhood $\mathcal{U} \subset \overline{\mathbf{G}}$ of Q. (This $\mathcal{U}$ is open inside $\overline{\mathbf{G}}$ but not open in the plane, as Q is a boundary point.)

(VI) According to Lemma 4.3.10(i), f_3 cannot be extended analytically to any neighborhood of Q. On the other hand, re–writing (4.4.3) as

$$\varphi_3\big(f_3(x,y)\big) = -\varphi_1\big(f_1(x,y)\big) - \varphi_2\big(f_2(x,y)\big),$$

we get an explicit formula for f_3 in $\mathcal{U}$:

$$f_3(x,y) = g_3\Big(-\varphi_1\big(f_1(x,y)\big) - \varphi_2\big(f_2(x,y)\big)\Big) .$$

By assumption (ii) the right hand side is defined beyond Q, hence provides an analytic extension of f_3. This is the required contradiction. ∎

4.5. Straight lines or unit circles?

In this chapter we restrict our attention to the real plane $\mathbb{R}^2$. Recognizing unit circles (and, especially, distinguishing them from straight lines) does not seem to be difficult. E.g., anyone can tell that in Figure 4.5.1(a)–(b), there can only be found circles and no straight lines. Similarly, few people

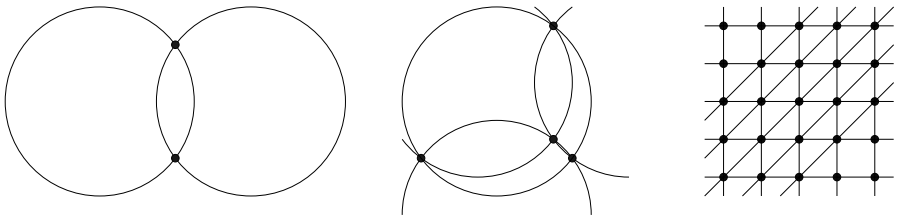

FIGURE 4.5.1. (a)–(b) unit circles; (c) straight lines?

would doubt that there is no unit circle in Figure 4.5.1(c), just straight lines. However, one should be more careful. How do we know that the lines are *really* straight? Perhaps they may be (arcs of) unit circles, provided that our "unit" is very large — so huge that their tiny little arcs do not even seem to be "bent". This is the moment when the points of the 5×5 lattice become important:

> *is it possible that 25 points and 15 unit circles are incident upon each other just like in Figure 4.5.1(c)?*

Unfortunately, we do not know the answer to this simple question. However, we are going to show that, for any $n > n_0$, the n^2 points of an $n \times n$ lattice and $3n$ lines in a similar grid–like configuration (n horizontal, n vertical and n "diagonal" ones) can only have this prescribed incidence pattern if the lines are really straight and cannot if they are (arcs of) unit circles — and this holds even if we only require a near–quadratic number of incidences.

Theorem 4.5.1. *There exist an absolute constant $\eta \in (0,1)$ and a threshold n_0 with the following property.*
Let (a_1, b_1), (a_2, b_2), (a_3, b_3) be three distinct points in the Euclidean plane and Γ_1, Γ_2, Γ_3 be three families of unit circles, such that, for each $i \leq 3$, all circles of Γ_i pass through the common point (a_i, b_i). Then

$$\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n) \leq 2^{10} \cdot n^{2-\eta} + 3, \tag{4.5.1}$$

provided that $n > n_0$.

Remark 4.5.2. The conjecture that in this case $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n) = o\big(n^2\big)$, originates from Székely (see [**42**, Conjecture 3.41]).

Remark 4.5.3. For straight lines the situation is quite different from the one described in Theorem 4.5.1. A configuration like the one in Figure 4.5.1(c) gives $\approx 3n^2/4$ triple points — where the three points (a_1, b_1), (a_2, b_2), and (a_3, b_3) which are common to the corresponding families of curves, — can

be considered as points on the line at infinity.
Similarly, if we allow *arbitrary* (i.e. not just unit) circles then they can produce any incidence pattern that straight lines can: just apply a suitable inversion to any configuration of points and straight lines. Even certain other conic sections have this property, e.g., shifted copies $y = x^2 + ax + b$ of the parabola $y = x^2$: just apply the diffeomorphism $(x, y) \mapsto (x, x^2 + y)$ to any configuration of points and straight lines.

4.6. Proof of Theorem 4.5.1

Assume we are given three families Γ_1, Γ_2, Γ_3 of unit circles and three distinct points (a_1, b_1), (a_2, b_2), $(a_3, b_3) \in \mathbb{R}^2$, with the property that all curves in Γ_i pass through (a_i, b_i), for $i \leq 3$.

(1) During the proof we do not consider the three points (a_i, b_i) as triple points (though they might be). This will only add a "+3" at the end.
(2) Pick a sufficiently small positive δ so that the δ–neighborhoods $B_\delta(a_i, b_i)$ do not contain any triple point.
(3) We subdivide each Γ_i in the way described in Example 4.3.9, this subdivision is pictured in Figure 4.6.1. Thus we get three times eight subfamilies denoted by G_i^k with $i = 1, 2, 3$ and $k = 1, 2, \ldots 8$. This subdivision will effect the bound on $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n)$ only by a factor of 8^3.

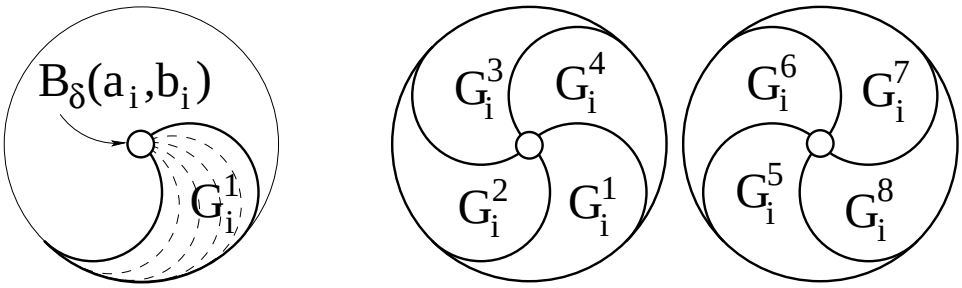

FIGURE 4.6.1. Subdivision of the family Γ_i into eight subfamilies. Left hand side: one of the subfamilies, G_i^1 (dotted arcs). Right hand side: all eight subfamilies.

(4) Each such $\Gamma_i^{(k)}$ only covers G_i^k once. Thus, as in Example 4.3.9, the family can be explicitly analytically parametrised.
(5) It is not difficult to find a trivariate polynomial equation $F(t_1, t_2, t_3) = 0$ that is satisfied by the parameters corresponding to any triple point. This is a rather straightforward calculation, our earlier manuscript had it. However, wishing to emphasize that we do not care for its actual form, we deleted it. (Actually, such polynomials can always be found in case of three *algebraically* parametrised families.)
(6) Each $\Gamma_i^{(k)}$ ($i \leq 3$, $k \leq 8$) possesses an envelope (a quarter circle of radius 2) that is not an envelope for any of the $\Gamma_j^{(l)}$ for $j \neq i$. Thus each triple $\langle \Gamma_1^{(k)}, \Gamma_2^{(l)}, \Gamma_3^{(m)} \rangle$, for $k, l, m \leq 8$, satisfies the

assumptions of the Main Theorem 4.4.1. Thus they cannot have more than $2n^{2-\eta}$ triple points for $n > n_0$ — where $\eta = \eta(\deg(F))$ and $n_0 = n_0(\deg(F))$ are as in Theorem 4.4.1.

(7) We conclude that, indeed, $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n) \leq 2^{10}n^{2-\eta} + 3$. ∎

Concluding remarks. We have given a sufficient condition for three one–parameter families of curves (or for three copies of a single family) to have "few", more specifically at most $n^{2-\eta}$ triple intersections.

How far below quadratic should it be? Since we have no reasonable estimate for $\eta > 0$, nothing is known about the exact order of magnitude. It may well be that the number of triple points is at most $n^{1+\varepsilon}$, for any $\varepsilon > 0$. We do not even know any families that satisfy the assumptions of Theorem 4.4.1 and can produce a super–linear number of triple points, say $n \log n$.

Which more-than-one parameter families of curves can determine a quadratic number of triple points? Our methods do not work in this generality, since Theorem 4.2.1 only applies to 1–parameter families.

We cannot help mentioning a related, beautiful, unsolved problem of Erdős. Assume that, in the projective plane, n straight lines define at least cn^2 *quadruple* points, i.e. points where at least four lines meet. Is it true that, for sufficiently large $n > n_0(c)$, there must exist a point where at least *five* of them intersect?

dc_650_12

CHAPTER 5

Triple Lines and Cubic Curves

5.1. Introduction

Given n point in the plane $\mathbb{R}^2$, a line is 3-*rich*, if it contains precisely 3 of the given points. One of the oldest problems of combinatorial geometry, the so-called Orchard Problem, is to maximise the number of 3-rich lines (see Jackson [**83**] and Sylvester [**152**]). Sylvester showed that the number of 3-rich lines is $n^2/6 + \mathcal{O}(n)$, and recently Green and Tao [**66**] have found the precise value of the maximum.

Theorem 5.1.1 (Orchard Problem. Green–Tao 2012)**.** *Suppose that $\mathcal{H}$ is a finite set of n points in the plane. Suppose that $n \geq n_0$ for some sufficiently large absolute constant n_0. Then there are no more then $\lfloor n(n-3)/6 \rfloor + 1$ lines that are 3-rich, that is they contain precisely 3 points of $\mathcal{H}$.*

Here we address the related problem of describing the structure of the *asymptotically near-optimal* configurations, i.e., of those for which the number of straight lines, which go through three or more points, has a quadratic (i.e., best possible) order of magnitude.

Definition 5.1.2. Let $\mathcal{H}$ be a subset of the plane $\mathbb{R}^2$. A straight line l is called a *triple line* with respect to $\mathcal{H}$ if there exist three distinct points $P_1, P_2, P_3 \in l \cap \mathcal{H}$. We shall also use the notation

$$\overset{\bullet\bullet\bullet}{\mathcal{H}} \overset{\text{def}}{=} \{l \ ; \ |l \cap \mathcal{H}| \geq 3\}.$$

We extended the notion of *triple line*, without any change in the definition, to subsets of the projective plane.

Note that $\overset{\bullet\bullet\bullet}{\mathcal{H}}$ is a set of *lines*, not a set of triples; e.g. if $\mathcal{H}$ is a collinear set of 3 or more points then $|\overset{\bullet\bullet\bullet}{\mathcal{H}}| = 1$.

Triple lines are not necessarily 3-rich (as they may be 4-rich, 5-rich, and so on), hence Theorem 5.1.1 does not directly bound the size of $\overset{\bullet\bullet\bullet}{\mathcal{H}}$. In any case, it is easy to find a (non-sharp) quadric upper bound. Indeed, each line with three points contains three segments of the $\binom{n}{2}$ which connect pairs of points of $\mathcal{H}$, hence

$$|\overset{\bullet\bullet\bullet}{\mathcal{H}}| \leq \frac{1}{3}\binom{n}{2} = \frac{n^2}{6} - n/6.$$

Definition 5.1.3 (quardic, cubic)**.** A *quadric* is a plane curve which is equal to the zero set of a polynomial of degree 2.
A *cubic* is a plane curve which is equal to the zero set of a polynomial of degree 3.

The following examples show four simple configurations for which the quadratic order of magnitude can really be attained. Two of them consist of three collinear point sets each, the third one is located on a conic and a straight line, while the fourth one on a cubic.

Example 5.1.4. If $\mathcal{H}_1$, $\mathcal{H}_2$, $\mathcal{H}_3$ are three copies of an arithmetic progression on three equidistant parallel lines then $|\overline{\bullet\bullet\bullet}(\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3)| \approx N^2/18$, where N denotes the total number of points and $\overline{\bullet\bullet\bullet}(\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3)$ denotes the set of lines l such that there exist three distinct points $P_i \in l \cap \mathcal{H}_i$ for $i = 1, 2, 3$.
(It is slightly better to place a point set of "double density" on the middle line.)

Example 5.1.5. Let P_1, P_2, P_3 be the vertices of a non–degenerate triangle, and $\mathcal{H}_i$ $(i = 1, 2, 3)$ point sets on the line through the vertices P_{i-1} and P_{i+1}, defined by

$$\mathcal{H}_i = \Big\{ X \ ; \ \frac{\overline{P_{i-1}X}}{\overline{XP_{i+1}}} \in \{\pm 1, \pm 2^{\pm 1}, \pm 4^{\pm 1}, \dots, \pm 2^{\pm(n-1)}\} \Big\}, \tag{5.1.1}$$

where $i \pm 1$ is used mod 3 in the indices of the P_i. (See Figure 5.1.1.)

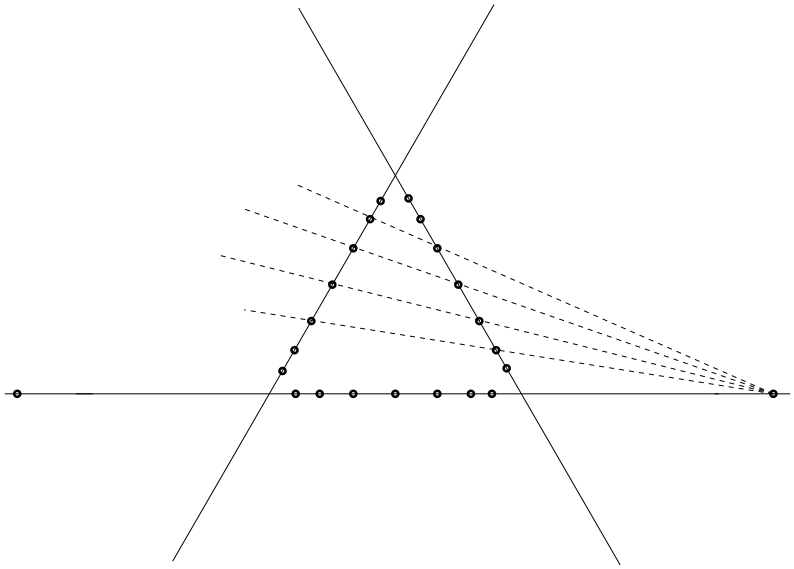

FIGURE 5.1.1. Portion of a triangular configuration with some triple lines marked.

Here again $|\overline{\bullet\bullet\bullet}(\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3)| \approx N^2/18$, where N denotes the total number of points. (The observant reader may have noticed that we allowed (-1) among the ratios, i.e., X may be a point at infinity.)

Example 5.1.6. The $\binom{n}{2}$ segments which connect pairs of vertices of a regular n–gon C only determine n distinct slopes. Let D be the set of points on the line at infinity which correspond to these directions. Then $|\overline{\bullet\bullet\bullet}(C, D)| \approx N^2/8$, where $N = |C \cup D| = 2n$ and $\overline{\bullet\bullet\bullet}(C, D)$ stands for $\overline{CCD}$.

Example 5.1.7. The point set $\mathcal{H} = \{(i, i^3) \ ; \ i = -n, \ldots, n\}$ on the curve $y = x^3$ satisfies $|\overset{\bullet\bullet\bullet}{\mathcal{H}}| \approx N^2/8$, where $N = 2n + 1$. This can easily be demonstrated by making use of the fact that three points (a, a^3), (b, b^3) and (c, c^3) are collinear iff $a + b + c = 0$.

The goal of Chapter 5 is to show that point sets with many triple lines are, from several points of view, closely related to cubics.

5.2. Problems and results

A conjecture. Since all the above examples with a quadratic order of magnitude of the triple lines involve cubic curves (some of which are degenerate), it is natural to believe the following.

Conjecture 5.2.1. *If $|\overset{\bullet\bullet\bullet}{\mathcal{H}}| \geq c|\mathcal{H}|^2$ then ten or more points of $\mathcal{H}$ lie on a (possibly degenerate) cubic, provided that $|\mathcal{H}| > n_0(c)$.*

Here the "magic number" 10 is the least non-trivial value since any nine points of $\mathbb{R}^2$ lie on a cubic. Perhaps even a stronger version may hold: *for every $c > 0$ and positive integer k there exist $c^* = c^*(c, k) > 0$ and $n_0 = n_0(c, k)$, such that, if $|\overset{\bullet\bullet\bullet}{\mathcal{H}}| \geq c|\mathcal{H}|^2$ then there is a con-cubic $\mathcal{H}^* \subset \mathcal{H}$ with $|\mathcal{H}^*| \geq k$ and $|\overset{\bullet\bullet\bullet}{\mathcal{H}^*}| \geq c^*|\mathcal{H}^*|^2$, provided that $|\mathcal{H}| \geq n_0$.*

It is very likely that in place of k above, even $c^*|\mathcal{H}|^\alpha$ con-cubic points exist (for some $c^* = c^*(c) > 0$ and $\alpha = \alpha(c) > 0$). An example with only $O(\sqrt{|\mathcal{H}|})$ such points is a $k \times k$ square or parallelogram lattice where the points of three parallel lines provide the set located on a (degenerate) cubic. Similarly, projections of d dimensional cube lattices to $\mathbb{R}^2$ form structures with only $O(|\mathcal{H}|^{1/d})$ con-cubic points.

Moreover, if we assume that $\mathcal{H}$ has no four–in–a–line and $|\overset{\bullet\bullet\bullet}{\mathcal{H}}| \geq c|\mathcal{H}|^2$, then perhaps as many as $c^*|\mathcal{H}|$ of its points will lie on an irreducible cubic.

Results. In order to support the above conjecture, we settle various special cases in the affirmative. Our main result is the following.

Theorem 5.2.2. *In $\mathbb{R}^2$, if irreducible algebraic curve of degree d contains a set $\mathcal{H}$ of n points with $|\overset{\bullet\bullet\bullet}{\mathcal{H}}| \geq cn^2$ then the curve is a cubic — provided that $n > n_0(c, d)$.*

Two simple applications of the forthcoming slightly more general Theorem 5.4.1 are the following.

Theorem 5.2.3. *In $\mathbb{R}^2$, no irreducible algebraic curve of degree d can accommodate n points with cn^2* quadruple *lines if $n > n_0(c, d)$.*

Theorem 5.2.4. *In $\mathbb{R}^2$, if a set of n points located on an irreducible algebraic curve of degree d only determines Cn distinct directions then the curve is a conic — provided that $n > n_0(d, C)$.*

The above theorems are of algebraic geometric nature, therefore it is natural to ask analogous questions in complex geometry (i.e. when the point set and the algebraic curves live in $\mathbb{C}^2$). However, in Chapter 5 we restrict our attention to the real plane $\mathbb{R}^2$.

In some other results (see Section 5.5) we allow part of the points (a positive proportion) to be arbitrary and only restrict the rest of them to a conic. In this case it will turn out that a large subset of the first part must be collinear. (Here again, the conic and the straight line, together, form a degenerate cubic.) The following is the essence of Theorems 5.5.1 and 5.5.2.

> *Let $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$ and assume that $\mathcal{H}_1$ lies on a (possibly degenerate) conic Γ while $\mathcal{H}_2 \cap \Gamma = \emptyset$. If $n \leq |\mathcal{H}_1|, |\mathcal{H}_2| \leq Cn$ and $|\overbrace{\mathcal{H}_1\mathcal{H}_2}| \geq cn^2$ then some c^*n points of $\mathcal{H}_2$ are collinear. (Here $c^* = c^*(c, C)$ does not depend on n.)*

We also mention a theorem of Jamison [**85**] which can be considered as another result in the direction of our Conjecture 5.2.1: if the diagonals and sides of a convex n–gon only determine n distinct slopes (which is smallest possible), then the vertices of the polygon all lie on an ellipse. In terms of triple lines (and a degenerate cubic formed by a straight line and an ellipse) this can be formulated as follows:

> *(Jamison's Theorem) if $\mathcal{H}_1$ is the vertex set of a convex polygon and $\mathcal{H}_2$ lies on the line at infinity with $|\mathcal{H}_1| = |\mathcal{H}_2| = n$ then $|\overbrace{\mathcal{H}_1\mathcal{H}_2}| = \binom{n}{2}$ implies that $\mathcal{H}_1$ lies on an ellipse.*

A similar statement was proven by Wettl [**174**] for finite projective planes.

The structure of Chapter 5. The aforementioned results (usually in stronger form) are presented in detail in the last two sections. Before that, we list some basic facts on the relation between continuous curves, collinearity and Abelian groups, concluding in the fundamental observation Lemma 5.3.8.

5.3. Collinearity and groups

Collinearity on cubics.

Definition 5.3.1. Let $\Gamma_1, \Gamma_2, \Gamma_3$ be three (not necessarily distinct) *Jordan curves* (i.e., bijective continuous images of an interval or a circle) in the projective plane, and $\langle \mathcal{A}, \oplus \rangle$ an Abelian topological group. We say that *collinearity between Γ_1, Γ_2 and Γ_3 can be described by the group operation $\oplus$*, if, for $i = 1, 2, 3$, there are homeomorphic monomorphisms (i.e., continuous injections whose inverses are also continuous)

$$f_i : \Gamma_i \to \mathcal{A}$$

— in other words, "parametrisation" of the Γ_i with $\mathcal{A}$ — such that three distinct points $P_1 \in \Gamma_1$, $P_2 \in \Gamma_2$, $P_3 \in \Gamma_3$ are collinear if and only if

$$f_1(P_1) \oplus f_2(P_2) \oplus f_3(P_3) = 0 \in \mathcal{A}.$$

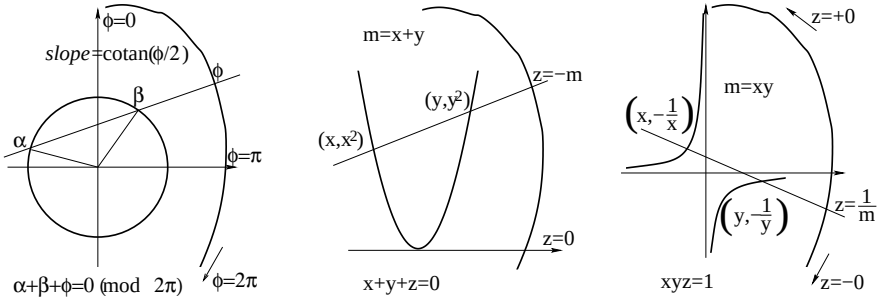

FIGURE 5.3.1. Parametrisation of reducible cubics: a conic plus the line at infinity. (Due to lack of sufficient space the line at infinity is depicted as a bent curve.)

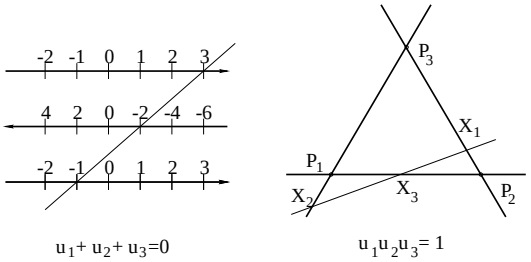

FIGURE 5.3.2. Parametrisation of reducible cubics: three straight lines. In case of a triangle, $u_i \stackrel{\text{def}}{=} f(X_i) = \overline{X_iP_{i-1}} \,/\, \overline{X_iP_{i+1}}$.

The curves we consider will usually be irreducible components of algebraic curves in $\mathbb{R}^2$ — or subsets thereof. However, sometimes we must also study general continuous curves, as well.

In what follows we denote the set of regular points of an algebraic curve Γ by $\mathrm{Reg}(\Gamma)$. The connected components of $\mathrm{Reg}(\Gamma)$ are Jordan curves.

Proposition 5.3.2. *Let $\mathcal{C}$ be a cubic curve in the projective plane. If Γ_1, Γ_2, Γ_3 are (not necessarily distinct) connected components of Reg($\mathcal{C}$), then collinearity between them can be described by commutative group operation — unless two of the Γ_i are identical straight lines.*

Indeed, for reducible cubics, Figures 5.3.1 and 5.3.2 show appropriate parametrisation in the real plane. (Any other reducible cubic is projective equivalent to one of these.) The groups used are $\langle \mathbb{R}, + \rangle / 2\pi\mathbb{Z}$, $\langle \mathbb{R}, + \rangle$, $\langle \mathbb{R} \setminus \{0\}, \cdot \rangle$ in Figure 5.3.1 and $\langle \mathbb{R}, + \rangle$, $\langle \mathbb{R} \setminus \{0\}, \cdot \rangle$ in Figure 5.3.2, respectively. If $\Gamma_1 = \Gamma_2 = \Gamma_3 = \mathcal{C} = \{(x, x^3) \,;\, x \in \mathbb{R}\}$ then the parametrisation $f(x, x^3) = x$ works well. It is also well-known that for irreducible cubics (i.e. elliptic curves), suitable parametrisation exist (see, e.g., in [**134**]).

Remark 5.3.3. Note that in all cases only regular points are parametrised. This will make no confusion since singular (e.g., multiple) points of a cubic never occur in proper collinear triples.

Collinearity on continuous curves. Throughout this section we consider the graphs of three continuous real functions.

Definition 5.3.4. We call α, β and γ a *standard system of continuous real functions* if

(i) they are defined in a neighbourhood $\mathcal{D}$ of 0;
(ii) $\alpha(x) < \beta(x) < \gamma(x)$ for all $x \in \mathcal{D}$;
(iii) any straight line through any point of the graph of any of the three functions intersects the other two graphs in at most one point each.

For such functions α, β and γ we denote their graphs (which are Jordan arcs) by $\overline{\alpha}$, $\overline{\beta}$ and $\overline{\gamma}$.

Remark 5.3.5. Assumption (iii) is not very strong a requirement; e.g., if the functions are differentiable at 0 (elsewhere they may not even be smooth) then $\mathcal{D}$ can be restricted to a sufficiently small neighbourhood of 0 so that (iii) be satisfied there.

Proposition 5.3.6. *Let $P(x, \beta(x))$ be a point of the "middle" graph $\overline{\beta}$. Connect it with lines to the two points $A_0(0, \alpha(0))$ and $C_0(0, \gamma(0))$; moreover, denote by $C(P)$ and $A(P)$ the points of intersection of these lines with the graphs $\overline{\gamma}$ and $\overline{\alpha}$, respectively (if they exist). Finally, let $B(P)$ be the intersection of the line through $A(P)$ and $C(P)$ with the graph $\overline{\beta}$. Then*

(i) if x is sufficiently close to 0 then $A(P)$, $B(P)$ and $C(P)$ really exist; and the composite mappings

$$x \;\mapsto\; P = P(x, \beta(x)) \;\mapsto\; \begin{cases} A(P) \text{ or} \\ B(P) \text{ or} \\ C(P) \end{cases}$$

are continuous functions $\mathbb{R} \to \mathbb{R}^2$;

(ii) for every point $\hat{B}$ of the graph $\overline{\beta}$, sufficiently close to the y–axis, there is a P for which $\hat{B} = B(P)$.

The straightforward proof using straightforward calculus — together with the Intermediate Value Theorem for (ii) — is left to the reader. ∎

Next we shall study when will collinearity between $\overline{\alpha}$ $\overline{\beta}$ and $\overline{\gamma}$ be described by an Abelian topological group $\mathcal{A}$, so we will search for parametrisations $f_\alpha : \overline{\alpha} \to \mathcal{A}$, $f_\beta : \overline{\beta} \to \mathcal{A}$ and $f_\gamma : \overline{\gamma} \to \mathcal{A}$. Part (iii) of Definition 5.3.4 also implies that the curves $\overline{\alpha}$, $\overline{\beta}$ and $\overline{\gamma}$ must be pairwise disjoint. That is why, in what follows, we shall only use one notation

$$f := (f_\alpha \cup f_\beta \cup f_\gamma) \;:\; (\overline{\alpha} \cup \overline{\beta} \cup \overline{\gamma}) \to A$$

in place of three.

Lemma 5.3.7 (Parameter–halving lemma)**.** *Let α, β and γ form a standard system of continuous real functions. Moreover, let $B_0 = (0, \beta(0))$ and A_0, C_0, P, $A = A(P)$, $B = B(P)$ and $C = C(P)$ be as above. Assume that collinearity between the three graphs is described by a group operation $\langle \mathcal{A}, \oplus \rangle$ and mapping (parametrisation) f. Then*

(i) if

$$f(P) = f(B_0) \oplus p \quad \textit{and}$$
$$f(B) = f(B_0) \oplus b$$

then $p = b/2$, *i.e.,* $b = p \oplus p$.

(ii) if B *is sufficiently close to* B_0 *then there really exists a* P *for which* $f(P) = f(B_0) \oplus b/2$.

Proof (i) Note that

$$f(A_0) \oplus f(B_0) \oplus f(C_0) = 0 \in \mathcal{A}.$$

Moreover, the collinearity of the triples C_0PA and CPA_0 imply

$$f(A) = f(A_0) \ominus p;$$
$$f(C) = f(C_0) \ominus p,$$

respectively; therefore

$$f(B) = p \oplus p \ominus f(A_0) \ominus f(C_0) =$$
$$= p \oplus p \oplus f(B_0),$$

whence the required identity.

(ii) is obvious from Proposition 5.3.6(ii). ∎

A fundamental lemma. The forthcoming Lemma 5.3.8 will work as our first tool for proving Theorem 5.2.2 and the slightly more general Theorem 5.4.1. The basic idea is to use the well-known construction of the group structure on cubics. If we know a few points on a cubic, then just by drawing specific lines and marking specific intersection points we can construct infinitely many new points on that cubic.

The essence of the following statement is that only on cubics can Abelian groups describe collinearity.

Lemma 5.3.8. *Let* α, β, γ *be a standard system of continuous functions defined in a neighbourhood of 0. Assume that collinearity between the three graphs is described by a group operation. Then their union* $\overline{\alpha} \cup \overline{\beta} \cup \overline{\gamma}$ *is contained in a (possibly reducible) cubic.*

For the proof we need certain special structures; they will be the topic of the next subsection. The proof itself comes then in the subsection afterwards.

Ten point configurations and cantilevers. Two types of point-line configurations will play special roles in what follows. The first one consists of ten points and a certain structure of triple lines while the latter will extend the former one.

Given $\overline{\alpha}$, $\overline{\beta}$, $\overline{\gamma}$ as in Lemma 5.3.8, we define *ten point configurations* as follows.

Denote, again, by A_0, B_0 and C_0 the points of intersection of the y–axis with the three graphs, respectively.

Choose B_1 on $\overline{\beta}$ sufficiently close to B_0 in order to make sure that all the forthcoming points exist. (This will be described later in more detail.) Let A_1 (resp. C_1) be the point of intersection of $\overline{\alpha}$ with the line through B_1 and C_0 (resp. that of $\overline{\gamma}$ with the line through B_1 and A_0). Define B_2 to

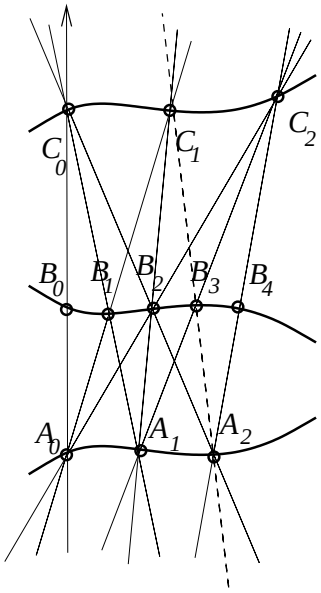

FIGURE 5.3.3. The straight line $A_2B_3C_1$ is not used in the definition of the points.

be the point of intersection of $\overline{\beta}$ with the line through A_1 and C_1. Let A_2 (resp. C_2) be the point of intersection of $\overline{\alpha}$ with the line through B_2 and C_0 (resp. that of $\overline{\gamma}$ with the line through B_2 and A_0).

The definition of B_3 is asymmetric: it will be the intersection of $\overline{\beta}$ with the line through A_1 and C_2. Finally, B_4 is, again, defined in a symmetric manner: the intersection of $\overline{\beta}$ with the line through A_2 and C_2 (see Figure 5.3.3). Note that by iterated application of Proposition 5.3.6, the rest of the points will all exist if B_1 is close enough to B_0.

The observant reader may have noticed that we defined eleven points altogether (instead of just ten). However, B_0 will NOT be in our configuration.

Definition 5.3.9. Given $\overline{\alpha}$, $\overline{\beta}$, $\overline{\gamma}$ as in Lemma 5.3.8, we call the above

$$\langle A_0, A_1, A_2, B_1, B_2, B_3, B_4, C_0, C_1, C_2 \rangle$$

a *ten point configuration* defined by B_1.

Proposition 5.3.10. *If α, β, γ is a standard system of continuous real functions and collinearity between their graphs is described by $\langle \mathcal{A}, \oplus \rangle$ and mapping f then*

(i) A_2, B_3 and C_1 are collinear.
(ii) More generally, A_i, B_j and C_k are collinear iff $i + k = j$.
(iii) There is a $\Delta \in \mathcal{A}$ such that $f(A_i) = f(A_0) \ominus i\Delta$, $f(B_i) = f(B_0) \oplus i\Delta$, and $f(C_i) = f(C_0) \ominus i\Delta$.

Proof Indeed, statement (ii) — with the exception of (i) — holds by definition. For $\Delta \stackrel{\text{def}}{=} f(B_1) \ominus f(B_0)$, this implies statement (iii) by group identities. Finally, (i) follows from (iii), using $f(A_0) \oplus f(B_0) \oplus f(C_0) = 0$, which, together with (iii), implies $f(A_2) \oplus f(B_3) \oplus f(C_1) = 0$. ∎

Lemma 5.3.11 (Ten point Lemma)**.** *Let $\overline{\alpha}$, $\overline{\beta}$, $\overline{\gamma}$ be a as in Lemma 5.3.8. Assume, moreover, that a ten point configuration defined on them is contained in two (possibly reducible) cubics $\mathcal{C}_1$ and $\mathcal{C}_2$. Then $\mathcal{C}_1 = \mathcal{C}_2$.*

Proof According to the definition of a standard system of continuous functions, if a straight line l contains two points of any of the three graphs then l is disjoint from the other two. This leaves us three possibilities for a cubic $\mathcal{C}_j$ $(j = 1, 2)$:

Type 1. three straight lines, one through the A_i, one through the B_i, and one through the C_i;
Type 2. a straight line through all (three or four) points of one of the graphs and a non-degenerate conic through the rest of them;
Type 3. an irreducible cubic through all the points.

According to Bézout's Theorem [**56**], two distinct irreducible algebraic curves of degree k and m, respectively, can only intersect in at most km points. This immediately implies the Lemma. Indeed, if we assume $\mathcal{C}_1 \neq \mathcal{C}_2$ for a contradiction, then e.g., if $\mathcal{C}_1$ is of type 2 and $\mathcal{C}_2$ of type 3 then either $\mathcal{C}_2$ and a straight line component of $\mathcal{C}_1$ intersect in four or more points, or $\mathcal{C}_2$ and a conic component of $\mathcal{C}_1$ intersect in seven or more points — a contradiction anyway. (The other pairs of types are easier.) ∎

Lemma 5.3.12 (Nine Point Lemma)**.** *Let $\overline{\alpha}$, $\overline{\beta}$, $\overline{\gamma}$ be a as in Lemma 5.3.8, consider a ten point configuration on them. If a (possibly reducible) cubic $\mathcal{C}$ contains, with the exception of B_3, the other nine points, then it must also contain B_3. Moreover, all ten points must belong to Reg($\mathcal{C}$).*

Proof Define $\delta \stackrel{\text{def}}{=} f(A_0) \oplus f(B_1) \oplus f(C_0) \in \mathcal{A}$. Then $\delta \neq 0$ since A_0, B_1 and C_0 are not collinear. What is $X \in \overline{\beta}$ for which $f(X) = 3\delta$? According to Proposition 5.3.10, it must be the point of intersection of the two straight lines $\overline{C_1 A_2}$ and $\overline{C_2 A_1}$. Finally, lines passing through a singular point $P \in \mathcal{C}$, if it has any, may contain at most two points of $\mathcal{C}$, so the lines in our ten point configuration may not pass through P. In particular, P cannot belong to a ten point configuration. ∎

Remark 5.3.13. Note that Lemmas 5.3.11 and 5.3.12 also imply that two cubics must coincide if they both contain the nine points (with the exception of B_3). However, we shall not need this fact.

Now we extend ten point configurations to what we call *cantilevers*.
(We hope that the shape of these structures will really justify this non-conventional notion.)

Starting from a ten point configuration on α, β, γ, we proceed recursively as follows.

Assume that B_i and B_{i+1} have already been defined for an $i \geq 3$. Then let C_i be the intersection of the lines $\overline{A_0 B_i}$ and $\overline{A_1 B_{i+1}}$ while A_i the intersection of the lines $\overline{C_0 B_i}$ and $\overline{C_1 B_{i+1}}$. Finally, define B_{i+2} to be the intersection of $\overline{A_2 C_i}$ and $\overline{C_2 A_i}$. (See Figure 5.3.4.) It is important to note that the construction of cantilevers use only the ten points, and does not depend on the three curves.

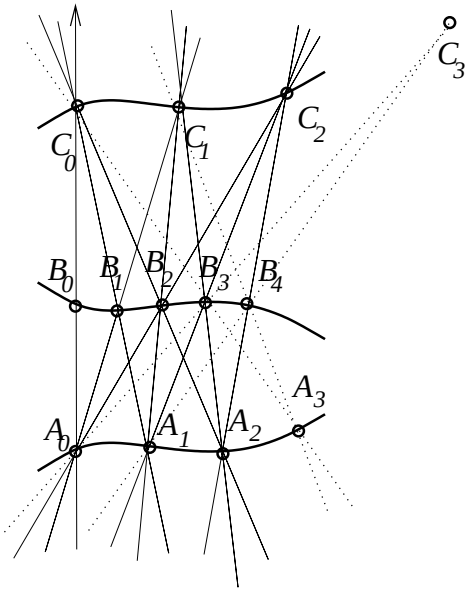

FIGURE 5.3.4.

Remark 5.3.14. Formally, here we work in the projective plane and even allow points of intersection located on the line at infinity. However, whenever we apply this construction, all points will lie on the curves $\overline{\alpha}$, $\overline{\beta}$, and $\overline{\gamma}$.

Lemma 5.3.15. *If the straight lines $\overline{A_0B_i}$ and $\overline{A_1B_{i+1}}$ intersect $\overline{\gamma}$ then this must happen at C_i, and similarly for $\overline{C_0B_i}$, $\overline{C_1B_{i+1}}$, $\overline{\alpha}$ and A_i. Moreover, if the above intersections all exist (and coincide with the C_i and the A_i, respectively), then B_{i+2} is located on $\overline{\beta}$.*

Proof Denote by X and Y the points of intersection of $\overline{\gamma}$ with $\overline{A_0B_i}$ and $\overline{A_1B_{i+1}}$, respectively. What is $f(X)$ then? By Proposition 5.3.10,

$$f(X) = \ominus f(A_0) \ominus f(B_i) = \ominus f(A_0) \ominus f(B_0) \ominus i\Delta = f(C_0) \ominus i\Delta.$$

Similarly, $f(Y) = f(C_0) \ominus (i+1-1)\Delta = f(X)$, whence $X = Y$. Therefore, also C_i must coincide with these points.
A similar argument proves the statement on B_{i+2}, too, since in that case the lines which define it must always intersect $\overline{\beta}$. ∎

Lemma 5.3.16. *If a cubic $\mathcal{C}$ contains the nine points A_0, A_1, A_2, B_1, B_2, B_4, C_0, C_1, C_2 of a ten point configuration then the entire cantilever (of infinite length) built from this configuration is contained in Reg($\mathcal{C}$).*

Proof By Lemma 5.3.12 the entire ten point configuration is contained in Reg($\mathcal{C}$). Let Γ_1, Γ_2, and Γ_3 denote the connected components of Reg($\mathcal{C}$) containing A_0, B_1, and C_0, respectively. By Proposition 5.3.2 the collinearity between the Γ_i is described by a group operation, let f_1, f_2, f_3 denote the parametrisations. In this case (i.e. for cubics) all f_i are bijections, hence they have inverse functions.

Consider the group element $\Delta = f_3(C_1) \ominus f_3(C_0)$. For all $n \geq 0$ we define the following points on $\mathcal{C}$:

$$\begin{aligned} A'_n &= f_1^{-1}\big(f_1(A_0) \ominus n\Delta\big) \\ B'_n &= f_2^{-1}\big(f_2(B_1) \oplus (n-1)\Delta\big) \\ C'_n &= f_3^{-1}\big(f_3(C_1) \ominus (n-1)\Delta\big) \end{aligned}$$

Plugging in $n = 0$ and $n = 1$ we obtain that

$$A'_0 = A_0, \quad B'_1 = B_1, \quad C'_0 = C_0, \quad C'_1 = C_1.$$

By assumption A_0, B_1, C_2 are collinear, hence $f_1(A_0) \oplus f_2(B_1) \oplus f_3(C_2) = 0$. This implies that

$$f_1(A'_i) \oplus f_2(B'_j) \oplus f_3(C'_k) = \ominus i\Delta \oplus (j-1)\Delta \ominus (k-1)\Delta = (i+k-j)\Delta$$

hence A'_i, B'_j, C'_k are collinear iff $i + k = j$.

Moreover, if a line can intersect $\mathcal{C}$ in at most three points, and if two of the intersection points are regular then all of them must be regular. Apply this to the line $\overline{C_0B_1} = \overline{C'_0B'_1}$. The third intersection point of this line with $\mathrm{Reg}(\mathcal{C})$ must be A_1 by Proposition 5.3.10, but above we proved it is A'_1. Therefore $A'_1 = A_1$. Similarly, the third intersection point of the line $\overline{A_1C_1} = \overline{A'_1C'_1}$ with $\mathrm{Reg}(\mathcal{C})$ must be B_2 on the one hand, and B'_2 on the other hand, which implies $B_2 = B'_2$. Finally apply the same argument to the lines $\overline{C_0B_2} = \overline{C'_0B'_2}$ and $\overline{A_0B_2} = \overline{A'_0B'_2}$ to obtain that $A_2 = A'_2$ and $C_2 = C'_2$.

To prove the lemma it is enough to show that $A'_n = A_n$, $B'_n = B_n$ and $C'_n = C_n$ for all $n \geq 1$. We prove it by induction on n. However, it is easier to do the induction with a slightly stronger statement. So we shall prove that

$$A'_n = A_n\ , \quad B'_{n+1} = B_{n+1}, \quad B'_{n+2} = B_{n+2}\ , \quad C'_n = C_n$$

for all $n \geq 0$. For $n = 0$ we have already seen this. Assume now that it is true for $n-1$. Consider the intersection point of the lines $\overline{C_0B_{n+1}} = \overline{C'_0B'_{n+1}}$ and $\overline{C_1B_{n+2}} = \overline{C'_1B'_{n+2}}$. On the one hand it must be A_{n+1}, on the other hand it is A'_{n+1}, hence $A'_{n+1} = A_{n+1}$. Similarly, the intersection point of the lines $\overline{A_0B_{n+1}} = \overline{A'_0B'_{n+1}}$ and $\overline{A_1B_{n+2}} = \overline{A'_1B'_{n+2}}$ must be $C'_{n+1} = C_{n+1}$. Finally, the intersection point of $\overline{C_2A_{n+1}} = \overline{C'_2A'_{n+1}}$ and $\overline{A_2C_{n+1}} = \overline{A'_2C'_{n+1}}$ must be $B_{n+3} = B'_{n+3}$. This completes the induction step. ∎

Proof of Lemma 5.3.8. It suffices to show that, for any x_0 in the (common) domain $\mathcal{D}$ of the functions α, β, and γ, there exists a cubic $\mathcal{C}$ which contains the three graphs *restricted to a sufficiently small neighbourhood* of x_0. Indeed, if we have such a neighbourhood (for each x_0) then it is possible to extend any of them as follows. Let $x_1 \in \mathcal{D}$ be one of the endpoints of this neighbourhood (interval) and consider a cubic $\mathcal{C}_1$ which contains the graphs in a neighbourhood of x_1. Within the intersection of the two intervals one can find a ten point configuration contained both by $\mathcal{C}$ and $\mathcal{C}_1$. By the Ten Point Lemma (Lemma 5.3.11), $\mathcal{C} = \mathcal{C}_1$, i.e., we have a longer neighbourhood of x_0. Thus the *maximal* such neighbourhood must be $\mathcal{D}$ itself.

Now we find an appropriate cubic in a neighbourhood of (without loss of generality) $x_0 = 0$. To start with, we select a ten point configuration, also include B_0, and extend it to the other side as follows. Start "backwards"

from the collinear triple A_2, B_4, C_2 and define (using B_3 in place of the original B_1) a $5+9+5$ point cantilever — with A_0, B_0 and C_0 in the "middle". We shall denote this structure by $\mathcal{H}$.

Define $B_{1/2}$ as in the Parameter Halving Lemma (Lemma 5.3.7) and, starting from A_0, B_0 and C_0, using this $B_{1/2}$ as reference point, define a cantilever with points A_i $(i=0,\ldots,4)$, B_i $(i=0,\ldots,8)$ and C_i $(i=0,\ldots,4)$. Of course, the new points will include the old ones, as well, by Proposition 5.3.10(iii). Also continue the structure "to the left" and denote this refined (halved) cantilever of 35 points by $\mathcal{H}_1$. Keep on defining $B_{1/2^n}$ and $\mathcal{H}_n$ by recursive halving, where the latter consists of $(2^{n+2}+1)+(2^{n+3}+1)+(2^{n+2}+1)=2^{n+4}+3$ points.

For each n, consider a cubic $\mathcal{C}_n$ which passes through A_0, $A_{1/2^n}$, $A_{2/2^n}$, $B_{1/2^n}$, $B_{2/2^n}$, $B_{4/2^n}$, C_0, $C_{1/2^n}$, and $C_{2/2^n}$. By Lemma 5.3.16 this cubic contains all points of $\mathcal{H}_n$. In particular, all $\mathcal{C}_n$ must contain the ten point configuration we started with, hence all these cubics are identical by the Ten Point Lemma (Lemma 5.3.11).

At this point we have a cubic $\mathcal{C}$ for which

$$\bigcup_n \mathcal{H}_n \subset \mathcal{C}.$$

On it, the halving process (starting from $\mathcal{H}_0$) gives exactly the same $\mathcal{H}_n$, whence the parameters which occur in $\cup_n \mathcal{H}_n$ are dense somewhere in an open set $\mathcal{U}$ of the topological group A. Hence so is the point set itself in three corresponding arcs of $\mathcal{C}$ (i.e., in the homeomorphic pre–images of $\mathcal{U}$). By the continuity of α, β, γ (and $\cup_n \mathcal{H}_n \subset \mathcal{C}$), these arcs are completely on $\mathcal{C}$, as well, thus providing the required common parts. ∎

Surfaces and groups. Let $F \in \mathbb{R}[x,y,z]$ be a polynomial of three real variables. Denote by

$$S = S_F \stackrel{\text{def}}{=} \{(x,y,z) \in \mathbb{R}^3 \ ; \ F(x,y,z)=0\}$$

its zero set, i.e., the algebraic surface described by the equation $F=0$. The degree of S_F is the (total) degree of its defining polynomial F.

Definition 5.3.17. We say that a surface $S \subset \mathbb{R}^3$ is *described by a commutative group operation* $\langle \mathcal{A}, \oplus \rangle$ if there are mappings ("parametrisations") $f_i : \mathbb{R} \mapsto \mathcal{A}$ for $i=1,2,3$ such that

$$(x_1,x_2,x_3) \in S \ \Leftrightarrow \ f_1(x_1) \oplus f_2(x_2) \oplus f_3(x_3) = 0.$$

E.g., the ball of equation $x^2+y^2+z^2=1$ is described by the additive group through the mappings $f_i(t) = t^2 - 1/3$ $(i=1,2,3)$.

One of the main ingredients of our proof is Theorem 5.3.18 below, proven in [**48**].

Assume we consider a plane $\alpha x + \beta y + \gamma z = \delta$, intersecting the cube $[0,n]^3$. If the coefficients $\alpha, \beta, \gamma, \delta$ are rationals with small numerators and denominators then this plane will contain $\sim n^2$ lattice points. If we apply independent uni-variate transformations in the three coordinates, x, y, z, then we can easily produce 2-dimensional surfaces — described by some

equation $f(x)+g(y)+h(z)=\delta$ — containing a quadratic number of points from a product set $X\times Y\times Z$, where $|X|=|Y|=|Z|=n$. The main result of [**48**] asserts that if some appropriate algebraicity conditions hold then (apart from being a cylinder) this is the only way for a surface $F(x,y,z)=0$ to contain a near–quadratic number of points from such a product set $X\times Y\times Z$.

As usual, we call a function of one or two variable(s) *analytic* at a point if it can be expressed as a convergent power series in a neighbourhood. Also, it is analytic on an open set if it is analytic at each of its points.

For the convenience of the reader we restate here Theorem 1.1.3, in a form slightly adapted to our current situation.

Theorem 5.3.18 (Surface Theorem)**.** *For any positive integer d there exist positive constants $\eta=\eta(c,d)$, $\lambda=\lambda(c,d)$ and $n_0=n_0(c,d)$ with the following property.*
If $V\subset\mathbb{R}^3$ is an algebraic surface (i.e. each component is two dimensional) of degree $\le d$ then the following are equivalent:

(a) For at least one $n>n_0(c,d)$ there exist $X,Y,Z\subset\mathbb{R}$ such that $|X|=|Y|=|Z|=n$ and

$$|V\cap(X\times Y\times Z)|\ge cn^{2-\eta};$$

(b) Let D denote the interval $(-1,1)$. Then either V contains a cylinder over a curve $F(x,y)=0$ or $F(x,z)=0$ or $F(y,z)=0$ or, otherwise, there are one-to-one analytic functions $f,g,h:D\to\mathbb{R}$ with analytic inverses such that V contains the $f\times g\times h$-image of a part of the plane $x+y+z=0$ near the origin:

$$V\supseteq\Big\{\Big(f(x),g(y),h(z)\Big)\in\mathbb{R}^3\ ;\ x,y,z\in D\ ;\ x+y+z=0\Big\};$$

(c) The statement in (b) *can be localised as follows. There is a finite subset $H\subset\mathbb{R}$ and an irreducible component $V_0\subseteq V$ such that whenever $P\in V_0$ is a point whose coordinates are not in H, then one may require that $\Big(f(0),g(0),h(0)\Big)=P$.*

This result indicates a significant "jump": either V has the special form described in (b), in which case a quadratic order of magnitude is possible, by (b)⇒(c); or, else, we cannot even exceed $n^{2-\eta}$, by (a)⇒(b).

5.4. Theorems on curves

Here we present some results on point sets located on algebraic curves and satisfying certain requirements.

The first one (Theorem 5.4.1) is a "gap version" of Theorem 5.2.2. It states that there is a significant difference between cubics and other algebraic curves: on a cubic, n points can determine as many as cn^2 triple lines; otherwise even as few as $n^{2-\eta}$ are impossible for n large enough.

The other result is related to a problem of Erdős. He asked if a point set with cn^2 *quadruple* lines must also contain a five-in-a-line. In Theorem 5.4.3 we settle this in the affirmative, under the additional assumption that the points lie on an algebraic curve.

Finally, Theorem 5.4.4 concerns point sets which determine few distinct directions.

Many triple lines force cubics. Our first main result states that, of all algebraic curves, only cubics can accommodate n points with $cn^{2-\eta}$ triple lines. This is probably far from being best possible; perhaps even the existence of as few as $cn^{1+\delta}$ such lines will also imply the same statement, for any $\delta > 0$ and $n > n_0(c, \delta)$.

Theorem 5.4.1. *For every $c > 0$ and positive integer d there exist $\eta = \eta(c, d)$ and $n_0 = n_0(c, d)$ with the following property. Let Γ_1, Γ_2, Γ_3 be (not necessarily distinct) irreducible algebraic curves of degree at most d in the plane $\mathbb{R}^2$. Assume that $n > n_0$ and*

(i) no two Γ_i are identical straight lines;
(ii) $\mathcal{H}_i \subset \Gamma_i$ with $|\mathcal{H}_i| \leq n$ $(i = 1, 2, 3)$;
(iii) $|\overline{\mathcal{H}_1\mathcal{H}_2\mathcal{H}_3}| \geq cn^{2-\eta}$.

Then $\Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ is a cubic.

Remark 5.4.2. If we have an *arbitrary* (i.e., possibly reducible) algebraic curve Γ of degree d and a point set $\mathcal{H}$ with many triple lines on it, then by the Pigeonhole Principle, some (at most three) irreducible components of Γ will contain a subset of $\mathcal{H}$ which still determines at least $|\overset{\bullet\bullet\bullet}{\mathcal{H}}|/d^3$ distinct triple lines. Therefore, the union of these components must be a cubic, according to the aforementioned Theorem.

Proof of Theorem 5.4.1. Let the curves Γ_1, Γ_2, Γ_3 be defined by the polynomial equations $F_1(x, y) = 0$, $F_2(x, y) = 0$, $F_3(x, y) = 0$, respectively. Three points $P_i(x_i, y_i) \in \Gamma_i$ $(i = 1, 2, 3)$ are collinear iff

$$F(x_1, y_1, x_2, y_2, x_3, y_3) \stackrel{\text{def}}{=} \begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix} = 0.$$

Eliminating the y_i from the system of the four equations

$$\big\{F(x_1, y_1, x_2, y_2, x_3, y_3) = 0\big\} \cup \big\{F_i(x_i, y_i) = 0 \quad (i = 1, 2, 3)\big\}, \tag{5.4.1}$$

we get a polynomial relation $f(x_1, x_2, x_3) = 0$. In other words, the projection to $\mathbb{R}^3$ (i.e., to the subspace spanned by the x_i coordinates) of the two dimensional algebraic variety defined by (5.4.1) in $\mathbb{R}^6$, will be contained in the zero-set of a single polynomial equation $f(x_1, x_2, x_3) = 0$.

Let $\eta = \eta(c, d)$ be as in Theorem 5.3.18. Denoting the set of the x coordinates of $\mathcal{H}_i$ by X_i $(i = 1, 2, 3)$, we have that the surface $S_f = \{f = 0\}$ contains at least $cn^{2-\eta}$ points of $X_1 \times X_2 \times X_3$.

In other words, (a) of the Surface Theorem 5.3.18 is satisfied for $V = S_f$ and the X_i. Since S_f cannot contain a cylinder by assumption (i), there exists an irreducible component $V_0 \subset S_f$ for which also (b) — localised as in (c) of the same Theorem — holds.

Pick a generic point $P(a_1, a_2, a_3) \in V_0 \subset S_f$. By the definition of the surface, there exist b_1, b_2, $b_3 \in \mathbb{R}$ such that, on the one hand, $Q_i(a_i, b_i) \in \Gamma_i$ for $i = 1, 2, 3$, while on the other hand, these Q_i are collinear. We can also assume without loss of generality, that these three points are distinct, they are regular points of $\Gamma_1 \cup \Gamma_2 \cup \Gamma_3$, and the straight line l which contains them is not tangent to Γ_i at Q_i $(i = 1, 2, 3)$. [Indeed, V_0 is two dimensional

by Theorem 5.3.18(b) while the points to be excluded form a finite number of one dimensional curves.]

Moreover, by (b) and (c) there, collinearity between sufficiently small arcs of the Γ_i around the Q_i is described by $\langle \mathbb{R}, + \rangle$. Now if we rotate and/or shift the plane so that l becomes the y axis then, according to Remark 5.3.5, in a sufficiently small neighbourhood of 0, the (rotated) Γ_i coincide with the graphs of a standard system of continuous functions. Thus we can use Lemma 5.3.8 to conclude that a suitable cubic $\mathcal{C}$ contains a non-empty open arc of each Γ_i. Thus also the union $\Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ of the three *irreducible* curves is contained in $\mathcal{C}$.

Finally, they cannot all be contained in a curve of degree < 3 since in that case they could not define many triple lines. Therefore, $\Gamma_1 \cup \Gamma_2 \cup \Gamma_3 = \mathcal{C}$. ∎

Four-in-a-line. Erdős [**52**] posed the problem whether a set of n points which contains cn^2 collinear four-tuples must also contain five collinear points. To our best knowledge, no progress has been made on this question so far.

In 1995, M. Simonovits asked the following. *Is it possible to find n points on an irreducible algebraic curve of degree 4 which determine cn^2 four-in-a-line?* (Of course, such a set can contain no five-in-a-line.) We show here that the answer is in the negative, even in a more general setting.

Theorem 5.4.3. *If an algebraic curve Γ of degree d accommodates a set $\mathcal{H}$ of n points with $cn^{2-\eta}$ distinct quadruple lines, where $\eta = \eta(c,d)$ is the same as in Theorem 5.4.1, then Γ contains four straight lines, each with $\geq c'(c,d) \cdot n^{1-\eta}$ points of $\mathcal{H}$, provided that $n > n_0(c,d)$.*

Proof Γ has at most d irreducible components. Classify the $cn^{2-\eta}$ collinear four-tuples (located on distinct straight lines) according to which point lies on which component. By the Pigeonhole Principle, some four (not necessarily distinct) components Γ_1, Γ_2, Γ_3 and Γ_4 generate $cn^{2-\eta}/d^4 = c'(c,d)n^{2-\eta}$ quadruple lines. By Theorem 5.4.1, any three of the Γ_i must form a cubic. However, this is only possible if they are distinct straight lines. ∎

Few directions. In [**45**], it was shown that if the graph of a polynomial $f \in \mathbb{R}[x]$ contains n points whose $\binom{n}{2}$ connecting lines only determine a linear number (at most Cn) distinct directions then the polynomial f is quadratic. (Some historic remarks and earlier results concerning sets which determine few directions can also be found there.)

Here we extend this to general algebraic curves.

Theorem 5.4.4. *For every $C > 0$ and positive integer d there is an $n_0 = n_0(C,d)$ with the following property.*
Let Γ_1 and Γ_2 be two (not necessarily distinct) irreducible algebraic curves, $n > n_0$, and $\mathcal{H}_i \subset \Gamma_i$ with $|\mathcal{H}_i| = n$ $(i = 1, 2)$. Assume that among the directions of the straight lines $\overline{P_1P_2}$, for $P_i \in \mathcal{H}_i$ and $P_1 \neq P_2$, at most Cn are distinct. Then $\Gamma_1 \cup \Gamma_2$ is a (possibly degenerate) conic.

Proof Let Γ_3 be the line at infinity and $\mathcal{H}_3$ the set of the $\leq Cn$ directions on it. (If someone prefers no points at infinity, they can apply a projective

mapping before proceeding further.) By assumption, $|\overline{\mathcal{H}_1\mathcal{H}_2\mathcal{H}_3}| \geq \binom{n}{2} > n^{2-\eta}$ if n is large. Hence, by Theorem 5.4.1, $\Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ is a cubic. Therefore, $\Gamma_1 \cup \Gamma_2$ is a conic. ∎

5.5. Straight lines and conics

Theorem 5.5.1. *Let $n \leq |\mathcal{H}_1|, |\mathcal{H}_2|, |\mathcal{H}_3| \leq Cn$ and assume that $\mathcal{H}_1$ and $\mathcal{H}_2$ lie on the distinct straight lines l_1 and l_2, respectively, while $\mathcal{H}_3 \cap l_1 = \mathcal{H}_3 \cap l_2 = \emptyset$.*
If, moreover, $|\overline{\mathcal{H}_1\mathcal{H}_2\mathcal{H}_3}| \geq cn^2$, then some c^*n of the points of $\mathcal{H}_3$, too, must be collinear. (Here $c^* = c^*(c, C)$ does not depend on n.)

Proof Apply a projective transform π which maps l_1 to the line at infinity. Then some cn^2 pairs of points of $\pi(\mathcal{H}_2) \times \pi(\mathcal{H}_3)$ determine at most $|\pi(\mathcal{H}_1)| = |\mathcal{H}_1| \leq Cn$ distinct directions, while $\pi(\mathcal{H}_2)$ is still collinear. By a result in [**41**] (see Theorem 3 there), also $\pi(\mathcal{H}_3)$ — hence $\mathcal{H}_3$, too — must contain c^*n collinear points. ∎

The following Theorem 5.5.2 is the "elder brother" of Theorem 5.5.1 in the sense that now we start from a non-degenerate conic while the two lines l_1, l_2 above can be considered as a degenerate one.

Theorem 5.5.2. *Let $C > 1$ be arbitrary and $\mathcal{H}_1$, $\mathcal{H}_2 \subset \mathbb{R}^2$. Assume that*
(a) $n \leq |\mathcal{H}_1|, |\mathcal{H}_2| \leq Cn$;
(b) $\mathcal{H}_2$ lies on a non-degenerate conic which contains no point of $\mathcal{H}_1$;
(c) $|\overline{\mathcal{H}_1\mathcal{H}_2}| \geq n^2$.
Then some c^*n of the points of $\mathcal{H}_1$ must be collinear (where $c^* = c^*(C)$ does not depend on n.)

Proof First, without loss of generality, we may assume that every point of $\mathcal{H}_1$ is incident upon at least n triple lines. (Otherwise keep on deleting those with less than $n/(2C)$ such lines and finally, use the new values of $n' = n/(2C)$, $C' = 2C^2$.)

Moreover, we may assume that the conic which contains $\mathcal{H}_2$, is the parabola $y = x^2$. (Else we apply a projective mapping which maps it to that curve. This can also be done such a way that no point of $\mathcal{H}_1$ is mapped to the line at infinity and the x–coordinates of the points in $\mathcal{H}_1 \cup \mathcal{H}_2$ become all distinct.)

Denote the coordinates of the points of $\mathcal{H}_1$ by (a_i, b_i) and the set of the x–coordinates of the points of $\mathcal{H}_2$ by X, i.e.,

$$\mathcal{H}_1 = \{(a_i, b_i) \mid i = 1, 2, \ldots, |\mathcal{H}_1|\};$$
$$\mathcal{H}_2 = \{(x, x^2) \mid x \in X\},$$

where, of course, $|X| = |\mathcal{H}_2|$.

Proposition 5.5.3. *Two distinct points (x, x^2), (y, y^2) of $\mathcal{H}_2$ and a point $(a_i, b_i) \in \mathcal{H}_1$ are collinear iff*

$$xy - a_i x - a_i y + b_i = 0. \text{ ∎}$$

The above equations can be considered as functions of type $X \mapsto X$:

$$y = f_i(x) \stackrel{\text{def}}{=} \frac{a_i x - b_i}{x - a_i}.$$

These projective mappings f_i are "vertical projections" (to X) of the involutions of the parabola, with centres (a_i, b_i).

We started with the assumption that every point of $\mathcal{H}_1$ is incident upon at least n triple lines. Therefore, each f_i maps at least n elements of X to elements of X. According to [**43**] Theorem 29 (the "Image Set Theorem"), some c^*n of the f_i must be collinear — if we represent them as elements of the three dimensional projective space. In other words, in that space at least c^*n points of projective coordinates $(a_i, -b_i, 1, -a_i)$ are combinations of as few as two of them, say $(a_1, -b_1, 1, -a_1)$ and $(a_2, -b_2, 1, -a_2)$. Considering the (constant) third coordinates, this is only possible if — even as four dimensional vectors — $(a_i, -b_i, 1, -a_i) = \lambda_i(a_1, -b_1, 1, -a_1) + (1 - \lambda_i)(a_2, -b_2, 1, -a_2)$, for suitable reals λ_i. We conclude that also the corresponding c^*n original points $P_i(a_i, b_i) \in \mathcal{H}_1 \subset \mathbb{R}^2$ must be collinear. ∎

5.6. Concluding remarks

Beyond Conjecture 5.2.1 the following remain open.

Problem 5.6.1. Let $\delta > 0$ be arbitrary. Does the conclusion "$\Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ is a cubic" of Theorem 5.4.1 hold if, in place of (iii), we only assume

$$\text{(iii*)} \qquad |\overline{\mathcal{H}_1\mathcal{H}_2\mathcal{H}_3}| \geq n^{1+\delta}$$

— provided that $n > n_0 = n_0(\delta, d)$?

Problem 5.6.2. Does Theorem 5.4.3 hold with $n^{1-\eta/2}$ in the statement (in place of $n^{1-\eta}$)?

Problem 5.6.3. Let $\delta > 0$ be arbitrary. Does the conclusion "$\Gamma_1 \cup \Gamma_2$ is a conic"of Theorem 5.4.4 hold if we only assume that the lines $\overline{P_1P_2}$ only determine $\leq n^{2-\delta}$ distinct directions — in place of Cn — provided that $n > n_0 = n_0(\delta, d)$?

dc_650_12

CHAPTER 6

The Weiss Conjecture

6.1. Introduction

A graph Γ is said to be *G-vertex-transitive* if G is a subgroup of $\mathrm{Aut}(\Gamma)$ acting transitively on the vertex set $V\Gamma$ of Γ. We say that a G-vertex-transitive graph Γ is *G-locally primitive* if the stabiliser G_α of the vertex α induces a primitive permutation group on the set $\Gamma(\alpha)$ of vertices adjacent to α. In 1978 *Richard Weiss* [**170**] *conjectured* that for a finite connected G-vertex-transitive, G-locally primitive graph Γ, the size of G_α is bounded above by some function depending only on the valency of Γ. In spirit this conjecture is similar to the 1967 conjecture of Charles Sims [**143**], that (stated in the graph theoretic context) for a G-vertex-primitive graph or digraph Γ, the size of the stabiliser of a vertex is bounded above by some function of the valency of Γ. In summary, in the conjecture of Weiss the graph Γ is assumed to be locally primitive, connected, and vertex-transitive; in the conjecture of Sims the graph Γ is assumed to be locally transitive and vertex-primitive (and hence connected). Despite the fact that the Sims' Conjecture has been proved true in [**28**], the truth of the Weiss Conjecture is still unsettled and only partial results are known, much of it focussing on the 'locally 2-transitive' case [**82, 163, 164, 171, 172, 173**] apart from the normal quotient reduction results in [**34, 126**].

In Chapter 6 we discuss the Weiss Conjecture and we prove it for groups with composition factors of bounded rank.

Definition 6.1.1. Define $\mathrm{BCP}(r)$ to be the class of finite groups G such that there is no section H/K of G, where $K < H \leq G$ and K normal in H, isomorphic to the alternating group $\mathrm{Alt}(r+1)$.

The class of $\mathrm{BCP}(r)$-groups was first considered by Babai, Cameron and Pálfy [**4**]. They showed that primitive $\mathrm{BCP}(r)$-groups of degree n have order at most $n^{f(r)}$. This result is an essential ingredient of many polynomial time algorithms for permutation groups related to the graph isomorphism problem [**86**]. The $\mathrm{BCP}(r)$-groups play also a very important role in the theory of subgroup growth of residually finite groups (see [**107**]). Chapter 6 uncovers a new application of this class of groups to the Weiss Conjecture.

We note that, a group $G \in \mathrm{BCP}(r)$ does not have as composition factor a simple group of Lie type of rank at least $r+1$, as such simple groups have $\mathrm{Alt}(r+1)$ as a section. On the other hand, if G has no simple groups of Lie type (resp. alternating groups) of rank (resp. degree) at least k as composition factors, then $G \in \mathrm{BCP}(r)$ for some $r \leq Ck$. This justifies our reference to these groups as having composition factors of bounded rank.

Theorem 6.1.2. *There exists a function $g : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ such that, for Γ a connected G-vertex-transitive, G-locally primitive graph of valency at most d, if G is a $\mathrm{BCP}(r)$-group, then a vertex stabiliser in G has size at most $g(r, d)$.*

Remark 6.1.3. In the light of this theorem the Weiss Conjecture asks whether the function g can be chosen not to depend on r.

Let Γ be a connected G-vertex-transitive graph of valency at most d. If G has a normal subgroup K with at least three orbits on $V\Gamma$, then the group $H = G/K$ is called an *intransitive head* of G. We write α^K for the K-orbit of the vertex α of Γ. The *normal quotient* Γ_K is the graph whose vertices are the orbits of K on $V\Gamma$, with an edge between two distinct vertices α^K and β^K in Γ_K, if and only if there is an edge of Γ between α' and β', for some $\alpha' \in \alpha^K$ and some $\beta' \in \beta^K$. It was proved in [**126**, Section 1] that Γ_K is an H-vertex-transitive graph of valency at most d. Furthermore, if Γ is G-locally primitive, then Γ_K is H-locally primitive, and the vertex stabilisers for H on Γ_K and for G on Γ are isomorphic groups. Thus for proving Theorem 6.1.2 it is sufficient to consider the case where there is no non-trivial normal quotient reduction. The groups G without non-identity normal subgroups K with at least three orbits on $V\Gamma$ (and hence admitting no reduction) are called *quasiprimitive* (if every non-identity normal subgroup of G is transitive) and *biquasiprimitive* (if G is not quasiprimitive and every non-identity normal subgroup of G has at most two orbits).

In [**34**] an analysis of G-locally primitive graphs with G quasiprimitive on vertices was undertaken, considering separately each of the eight types of quasiprimitive groups according to the quasiprimitive groups subdivision described in [**127**]. For six of the eight quasiprimitive types it was proved that $|G_\alpha|$ is bounded above by an explicit function of the valency, reducing the problem of proving the Weiss Conjecture for quasiprimitive groups G to the almost simple and product action types AS and PA ([**34**, Section 2]). The PA type was also examined in [**34**, Proposition 2.2] but unfortunately the proof contains an error. (We explain the mistake in Remark 6.1.9.)

In this chapter we actually prove a more general result from which, in the light of the comment above, Theorem 6.1.2 follows immediately.

Theorem 6.1.4. *There exists a function $g : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ such that, for Γ a connected G-vertex-transitive, G-locally primitive graph of valency at most d, if G has an intransitive head that is a $\mathrm{BCP}(r)$-group, then a vertex stabiliser in G has size at most $g(r, d)$.*

Proof of this result makes use of new results in two separate areas. First we apply new results of Praeger, Spiga and Verret [**129**] which reduce the proof of Theorem 6.1.4 to consideration of the case of finite simple groups G. Then we apply the Product theorem (see Theorem 2.1.4), which is due to Pyber, Szabó [**133**] and Breuillard, Green, Tao [**25**].

We mention that the results in [**129**] are for vertex-transitive graphs and in this context we prove the following more general version of Theorem 6.1.4.

Theorem 6.1.5. *There exists a function $g : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ such that, for Γ a connected G-vertex-transitive graph of valency at most d, if G has a*

BCP(r)*-group* G/K *as an intransitive head and* G/K *is quasiprimitive or biquasiprimitive on* $V\Gamma_K$*, then, for a vertex* α*, we have* $|G_\alpha/K_\alpha| \leq g(r,d)$*.*

If the graph Γ is G-locally primitive then by [**126**, Section 1] the subgroup K_α is trivial (and so Theorem 6.1.4 follows immediately from Theorem 6.1.5). As we noted in Remark 6.1.3, the Weiss Conjecture would assert that in this case the function g can be chosen not to depend on r. More generally we ask:

Question 6.1.6. What is the weakest local assumption that guarantees a bound on the size of G_α/K_α in terms of the valency d alone?

In Section 6.3 we give examples which show that some local assumption is needed even in the case when G is quasiprimitive. Namely we construct connected G-arc-transitive graphs Γ of valency $2r$ such that $G \cong \mathrm{Sym}\big((m+1)r-1\big)$ and $G_\alpha \cong \mathrm{Alt}(r)^{m-2} \times \mathrm{Alt}(r-1)$ for all $m \equiv 3 \mod 4$ and $r \geq 3$.

These examples also yield generating sets A of size $2(r!)^{m-1}$ in $\mathrm{Sym}\big((m+1)r-1\big)$ such that $|A^3| \leq 4r^2|A|$. This shows that the analogues of the Product theorem (valid for bounded rank families of simple groups of Lie type, see Theorem 2.1.4) do not hold for the family of finite symmetric groups.

It would be interesting to know whether there exist families of G-arc-transitive graphs of fixed valency with unbounded vertex stabilisers (where each G is isomorphic to some alternating or symmetric group) which are essentially different from the ones mentioned above. Motivated by the known examples we ask the following:

Question 6.1.7. Let Γ be a connected G-vertex-transitive graph or digraph. Is it true that the exponent of G_α is bounded in terms of the valency d?

A positive answer, even in the case when G is non-abelian simple, would be of great interest.

Theorem 6.1.5 raises another question:

Question 6.1.8. What is the weakest local assumption that guarantees a bound on the size of K_α in terms of the valency, for an intransitive normal subgroup K?

There are infinite families of vertex-transitive (and arc-transitive) graphs of fixed valency d with intransitive heads G/K such that K_α is unbounded. For example, in the case of wreath graphs $C_n[\overline{K}_{d/2}]$ of even valency d (that is, the lexicographic product of a cycle of length n with an edgeless graph on $d/2$ vertices), we have $G = \mathrm{Sym}(d/2) \,\mathrm{wr}\, D_{2n}$, $K = \mathrm{Sym}(d/2)^n$ and $K_\alpha = \mathrm{Sym}(d/2-1) \times \mathrm{Sym}(d/2)^{n-1}$.

Remark 6.1.9. In [**34**] the G-vertex-quasiprimitive, G-locally primitive graphs were analysed and an attempt was made to reduce the proof of the Weiss Conjecture in this case to the situation where G is an almost simple group. This approach cannot succeed, as shown in Example 6.2.7. This example demonstrates that there is no natural reduction to the case of G almost simple. A completely new combinatorial approach was needed, and developed in [**129**], to enable the Product theorem (2.1.4) to be applied.

Example 6.2.7 gives a connected graph Γ of valency 9 and a group $G \leq \mathrm{Aut}(\Gamma)$ with G quasiprimitive on vertices of product action type PA and a vertex stabiliser inducing $\mathrm{Sym}(3) \,\mathrm{wr}\, \mathrm{Sym}(2)$ in its primitive product action on the vertex neighbourhood. There is a naturally defined associated H-arc-transitive graph of valency 9 where H is almost simple with socle T, with T the simple direct factor of the socle of G. The (incorrect) proof of [**34**, Proposition 2.2] asserts that this graph is H-locally primitive. However for this graph the local action induced by a vertex stabiliser in H is $\mathrm{Sym}(3) \times \mathrm{Sym}(3)$ (having two intransitive normal subgroups $\mathrm{Sym}(3)$).

6.2. Proofs of the theorems

We start by deriving Theorems 6.1.2 and 6.1.4 from Theorem 6.1.5.

Proof of Theorems 6.1.2 and 6.1.4 from Theorem 6.1.5 . Let g be the function in the statement of Theorem 6.1.5. Assume that Γ is a connected G-vertex-transitive and G-locally primitive graph of valency at most d, and that G has an intransitive head G/K' that is a $\mathrm{BCP}(r)$-group. (For Theorem 6.1.2 take $K' = 1$.) Choose K maximal such that K has at least three orbits on the vertices of Γ with $K' \subseteq K$, we conclude that the action of G/K on the set of K-orbits is faithful and is either quasiprimitive or biquasiprimitive. Since G/K' is a $\mathrm{BCP}(r)$-group and $K' \subseteq K$, we have that G/K is a $\mathrm{BCP}(r)$-group. Let α be a vertex. By Theorem 6.1.5, $|G_\alpha|/|K_\alpha| \leq g(r,d)$. Since K_α is normal in G_α and since G_α induces a primitive action on the set $\Gamma(\alpha)$ of neighbours of α, either (i) K_α is transitive on $\Gamma(\alpha)$, or (ii) K_α fixes $\Gamma(\alpha)$ pointwise. We now use the fact that Γ is connected. In case (i), since G is vertex-transitive, K_β is transitive on $\Gamma(\beta)$ for all vertices β, and it follows from connectivity that K is edge-transitive and so has at most two orbits on vertices, contradicting the assumption that G/K is an intransitive head. In case (ii) by connectivity, K_α fixes every vertex of Γ and so $K_\alpha = 1$. Hence $|G_\alpha| \leq g(r,d)$. $\square$

Before embarking on the proof of Theorem 6.1.5 we recall the definition of *coset graph* and some elementary results.

Definition 6.2.1. [coset graph] Let G be a group, H a subgroup of G and A a subset of G. The coset digraph $\mathrm{Cos}(G,H,A)$ is the digraph with vertex set the right cosets of H in G and with arcs the ordered pairs (Hx, Hy) such that $Hyx^{-1}H \subseteq HAH$ (where $HAH = \{hsk \mid h,k \in H, a \in A\}$). Since $\mathrm{Cos}(G,H,A) = \mathrm{Cos}(G,H,HAH)$, replacing A by HAH, we may assume that A is a union of H-double cosets, that is, A is a disjoint union $\cup_{s\in S} HsH$ for some subset S of G.

It is immediate to check that $\mathrm{Cos}(G,H,A)$ is undirected if and only if $A = A^{-1}$, and $\mathrm{Cos}(G,H,A)$ is connected if and only if $G = \langle A \rangle$. Also the action of G by right multiplication of G/H induces a vertex-transitive automorphism group of $\mathrm{Cos}(G,H,A)$.

It was proved by Sabidussi [**141**] that every G-vertex-transitive graph Γ is isomorphic to some coset graph of G. More precisely, we have the following well-known result.

Proposition 6.2.2. *Let Γ be a G-vertex-transitive graph and α a vertex of Γ. Then there exists a union A of G_α-double cosets such that $\Gamma \cong$*

$\mathrm{Cos}(G, G_\alpha, A)$ *and with the action of* G *on* $V\Gamma$ *equivalent to the action of* G *by right multiplication on the right cosets of* G_α *in* G.

In the proof of Theorem 6.1.5 we will use two new results (which we report below), one combinatorial [**129**] (see Theorem 6.2.4), and the other group theoretic [**133**] (see Theorem 6.2.5). For stating Theorem 6.2.4 we need the following definition (see [**127**], and [**128**, Theorem 1.1]). Also we denote the set of functions $\mathbb{N} \to \mathbb{N}$ by $\mathrm{Func}(\mathbb{N})$.

Definition 6.2.3. If G is a quasiprimitive or biquasiprimitive permutation group with socle T^ℓ with T simple, then we call T the *socle factor* of G.

Theorem 6.2.4 (Theorems 4 and 5 in [**129**])**.** *There exists a function* $h : \mathbb{N} \to \mathbb{N}$ *such that, for* Γ *a connected* G*-vertex-transitive graph of valency at most* d *and* α *a vertex of* Γ*, if* G *is quasiprimitive or biquasiprimitive on vertices with socle factor* T*, then either*

(1) $|G_\alpha| \leq h(d)$*, or*

(2) Γ *and* G *uniquely determine two (possibly isomorphic) connected* T*-vertex-transitive graphs* Λ_1 *and* Λ_2 *of valency at most* $d(d-1)$*. Also there is a function* $p : \mathbb{N} \times \mathrm{Func}(\mathbb{N}) \times \mathrm{Func}(\mathbb{N}) \to \mathbb{N}$ *such that if, for each* $i = 1, 2$*,* $|T_{\lambda_i}| \leq g_i(d(d-1))$ *for* $\lambda_i \in V\Lambda_i$ *and for some functions* $g_i : \mathbb{N} \to \mathbb{N}$*, then* $|G_\alpha| \leq p(d, g_1, g_2)$*.*

For the convenience of the reader we restate Theorem 2.1.4, with notation adopted to the present situation. The result was proved simultaneously and independently by Breuillard–Green–Tao [**25**] and Pyber–Szabó [**133**] in 2010.

Theorem 6.2.5 (Product theorem)**.** *Let* T *be a simple group of Lie type of rank* r *and* A *a generating set of* T*. Then either* $T = A^3$ *or* $|A|^{1+\varepsilon(r)} \leq c(r)|A^3|$ *with positive constants* $c(r)$ *and* $\varepsilon(r)$ *depending only on the Lie rank* r *of the simple group* T*.*

In Lemma 6.2.6 we derive from Theorem 6.2.5 the proof of Theorem 6.1.5 in the preliminary case that the group T of automorphisms of the graph Γ is a $\mathrm{BCP}(r)$-group with T simple.

Lemma 6.2.6. *There exists a function* $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ *such that, for* Γ *a connected* T*-vertex-transitive graph of valency at most* d*, if* T *is a* $\mathrm{BCP}(r)$*-group with* T *simple, then a vertex stabiliser in* T *has size at most* $f(r, d)$*.*

Proof By Proposition 6.2.2, we may identify Γ with $\mathrm{Cos}(T, T_\alpha, A)$ and the action of T on Γ with the action of T by right multiplication on the right cosets of T_α in T, for some union A of T_α-double cosets.

Assume that T is abelian or a sporadic simple group or $\mathrm{Alt}(n)$ for $n \leq r$. Clearly $|T_\alpha| \leq \max\{r!, |M|\}$ where M is the Monster sporadic simple group.

Assume that T is a simple group of Lie type. As T is a $\mathrm{BCP}(r)$-group, then, as noted in the introduction, T has Lie rank at most r. Since Γ has valency $d_0 \leq d$ and the neighbours of the vertex T_α are the T_α-right cosets contained in A, we have $|A| = d_0|T_\alpha|$. We claim that $|A^3| \leq d_0^3|T_\alpha|$. Let x be in A^3 and write $x = a_1a_2a_3$ with a_1, a_2 and a_3 in A. By definition of the coset graph

$$(T_\alpha, T_\alpha a_3, T_\alpha a_2a_3, T_\alpha a_1a_2a_3)$$

is a path of length 3 from T_α to $T_\alpha x$ in Γ. Thus every vertex $T_\alpha x$ of Γ with x in A^3 is at distance at most 3 from T_α. Since Γ has valency d_0, the number of vertices at distance at most 3 from α is at most $1 + d_0 + d_0(d_0 - 1) + d_0(d_0-1)^2 \leq d_0^3$ for $d_0 \geq 2$. Since A is a union of right T_α-cosets, we obtain $|A^3| \leq d_0^3|T_\alpha|$, proving the claim.

Since Γ is a connected undirected graph, we have $T = \langle A \rangle$. From Theorem 6.2.5, we obtain that either $T = A^3$ or there exist positive constants, $c(r)$ and $\varepsilon(r)$, depending only on the Lie rank r of the simple group T, such that $|A|^{1+\varepsilon(r)} \leq c(r)|A^3|$. If $T = A^3$, then

$$|V\Gamma| = |T : T_\alpha| = \frac{|A^3|}{|T_\alpha|} \leq d_0^3 \leq d^3$$

and hence the number of vertices of Γ is bounded by a function of d. In particular, $|T_\alpha| \leq d^3!$. If $|A|^{1+\varepsilon(r)} \leq c(r)|A^3|$, then

$$(d_0|T_\alpha|)^{1+\varepsilon(r)} = |A|^{1+\varepsilon(r)} \leq c(r)|A^3| \leq c(r)d_0^3|T_\alpha|$$

which yields

$$|T_\alpha| \leq (c(r)d_0^{2-\varepsilon(r)})^{1/\varepsilon(r)} \leq (c(r)d^{2-\varepsilon(r)})^{1/\varepsilon(r)}$$

and thus in all cases $|T_\alpha| \leq f(r,d)$ where $f(r,d) = \max\{r!, |M|, d^3!, (c(r)d^{2-\varepsilon(r)})^{1/\varepsilon(r)}\}$.

Finally we are ready to prove Theorem 6.1.5.

Proof of Theorem 6.1.5 . Let h and p be the functions in the statement of Theorem 6.2.4 and f the function in the statement of Lemma 6.2.6. Define $g : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by $g(r,d) = \max\{h(d), p(d, f(r, d(d-1)), f(r, d(d-1)))\}$.

Assume that Γ is a connected G-vertex-transitive graph of valency at most d, that G has an intransitive head G/K that is a BCP(r)-group and that G/K is quasiprimitive or biquasiprimitive on $V\Gamma_K$. Let α be a vertex of Γ. We consider the action of G/K on the normal quotient graph Γ_K. Since Γ has valency at most d, we obtain that Γ_K has valency at most d. The socle of G/K is T^ℓ for some simple group T and integer ℓ (by [**127**] and [**128**, Theorem 1.1]). Note that the stabiliser in G of the vertex $B = \alpha^K$ of Γ_K is $G_B = KG_\alpha$, and that $|KG_\alpha/K|$ (the size of the stabiliser in the action of G/K on Γ_K) is equal to $|G_\alpha|/|K_\alpha|$.

Now we apply Theorem 6.2.4 to Γ_K and G/K. If Part (1) of Theorem 6.2.4 holds, then $|G_\alpha|/|K_\alpha| \leq h(d) \leq g(r,d)$. Hence we may assume that Part (2) of Theorem 6.2.4 holds for G/K and Γ_K. This gives two (possibly isomorphic) graphs Λ_i each of valency at most $d(d-1)$, admitting T acting vertex-transitively. Let λ_i be a vertex of Λ_i for $i = 1, 2$.

Since G/K is a BCP(r)-group, so is T. Now, for each $i = 1, 2$, we apply Lemma 6.2.6 to Λ_i and T, and we obtain $|T_{\lambda_i}| \leq f(r, d(d-1))$. Hence from Theorem 6.2.4 (2), we get $|G_\alpha|/|K_\alpha| \leq p(d, f(r, d(d-1)), f(r, d(d-1))) \leq g(r,d)$ and the theorem is proved. □

We conclude this section by giving the example described in Remark 6.1.9.

Example 6.2.7. The graph Γ will be a connected G-vertex-transitive, G-locally primitive graph of valency 9, where $G \leq \mathrm{Sym}(10) \,\mathrm{wr}\, \mathrm{Sym}(2)$ is

quasiprimitive of type PA with socle $\mathrm{Alt}(10)^2$. The naturally defined associated Sym(10)-arc-transitive graph will have valency 9 and the local action induced by Sym(10) is the product action of $\mathrm{Sym}(3) \times \mathrm{Sym}(3)$ which is imprimitive (having two intransitive normal subgroups Sym(3)).

Let $H = \mathrm{Sym}(10)$, $x = (1,2,3)(4,5,6)(7,8,9)$, $y = (1,4,7)(2,5,8)(3,6,9)$, $z = (2,3)(5,6)(8,9)$, $t = (4,7)(5,8)(6,9)$ and $\iota = (1,10)$. Write $K = \langle x, y, z, t\rangle$. Clearly, $K = \langle x, z\rangle \times \langle y, t\rangle \cong \mathrm{Sym}(3) \times \mathrm{Sym}(3)$. Let Δ be the H-set H/K and Λ the coset graph $\mathrm{Cos}(H, K, K\iota K)$. Since ι is an involution, Λ is undirected. Also, as $K \cap K^\iota = \langle z, t\rangle$ and $|K : \langle z, t\rangle| = 9$, the graph Λ has valency 9 and the local action is the natural product action of $\mathrm{Sym}(3) \times \mathrm{Sym}(3)$ of degree 9. Furthermore, it is easy to check that $H = \langle K, \iota\rangle$ and hence the graph Λ is a connected H-arc-transitive vertex-quasiprimitive graph.

Let W be the wreath product $H \operatorname{wr} \mathrm{Sym}(2) = (H \times H) \rtimes \langle\pi\rangle$ where $\pi^2 = 1$ and $(h_1, h_2)^\pi = (h_2, h_1)$ for $h_1, h_2 \in H$. Let T be the socle of H and $N = T^2$ the socle of W. Consider $G = N \rtimes \langle \pi, (\iota, \iota)\rangle$. The group $N\langle(\iota, \iota)\rangle = G \cap (H \times H)$ is the subgroup of index 2 of H^2 normalised by π. Note that each of $\pi, (\iota, \iota)$ has order 2 and $(\iota, \iota)^\pi = (\iota, \iota)$. So G/N is an elementary abelian group of order 4. Also, the projection of $N_G(T \times 1) = N\langle(\iota, \iota)\rangle$ onto the first coordinate of H^2 is the whole of H. Consider the subgroup $L = \langle (x, y), (y, x), (z, t), (t, z), \pi\rangle$ of G. Note that $\langle (x, y), (y, x), (t, z), (z, t)\rangle$ is a diagonal subgroup of $K \times K$ normalised by π. Furthermore $\langle (x, y), (z, t)\rangle^\pi = \langle (x, y)^\pi, (z, t)^\pi\rangle = \langle (y, x), (t, z)\rangle \cong \mathrm{Sym}(3)$ and also $\langle (x, y), (z, t)\rangle$ and $\langle (y, x), (t, z)\rangle$ commute. Therefore we have that $|L| = 72$ and L is isomorphic to $\mathrm{Sym}(3) \operatorname{wr} \mathrm{Sym}(2)$.

Let Ω be the G-set G/L. As $z, t \in \mathrm{Sym}(10) \setminus \mathrm{Alt}(10)$ and $(z, t) \in L$, we have $N\langle(\iota, \iota)\rangle = N\langle(z, t)\rangle \subseteq NL \subseteq G$. Also, since $\pi \in L$, we obtain $G = NL$. Clearly $N \cap L = \langle (x, y), (y, x), (zt, tz)\rangle$ has order 18. In particular, N is the unique minimal normal subgroup of G and is transitive on Ω since $G = NL$. Thus G is quasiprimitive on Ω. Since $N \cap L$ projects to proper nontrivial subgroups of T, the group G has quasiprimitive type PA (see [**127**]).

Note that the group $(K \cap T) \times (K \cap T) = \langle (x, 1), (y, 1), (zt, 1)\rangle \times \langle (1, x), (1, y), (1, zt)\rangle$ is normalised by (z, t) and π and hence by L. Consider $L^* = ((K \cap T) \times (K \cap T))L$ and Σ the system of imprimitivity of Ω corresponding to the overgroup L^* of L. As $z, t \in \mathrm{Sym}(10) \setminus \mathrm{Alt}(10)$, we have $L^* = ((K \cap T) \times (K \cap T))\langle(z, t), \pi\rangle$ and $(K \cap T) \times (K \cap T)$ is normal in L^*. Since $(K \cap T) \times (K \cap T)$ has order $18^2 = 324$ and $|L^* : ((K \cap T) \times (K \cap T))| = 4$, we have $|L^*| = 4 \cdot 324 = 1296$.

Since $L^* \cap N = (K \cap T) \times (K \cap T)$, we have that the N-space Σ is permutation equivalent to the N-space D^2 where $D = T/(K \cap T)$. As the action of T on D is equivalent to its action on Δ, we obtain that the N-space Σ is equivalent to Δ^2 with N acting in product action. Finally since Σ is G-invariant, the action of G on Σ is equivalent to its product action on Δ^2.

Let Γ be the coset graph $\mathrm{Cos}(G, L, L(tz\iota, \iota)L)$. Denote by α the vertex L of Γ and by β the vertex $L(tz\iota, \iota)$ of Γ. Since the involutions z, t and ι of H are pairwise commuting, the element $(tz\iota, \iota)$ is an involution of G interchanging α and β, and hence Γ is undirected. Furthermore, $(z, t)^{(tz\iota,\iota)} = (t, z) \in L$, $(t, z)^{(tz\iota,\iota)} = (t, z) \in L$ and $\pi^{(tz\iota,\iota)} = (tz\iota, \iota)\pi(tz\iota, \iota) = (tz\iota, \iota)(\iota, tz\iota)\pi =$

$(zt, zt)\pi \in L$. Hence $|L : L^{(zt\iota,\iota)}| \leq 9$. A similar computation with (x, y) and (y, x) shows that $|L \cap L^{(zt\iota,\iota)}| \geq 9$ and hence $L \cap L^{(zt\iota,\iota)} = \langle (t, z), (z, t), \pi \rangle$. This gives that $|G_\alpha : G_{\alpha,\beta}| = |L : L \cap L^{(zt\iota,\iota)}| = 9$ and so Γ has valency 9. Moreover, as $L \cap L^{(zt\iota,\iota)}$ is a Sylow 2-subgroup of L and any two distinct Sylow 2-subgroups of L generate the whole of L, we obtain that the action of G_α on $\Gamma(\alpha)$ is primitive. Then Γ is G-locally primitive with the claimed local action.

It is easy to show with the invaluable help of `Magma` [**9**] that $G = \langle L, (zt\iota, \iota) \rangle$, from which it follows that Γ is connected.

6.3. The main examples

In this section we construct connected G-arc-transitive graphs for which G_α can be arbitrarily large compared to the valency. Recall, that by the Thompson-Wielandt theorem [**161**], if G is a primitive group and d is the size of a suborbit then for some prime p the size of $G_\alpha / O_p(G_\alpha)$ is bounded by some function of d (this result is the starting point of the proof of the Sims Conjecture [**28**]). The examples given below show that even the analogue of the Thompson-Wielandt theorem fails for quasiprimitive groups.

As noted in the introduction, these examples also yield exponentially large generating sets of symmetric groups with very small growth.

Let m be an integer with $m \equiv 3 \mod 4$, $r \geq 3$ and $\Omega = \{1, \ldots, r - 1 + mr\}$. Let $\mathcal{P} = \{X_0, \ldots, X_m\}$ be the partition of Ω defined by

$$\begin{aligned} X_j &= \{1 + jr, \ldots, r + jr\}, \qquad \text{for } j = 0, \ldots, m - 1, \\ X_m &= \{1 + mr, \ldots, r - 1 + mr\}. \end{aligned} \tag{6.3.1}$$

In particular, for $j \in \{0, \ldots, m - 1\}$, we have $|X_j| = r$, and $|X_m| = r - 1$. For each $j \in \{0, \ldots, m\}$, write $\mathrm{Alt}(X_j)$ for the alternating group on X_j fixing point-wise $\Omega \setminus X_j$. Note that for $i, j \in \{0, \ldots, m\}$ with $i \neq j$, we have that $\mathrm{Alt}(X_i)$ centralises $\mathrm{Alt}(X_j)$. Set

$$H_0 = \prod_{j=0}^{m} \mathrm{Alt}(X_j) \cong \mathrm{Alt}(r)^m \times \mathrm{Alt}(r - 1). \tag{6.3.2}$$

Define the following permutation of Ω

$$\begin{cases} z + lr \mapsto z + (m - l - 1)r & \text{for } 1 \leq z \leq r, 0 \leq l \leq m - 1, \\ z + mr \mapsto z + mr & \text{for } 1 \leq z \leq r - 1. \end{cases} \tag{6.3.3}$$

Clearly h is an involution of $\mathrm{Sym}(\Omega)$ centralising $\mathrm{Alt}(X_{(m-1)/2})$ and $\mathrm{Alt}(X_m)$. Furthermore, for each $j \in \{0, \ldots, m - 1\}$, we have $X_j^h = X_{m-j-1}$ and hence $\mathrm{Alt}(X_j)^h = \mathrm{Alt}(X_{m-j-1})$. Therefore h normalises H_0. Set

$$H = \langle H_0, h \rangle \cong (\mathrm{Alt}(r)^m \rtimes C_2) \times \mathrm{Alt}(r - 1). \tag{6.3.4}$$

Define the following permutation of Ω

$$(6.3.5)\quad \begin{cases} z & \mapsto & z+mr & \text{for } 1 \le z \le r-1, \\ r & \mapsto & r & \\ z+lr & \mapsto & z+(l+1)r & \text{for } 1 \le z \le r, 1 \le l \le m-1, l \text{ odd}, \\ z+lr & \mapsto & z+(l-1)r & \text{for } 1 \le z \le r, 1 \le l \le m-1, l \text{ even}, \\ z+mr & \mapsto & z & \text{for } 1 \le z \le r-1. \end{cases}$$

Write $X_0' = X_0 \setminus \{r\}$. Clearly a is an involution of $\mathrm{Sym}(\Omega)$ with

$$\begin{array}{rclcll} X_j^a & = & X_{j+1} & \text{and} & X_{j+1}^a = X_j, & \text{for } 1 \le j \le m-1, j \text{ odd}, \\ X_j^a & = & X_{j-1} & \text{and} & X_{j-1}^a = X_j, & \text{for } 1 \le j \le m-1, j \text{ even}, \\ X_0'^a & = & X_m & \text{and} & X_m^a = X_0'. & \end{array}$$

In particular, a normalises the subgroup $\mathrm{Alt}(X_0') \times \mathrm{Alt}(X_1) \times \cdots \times \mathrm{Alt}(X_m)$ of H.

Given a G-arc-transitive graph Γ and v a vertex of Γ, we write $G_1(v)$ for the point-wise stabiliser of the neighbourhood $\Gamma(v)$ of v.

Theorem 6.3.1. *Let m, r, Ω, H and a be as above, $G = \mathrm{Sym}(\Omega)$ and $\Gamma = \mathrm{Cos}(G, H, HaH)$. Then Γ is a connected G-arc-transitive graph of valency $2r$ and for a vertex v, $G_v^{\Gamma(v)} \cong \mathrm{Alt}(r) \wr C_2$ (in its imprimitive action of degree $2r$) and $G_1(v) \cong \mathrm{Alt}(r)^{m-2} \times \mathrm{Alt}(r-1)$.*

For the generating set HaH of $\mathrm{Sym}(\Omega)$ we have $\left|(HaH)^3\right| \le 4r^2|HaH|$.

Proof We prove three claims from which the theorem will follow.

Claim 1. $|H : (H \cap H^a)| = 2r$, the core of $H \cap H^a$ in H is $\mathrm{Alt}(X_1) \times \cdots \times \mathrm{Alt}(X_{m-2}) \times \mathrm{Alt}(X_m)$ and the action of H on the right cosets of $H \cap H^a$ is equivalent to the imprimitive action of $\mathrm{Alt}(r) \wr C_2$ of degree $2r$.

Set $K = \mathrm{Alt}(X_0') \times \mathrm{Alt}(X_1) \times \cdots \times \mathrm{Alt}(X_{m-1}) \times \mathrm{Alt}(X_m)$. Note that K is the stabiliser in H of the point r of Ω. As a normalises K, we have $K \subseteq H \cap H^a$. The orbits of r under H and H^a are

$$r^H = \{1, \ldots, r, 1+(m-1)r, \ldots, r+(m-1)r\} = X_0 \cup X_{m-1}$$

and

$$\begin{array}{rcl} r^{H^a} & = & (r^a)^{Ha} = r^{Ha} = \{1, \ldots, r, 1+(m-1)r, \ldots, r+(m-1)r\}^a \\ & = & \{1+mr, \ldots, r-1+mr, r, 1+(m-2)r, \ldots, r+(m-2)r\} \\ & = & \{r\} \cup X_{m-2} \cup X_m. \end{array}$$

Let $g \in H \cap H^a$. We have $r^g \in r^H \cap r^{H^a} = \{r\}$ and hence g fixes r. Therefore $H \cap H^a \subseteq K$. This yields $K = H \cap H^a$ and $|H : (H \cap H^a)| = 2r$.

Finally, as $H \cap H^a$ is the stabiliser in H of the point r, we have that the core of $H \cap H^a$ in H is the point-wise stabiliser of the set $r^H = X_0 \cup X_{m-1}$, which is clearly $\mathrm{Alt}(X_1) \times \cdots \times \mathrm{Alt}(X_{m-2}) \times \mathrm{Alt}(X_m)$. ∎

It is easy to compute that

$$ha \;:\; \begin{cases} z & \mapsto & z+(m-2)r & \text{for } 1 \le z \le r, \\ z+lr & \mapsto & z+(m-l)r & \text{for } 1 \le z \le r, 1 \le l \le m-1, l \text{ odd}, \\ z+lr & \mapsto & z+(m-2-l)r & \text{for } 1 \le z \le r, 1 \le l \le m-3, l \text{ even}, \\ z+(m-1)r & \mapsto & z+mr & \text{for } 1 \le z \le r-1, \\ r+(m-1)r & \mapsto & r & \\ z+mr & \mapsto & z & \text{for } 1 \le z \le r-1. \end{cases}$$

Using this equation for the permutation ha, we obtain that $\langle ha\rangle$ has r orbits. Specifically, for each $z \in \{1,\dots,r-1\}$, the set $\{z+jr \mid j=0,\dots,m\}$ is an orbit of $\langle ha\rangle$ of size $m+1$. Also, $\{r+jr \mid j=0,\dots,m-1\}$ is an orbit of $\langle ha\rangle$ of size m.

Claim 2. $\mathrm{Alt}(\Omega) \subseteq \langle H, a\rangle$.

Set $L = \langle H, a\rangle$, $g = (ha)^{m+1}$ and $T = \langle H, g\rangle$. Since, for each $j \in \{0,\dots,m-1\}$, the group H is transitive on X_j and g is a cycle with support $\{r, 2r, \dots, mr\}$, we obtain that T is transitive on $\{1,\dots,mr\} = \Omega \setminus X_m$ and fixes point-wise X_m. Since $T \subseteq L$, $X_0'^a = X_m$ and $a \in L$, we obtain that L is transitive on Ω.

As T is transitive on $\Omega\setminus X_m$ and fixes point-wise X_m, using the definition of a, we see that $S = T^a$ acts transitively on $\Omega \setminus X_0'$ and fixes point-wise X_0'. For each $i \in \{2,\dots,r\}$, fix x_i an element of $\mathrm{Alt}(X_0) \subseteq H$ with $r^{x_i} = i$. Clearly, S^{x_i} fixes point-wise $\{1,\dots,i-1,i+1,\dots,r\}$ and acts transitively on $\{i, r+1,\dots,r-1+mr\}$. Therefore $\langle S^{x_i} \mid i=2,\dots,r\rangle \subseteq L$ is transitive on $\{2, l+1,\dots,r-1+mr\} = \Omega \setminus \{1\}$. This shows that L is 2-transitive. Since L contains a 3-cycle, we obtain that $\mathrm{Alt}(\Omega) \subseteq L$. ■

Claim 3. $G = \langle H, a\rangle$.

The permutation h is an involution fixing point-wise $X_{(m-1)/2} \cup X_m$. Therefore h has $(m-1)r/2$ cycles. If r is odd, then (as $m \equiv 3 \mod 4$) we obtain $h \notin \mathrm{Alt}(\Omega)$ and hence (from Claim 2) $G = \langle H, a\rangle$. We may thus assume that r is even. The permutation a is an involution fixing only the point r. Therefore a has $(r-2+mr)/2 = (m-1)r/2 + (r-1)$ cycles. Since $m \equiv 3 \mod 4$ and r is even, we have $a \notin \mathrm{Alt}(\Omega)$ and hence $G = \langle H, a\rangle$. ■

As a is an involution, from Claim 3 we have that Γ is a connected G-arc-transitive graph. Now Claim 1 gives that Γ has valency $2r$ and that, for a vertex v, $G_v^{\Gamma(v)} \cong \mathrm{Alt}(r) \wr C_2$ (in its imprimitive action of degree $2r$) and $G_1(v) = \mathrm{Alt}(X_1) \times \cdots \times \mathrm{Alt}(X_{m-2}) \times \mathrm{Alt}(X_m) \cong \mathrm{Alt}(r)^{m-2} \times \mathrm{Alt}(r-1)$.

Consider now the generating set HaH of $\mathrm{Sym}(\Omega)$. It has size $2r|H|$ and, as in the proof of Lemma 6.2.6, we see that $\left|(HaH)^3\right| \leq (2r)^3|H| = (2r)^2|HaH|$.

Remark 6.3.2. A more elaborate and general version of Theorem 6.3.1 is in [**147**]. Indeed, for any composite positive integer rs (where $r > 1$, $s > 1$) and any transitive permutation groups R of degree r and S of degree s, [**147**, Theorem 2] gives an infinite family of graphs Γ_m (where $m \geq rs$ is odd) of valency rs admitting a G_m-arc-transitive action (where $G_m = \mathrm{Alt}(rsm+r-1)$ or $\mathrm{Sym}(rsm+r-1)$) such that a vertex stabiliser in G_m induces $R \wr S$ (in its imprimitive action) on the neighbourhood of the vertex and with kernel $R^{m-2} \times R_1$ (where R_1 is a point stabiliser in R).

CHAPTER 7

Product Decomposition Conjecture

7.1. Introduction

Our starting point is the following conjecture of Liebeck, Nikolov and Shalev [**99**].

Conjecture 7.1.1. *There exists an absolute constant c such that if G is a finite simple group and S is a subset of G of size at least two, then G is a product of N conjugates of S for some $N \leq c \log |G| / \log |S|$.*

Note that we must have $N \geq \log |G| / \log |S|$ by order considerations, and so the bound above is best possible up to the value of the constant c.

The conjecture is an extension of a deep (and widely applied) theorem of Liebeck and Shalev. Indeed, the main result of [**104**] states that the above conjecture holds when S is a conjugacy class or, more generally, a normal subset (that is, a union of conjugacy classes) of G. In [**99**] Conjecture 7.1.1 is also proved for sets of bounded size.

Somewhat earlier Liebeck, Nikolov and Shalev [**97**] posed the following (still unproved) weaker conjecture.

Conjecture 7.1.2. *There exists an absolute constant c such that if G is a finite simple group and H is any nontrivial subgroup of G, then G is a product of N conjugates of H for some $N \leq c \log |G| / \log |H|$.*

Conjecture 7.1.2 itself represents a dramatic generalization of a host of earlier work on product decompositions of finite simple groups, most of which prove Conjecture 7.1.2 for particular subgroups H. For instance, in [**102**] it is proved that a finite simple group of Lie type in characteristic p is a product of 25 Sylow p-subgroups (see also [**6**] for a recent improvement from 25 to 5).

Further positive evidence for Conjecture 7.1.2 is provided by [**98**], [**106**] and [**112**] (when H is of type SL_n). Certain results of this type are essential to prove that finite simple groups can be made into expanders (see the announcement [**87**]).

The main purpose of this note is to prove Conjecture 7.1.1 for finite simple groups of Lie type of bounded rank. Put another way, we prove a version of Conjecture 7.1.1 in which the constant c depends on the rank of the group G. Our main result follows.

Theorem 7.1.3. *Fix a positive integer r. There exists a constant $c = c(r)$ such that if G is a finite simple group of Lie type of rank r and S is a subset of G of size at least two then G is a product of N conjugates of S for some $N \leq c \log |G| / \log |S|$.*

In [**99**] a weaker bound of the form $N \leq \left(\log|G|/\log|S|\right)^{c(r)}$ is obtained. Also, in [**97**], Theorem 7.1.3 is proved when S is a maximal subgroup of G.

As a byproduct of our proof we obtain two results of independent interest. In these results, and throughout Chapter 7, we denote by S^g the conjugate $g^{-1}Sg$ of a subset S of a group G by an element g of G, and, given a positive integer m, we denote by S^m the product $SS\cdots S$ of m copies of S. There should be no confusion between these two similar notations because the type of the exponent will always be given.

Theorem 7.1.4. *Fix a positive integer r. There exists a positive constant $\varepsilon = \varepsilon(r)$ such that if G is a finite simple group of Lie type of rank r and S is a subset of G then for some g in G we have $|SS^g| \geq |S|^{1+\varepsilon}$ or $S^3 = G$.*

The next theorem is similar, but concerns only normal subsets, in which case we obtain absolute constants.

Theorem 7.1.5. *There exists $\varepsilon > 0$ and a positive integer b such that if G is a finite simple group and S is a normal subset of G then $\left|S^2\right| \geq |S|^{1+\varepsilon}$ or $S^b = G$.*

Theorem 7.1.5 relates to a result of Shalev [**142**, Theorem 7.4], which we strengthen in Section 7.5.

Note that the theorem would not be true were we to consider sets that are not normal. For instance, take S to be a maximal parabolic subgroup in $G = PSL_n(q)$ with index $\frac{q^n-1}{q-1}$. Clearly $S^b = S$ for all positive integers b; on the other hand, for any positive number ε, and any g in G, we have $|SS^g| \leq |G| \leq |S|^{1+\varepsilon}$ once n is large enough. We conclude that neither of the given options can hold in this more general situation.

Theorems 7.1.4 and 7.1.5, and the remarks of the previous paragraph, lead us to make the following conjecture.

Conjecture 7.1.6. *There exists $\varepsilon > 0$ and a positive integer b such that if S is a subset of a finite simple group G then for some g in G we have $|SS^g| \geq |S|^{1+\varepsilon}$ or G is the product of b conjugates of S.*

Note that, by Theorems 7.1.3 and 7.1.4, Conjectures 7.1.1, 7.1.2 and 7.1.6 hold for all exceptional simple groups. Note too that all three conjectures could be phrased in terms of *translates* of the set S, rather than conjugates. This follows from the simple fact that a product of translates of S is equal to a translate of a product of conjugates of S. Similarly a product of conjugates of a translate of S is equal to a translate of a product of conjugates of S, a fact which will be useful in its own right.

It is possible that Conjecture 7.1.6 actually holds with $b = 3$. When $b = 2$ counterexamples are given by large non-real conjugacy classes (see the final section of [**142**] for some related issues). Further counterexamples are given by certain families of maximal subgroups (see for example [**100**, Corollary 2], which states that large enough simple unitary groups of odd dimension cannot be decomposed into the product of two proper subgroups).

We derive Theorems 7.1.3 and 7.1.4 as consequences of the Product theorem (Theorem 2.1.4, restated here in Section 7.2). Theorem 7.1.5 follows from a version of Conjecture 7.1.1 for normal subsets due to Liebeck and

Shalev [**104**] and an extension of Plünnecke's theorem [**160**, Theorem 6.27] to normal subsets of nonabelian groups (see Section 7.4).

In the final section we use a result of Petridis [**121**] to derive an analogue of the classical Doubling lemma, a special case of Plünnecke's theorem. We refer to the new result as the Skew doubling lemma; it can be thought of as a nonabelian version of the classical Doubling lemma. The Skew doubling lemma is applied to prove that Conjecture 7.1.1 implies Conjecture 7.1.6. In the other direction, a standard argument (similar to the proof of Corollary 7.2.8) shows that Conjecture 7.1.6 implies that a simple group G is a product of $(\log|G|/\log|S|)^c$ conjugates of S, a weaker version of Conjecture 7.1.1.

7.2. Proof of Theorem 7.1.4

We begin with a result of Petridis [**121**, Theorem 4.4], which extends work of Helfgott, Ruzsa and Tao [**73, 140, 136, 158**]. It relates to the Doubling lemma for abelian groups, which we return to in Section 7.4.

Lemma 7.2.1. *Let S be a finite subset of a group G. Suppose that there exist positive numbers J and K such that $|S^2| \leq J|S|$ and $|SgS| \leq K|S|$ for each g in S. Then $|S^3| \leq J^7K|S|$.*

Suppose now that G is a finite group, and let $\mathrm{minclass}(G)$ denote the size of the smallest nontrivial conjugacy class in G. Let $\mathrm{minclass}(S,G)$ denote the size of the smallest nontrivial conjugacy class in G that intersects S, and let $\deg_{\mathbb{C}}(G)$ denote the dimension of the smallest nontrivial complex irreducible representation of G.

As observed in [**114**], a result of Gowers [**64**] implies the following.

Proposition 7.2.2. *Let G be a finite group and let $k = \deg_{\mathbb{C}}(G)$. Take $S \subseteq G$ such that $|S| \geq \frac{|G|}{\sqrt[3]{k}}$. Then $G = S^3$.*

Now let $G = G_r(q)$ be a simple group of Lie type of rank r over $\mathbb{F}_q$, the finite field of order q. We need some facts about G. The first result can be deduced, for example, from [**89**, Tables 5.1 and Theorem 5.2.2].

Proposition 7.2.3. *We have $q^r \leq \mathrm{minclass}(G) < |G| \leq q^{8r^2}$.*

Proposition 7.2.4. *Let $k = \deg_{\mathbb{C}}(G)$. Then $|G| < k^{8r^2}$.*

Proof We use the lower bounds on projective representations given by Landazuri and Seitz [**93**], allowing for the slight errors corrected in [**89**, Table 5.3.A]. For $G \neq PSL_2(q)$, we see that $k \geq q$, and so the result follows from Proposition 7.2.3.

Now suppose that $G = PSL_2(q)$; then $|G| < q^3$ and $r = 1$. For $q \geq 5$ and $q \neq 9$, $k = \frac{1}{(2,q-1)}(q-1)$ and it is clear that $k^8 > q^3$. When $q = 4$ we have $k = 2$ and the result follows; likewise when $q = 9$ we have $k = 3$ and the result follows.

The next result was obtained independently in [**68**] and [**148**].

Proposition 7.2.5. *Each finite simple group G is $\frac{3}{2}$-generated; that is, for any nontrivial element g of G there exists h in G such that $\langle g,h\rangle = G$.*

Corollary 7.2.6. *Let G be a finite simple group and let S be a subset of G of size at least two. Then some translate of S generates G.*

Proof Let u and v be distinct elements of S. Since G is $\frac{3}{2}$-generated, there exists x in G such that $\langle vu^{-1}, x\rangle = G$. Therefore the translate $Su^{-1}x$, which contains x and $vu^{-1}x$, generates G.

Next we restate the Product theorem (2.1.4), our primary tool for proving Theorems 7.1.3 and 7.1.4.

Theorem 7.2.7. *[Product theorem] Fix a positive integer r. There exists a positive constant $\eta = \eta(r)$ such that, for G a finite simple group of Lie type of rank r and S a generating set of G, either $S^3 = G$ or $|S^3| \geq |S|^{1+\eta}$.*

We can now prove Theorem 7.1.4.

Proof [Proof of Theorem 7.1.4] Given a positive integer r, let η be the constant from Theorem 7.2.7. It suffices to prove Theorem 7.1.4 for sets S of size larger than some constant $L > 1$ that depends only on η, because if $|S| < L$, and $S^3 \neq G$, then, by the simplicity of G, there is an element g of G such that $|SS^g| \geq |S|+1$, and $|S|+1 \geq |S|^{1+\delta}$, where $\delta = \log(L+1)/\log L - 1$. In particular, we assume that $|S| \geq 8^{\frac{2}{\eta}}$, and we define $\varepsilon = \frac{1}{16}\min\{\eta, \frac{1}{24r^2}\}$.

Since G is $\frac{3}{2}$-generated, there exists an element g of G such that the set $T = S \cup \{g\}$ generates G. We can apply Theorem 7.2.7 to T to conclude that either $T^3 = G$ or $|T^3| \geq |S|^{1+\eta}$.

Now, T^3 is the union of the eight sets SSS, SSg, SgS, gSS, Sgg, gSg, ggS and $\{ggg\}$. Suppose that $|T^3| \geq |S|^{1+\eta}$. By the pigeon-hole principle at least one of the eight sets is larger than $\frac{1}{8}|S|^{1+\eta}$. We assumed earlier that $|S| \geq 8^{\frac{2}{\eta}}$, from which it follows that $\frac{1}{8}|S|^{1+\eta} > |S|^{1+\frac{\eta}{2}}$. Therefore one of the first seven of the eight sets is larger than $|S|^{1+\frac{\eta}{2}}$. All of these seven sets except SSS are equal to a translate of the product of one or two conjugates of S, so if any of these have size at least $|S|^{1+\frac{\eta}{2}}$ then $|SS^h| \geq |S|^{1+\frac{\eta}{2}}$ for some element h of G. If, on the other hand, $|SSS| \geq |S|^{1+\frac{\eta}{2}}$, then Lemma 7.2.1 (with $J = K = |S|^{\frac{\eta}{16}}$) implies that there is an element h of $S \cup \{1\}$ with $|SS^h| \geq |S|^{1+\frac{\eta}{16}}$. Thus in both cases there is an element h with $|SS^h| \geq |S|^{1+\varepsilon}$.

The remaining possibility is that $T^3 = G$. If $S^3 \neq G$ then Proposition 7.2.2 implies that $|S| \leq |G|/\sqrt[3]{k}$ where $k = \deg_{\mathbb{C}}(G)$. But Proposition 7.2.4 gives that $|S| \leq |G|^{1-\frac{1}{24r^2}}$, and this implies, in particular, that $|T^3| = |G| \geq |S|^{1+\frac{1}{24r^2}}$. The argument of the previous paragraph applies again, to give $|SS^h| \geq |S|^{1+\varepsilon}$ for some element h.

Note that we can immediately deduce the following result of [**97**] (which we will use later).

Corollary 7.2.8. *Fix a positive integer r. There exists a constant d such that if G is a finite simple group of Lie type of rank r and S is a subset of G of size at least two then G is a product of N conjugates of S for some $N \leq 3(\log|G|/\log|S|)^d$.*

Proof Let ε be the constant from Theorem 7.1.4, and define $d = \log_{1+\varepsilon} 2$. Let M be the integer part of $\log_{1+\varepsilon} \frac{\log|G|}{\log|S|}$. Theorem 7.1.4 implies

that G is the product of $3 \cdot 2^M$ conjugates of S, and

$$3 \cdot 2^M \leq 3 \left(\frac{\log |G|}{\log |S|} \right)^d .$$

The results in this section motivate a common generalisation of the Product theorem, and Conjecture 7.1.6, for groups of Lie type.

Conjecture 7.2.9. *There exists $\varepsilon > 0$ and a positive integer b such that the following statement holds. For each integer r there is a positive integer $c(r)$ such that if G is a finite simple group of Lie type of rank r and S a generating set of G, then either $|SS^g| \geq |S|^{1+\varepsilon}$ for some $g \in S^{c(r)}$, or else G is the product of b conjugates $S^{g_1}, \ldots, S^{g_b}$, where $g_1, \ldots, g_b \in S^{c(r)}$.*

It would be interesting to prove Conjecture 7.1.6 in the case when S is a subgroup of G. A rather general qualitative result in this direction was obtained by Bergman and Lenstra [**8**]. They show that if H is a subgroup of a group G satisfying $\left|HH^g\right| \leq K|H|$ for all g in G, then H is "close to" some normal subgroup N of G, in the sense that $\left|H : H \cap N\right|$ and $\left|N : H \cap N\right|$ are both bounded in terms of K.

7.3. Proof of Theorem 7.1.3

Given an element g of a group G we define

$$g^G = \{g^h \,:\, h \in G\},$$

and, for a subset Z of G,

$$Z^G = \{Z^h \,:\, h \in G\}.$$

We begin the proof of Theorem 7.1.3 with a simple combinatorial lemma, which enables us to deal with "small" sets.

Lemma 7.3.1. *Let S be a subset of a finite group G. There exist a positive integer m and m conjugates of S such that their product X satisfies*

$$|X| = |S|^m \geq \frac{\sqrt{\text{minclass}(SS^{-1}, G)}}{|S|} \geq \frac{\sqrt{\text{minclass}(G)}}{|S|}.$$

Proof Define $X_1 = S$ and, if possible, choose an element g of G such that $X_1^{-1}X_1 \cap gSS^{-1}g^{-1} = \{1\}$. Define $X_2 = X_1 g S g^{-1}$. Notice that if $x_L, x_R \in X_1, s_L, s_R \in S$, and $x_L g s_L g^{-1} = x_R g s_R g^{-1}$, then $x_R^{-1}x_L = g s_R s_L^{-1} g^{-1}$. Hence $x_R^{-1}x_L \in X_1^{-1}X_1 \cap gSS^{-1}g^{-1}$, and so $x_L = x_R$ and $s_L = s_R$. It follows that $|X_2| = |X_1||S|$. Now repeat this process with X_2 replacing X_1, and so on.

The process terminates with a set X of size $|S|^m$, which is a product of m conjugates of S, and such that $|X^{-1}X \cap gSS^{-1}g^{-1}| \geq 2$ for all g in G.

Let T be a set of smallest possible size that intersects every conjugate of $Z = SS^{-1}$ nontrivially, and write $t = |T|$. Let $n = |G : N_G(Z)|$, the number of G-conjugates of Z. By the pigeonhole principle there exists an element g of Z that lies in at least $\frac{n}{t}$ different conjugates of Z. Let us count the set

$$\Omega = \left\{(g', Z') \in g^G \times Z^G \,\middle|\, g' \in Z'\right\}$$

in two different ways.

First, since every conjugate of g lies in the same number of conjugates of Z, we know that $|g^G|\frac{n}{t} \leq |\Omega|$. On the other hand it is clear that $|\Omega| \leq n|Z|$. Putting these together we obtain that $|g^G|\frac{n}{t} \leq n|Z|$. Therefore

$$t \geq \frac{|g^G|}{|Z|} \geq \frac{\operatorname{minclass}(SS^{-1}, G)}{|S|^2}$$

and using $|X|^2 \geq |X^{-1}X| \geq t$ our statement follows.

Remark 7.3.2. Lemma 7.3.1 and Proposition 7.2.3 imply that if G is a simple group of Lie type of rank r and S a subset of size less that $q^{r/4}$ then we have $\left|SS^g\right| = |S|^2$ for some g in G.

We are now ready to prove Theorem 7.1.3.

Proof [Proof of Theorem 7.1.3] As observed above, a product of conjugates of a translate of S is equal to the translate of a product of conjugates of S. By Corollary 7.2.6, a translate of S generates G. Therefore we assume that S generates G.

Suppose that $|S| \geq \left|\operatorname{minclass}(G)\right|^{1/4}$; then $|G| < |S|^{32r}$ by Proposition 7.2.3. Now Corollary 7.2.8 implies that G is a product of fewer than $3(32r)^d$ conjugates of S. The theorem holds in this case with $c = 3(32r)^d$.

Suppose instead that $|S| < |\operatorname{minclass}(G)|^{1/4}$. By Lemma 7.3.1 we can choose conjugates $S_1, \ldots, S_m$ of S such that the set $X = S_1 \cdots S_m$ satisfies $|X| = |S|^m$ and

$$|X| \geq \frac{\sqrt{|\operatorname{minclass}(G)|}}{|S|} \geq \left|\operatorname{minclass}(G)\right|^{1/4}.$$

It follows from the first part of the proof that G is a product of fewer than $c \log |G| / \log |X|$ conjugates of X. Therefore G is a product of fewer than $mc \log |G| / \log |X|$ conjugates of S and, since $\log |X| = m \log |S|$, the result follows.

7.4. Plünnecke-Ruzsa estimates for nonabelian groups

The following basic result in additive combinatorics is due to Plünnecke [**123, 124**] (see also [**160**, Section 6.5]).

Theorem 7.4.1. *Let A and B be finite sets in an abelian group G and suppose that $|AB| \leq K|A|$ where K is a positive number. Then for any positive integer m there exists a nonempty subset X of A such that*

$$|XB^m| \leq K^m|X|.$$

In particular, $|B^2| \leq K|B|$ implies that $|B^m| \leq K^m|B|$ for $m = 1, 2, \ldots$.

The last statement ("In particular...") is called the Doubling lemma; it does not hold for nonabelian groups, however, as we saw in Lemma 7.2.1, there are useful analogues in this context due to Helfgott, Petridis, Ruzsa and Tao [**73, 121, 140, 136, 158**]. Petridis also proved the following lemma [**121**, Proposition 2.1].

Lemma 7.4.2. *Let X and B be finite sets in a group. Suppose that*

$$\frac{|XB|}{|X|} \leq \frac{|ZB|}{|Z|}$$

for all $Z \subseteq X$. Then, for all finite sets C,

$$|CXB| \leq \frac{|CX||XB|}{|X|}.$$

Using this lemma we can extend Plünnecke's theorem to normal subsets of nonabelian groups. The statement and proof mimic [**121**, Theorem 3.1], which is a stronger version of Theorem 7.4.1.

Theorem 7.4.3. *Let A and B be finite sets in a group G with B normal in G. Suppose that $|AB| \leq K|A|$ for some positive number K. Then there exists a nonempty subset X of A such that*

$$|XB^m| \leq K^m|X|$$

for $m = 1, 2, \ldots$. In particular, $|B^2| \leq K|B|$ implies that $|B^m| \leq K^m|B|$ for $m = 1, 2, \ldots$

Proof We proceed by induction on m. First choose $X \subseteq A$ such that

$$\frac{|XB|}{|X|} \leq \frac{|ZB|}{|Z|}$$

for all $Z \subseteq A$. Then

$$|XB| \leq |X|\frac{|AB|}{|A|} \leq K|X|,$$

so the result is true for $m = 1$.

Now suppose that $|XB^m| \leq K^m|X|$ for some positive integer m. Normality of B implies that $|XB^{m+1}| = |B^mXB|$, and then Lemma 7.4.2 gives

$$|XB^{m+1}| = |B^mXB| \leq \frac{|B^mX||XB|}{|X|} \leq K^{m+1}|X|.$$

This verifies the inductive step, and completes the proof of the theorem.

Following an argument of Petridis (see the proof of [**121**, Theorem 1.2]) we observe that the Plünnecke-Ruzsa estimates [**160**, Corollary 6.29] can also be generalised using Theorem 7.4.3.

Corollary 7.4.4. *Suppose that A and B are subsets of a group G, with B normal in G, and $|AB| \leq K|A|$. Then*

$$|B^mB^{-n}| \leq K^{m+n}|A|$$

for all positive integers m and n.

Theorem 7.4.3 suggests that certain techniques in additive combinatorics concerning subsets of abelian groups can be applied to normal subsets of nonabelian groups. The next example – which is a consequence of Plünnecke's theorem, and generalises [**140**, Corollary 2.4] – supports this suggestion.

Theorem 7.4.5. *Let A and B be subsets of a group G with B normal in G, and suppose that $|AB^j| \leq K|A|$ for some positive integer j. If $m \geq j$ then*

$$|B^m| \leq K^{\frac{m}{j}}|A|.$$

Proof [Sketch of proof] We use the notation of [**160**, Section 6.5]. Construct the m-tuple of directed bipartite graphs

$$(G_{A,B}, G_{AB,B}, \ldots, G_{AB^{m-1},B}).$$

This m-tuple is a Plünnecke graph. Now Plünnecke's theorem [**160**, Theorem 6.27] yields the result immediately.

7.5. Proof of Theorem 7.1.5

In this section we prove Theorem 7.1.5 and generalise some related results of Shalev. We will need the following theorem of Liebeck and Shalev [**104**].

Theorem 7.5.1. *There exists an absolute positive constant a such that, if G is a finite simple group and S is a nontrivial normal subset of G, then $G = S^m$, where $m \leq a\frac{\log|G|}{\log|S|}$.*

Proof [Proof of Theorem 7.1.5] Let a be the absolute constant from Theorem 7.5.1. Choose a positive integer b larger than $2a$. Suppose first that $|S| \geq \sqrt{|G|}$. Then Theorem 7.5.1 implies that $G = S^m$ where

$$m \leq \frac{a\log|G|}{\log|S|} \leq 2a \leq b,$$

and hence $S^b = G$.

Now suppose that $|S| \leq \sqrt{|G|}$. Then

$$\frac{\log|S|}{a\log|G|} \geq \frac{\log|S|}{2a(\log|G| - \log|S|)} = \frac{\log|S|}{2a(\log(|G|/|S|)}.$$

Theorem 7.5.1 implies, once again, that for some $m \leq \frac{a\log|G|}{\log|S|}$ we have $G = S^m$. Hence, applying Theorem 7.4.3 to the normal subset S, we see that

$$\frac{|S^2|}{|S|} \geq \left(\frac{|S^m|}{|S|}\right)^{\frac{1}{m}} \geq \left(\frac{|G|}{|S|}\right)^{\frac{\log|S|}{a\log|G|}} \geq \left(\frac{|G|}{|S|}\right)^{\frac{\log|S|}{2a(\log(|G|/|S|)}} = |S|^{\frac{1}{2a}} \geq |S|^{\frac{1}{b}},$$

and this completes the proof.

The next result is a strengthening of [**142**, Theorem 7.4].

Proposition 7.5.2. *For every $\delta > 0$ there exists $\varepsilon > 0$ such that for any finite simple group G and subsets A and B of G with B normal in G and $|A| \leq |G|^{1-\delta}$ we have*

$$|AB| \geq |A||B|^{\varepsilon}.$$

Proof We assume that A is nonempty and B is nontrivial, otherwise the result is immediate.

By Theorem 7.5.1, $G = B^m$, where $m \leq a\frac{\log|G|}{\log|B|}$. Let $K = |AB|/|A|$. Then, by Theorem 7.4.3, there is a nonempty subset X of A such that $|XB^m| \leq K^m|X|$. It follows that

$$|G| = |B^m| = |XB^m| \leq K^m|X| \leq K^m|A|.$$

Since $|A| \leq |G|^{1-\delta}$ and $m \leq a\frac{\log|G|}{\log|B|}$ we can rearrange this inequality to give

$$|G|^{\delta} \leq K^{a\frac{\log|G|}{\log|B|}}.$$

This is equivalent to $|B|^{\frac{\delta}{a}} \leq K$, which, with $\varepsilon = \frac{\delta}{a}$, is the required result.

Proposition 7.5.2 constitutes the expansion result for B^2 that was partially proven in [**142**, Proposition 10.4]. Furthermore it goes some way towards a proof of [**142**, Conjecture 10.3] although what remains is the more difficult part of the conjecture.

We can strengthen [**142**, Proposition 10.4] in a different direction as follows.

Proposition 7.5.3. *For every $\delta > 0$ and positive integer r there exists $\varepsilon > 0$ such that for any finite simple group G of Lie type of rank r and any set $S \subseteq G$ such that $|S| \leq |G|^{1-\delta}$, there exists g in G such that*

$$|SS^g| \geq |S|^{1+\varepsilon}.$$

Proof Given $\delta > 0$ and a positive integer r, let ε be the positive constant from Theorem 7.1.4. Now choose any subset S of G such that $|S| \leq |G|^{1-\delta}$. According to Theorem 7.1.4, either $|SS^g| \geq |S|^{1+\varepsilon}$ or else $S^3 = G$. In the former case the result is proven. In the latter case we apply Lemma 7.2.1 with $J = K = (|S^3|/|S|)^{1/10}$ to deduce the existence of an element g of G with $|SgS| > K|S|$. Then, using $S^3 = G$ and $|G| \geq |S|^{1+\delta}$, it follows that

$$|SgS| > \left(\frac{|S^3|}{|S|}\right)^{\frac{1}{10}} |S| \geq |S|^{1+\frac{\delta}{10}}.$$

Provided that ε is chosen to be smaller than $\frac{\delta}{10}$, the inequality $|SS^g| \geq |S|^{1+\varepsilon}$ is again satisfied.

7.6. The Skew doubling lemma

The next result is another analogue of the Doubling lemma for non-abelian groups, which we call the *Skew doubling lemma*.

Lemma 7.6.1 (Skew doubling lemma)**.** *If S is a finite subset of a group G such that, for some positive number K, $|SS^g| \leq K|S|$ for every conjugate S^g of S, then*

$$|S_1 \cdots S_m| \leq K^{14(m-1)}|S|$$

for $m = 1, 2, \ldots$, where each of $S_1, \ldots, S_m$ is any conjugate of either S or S^{-1}.

To prove Lemma 7.6.1 we will use Lemma 7.2.1 and the following result, Ruzsa's triangle inequality [**139**] (see also [**160**, Section 2.3]).

Lemma 7.6.2. *Let U, V and W be finite subsets of a group G. Then*

$$\frac{|VW^{-1}|}{|U|} \leq \frac{|UV^{-1}|}{|U|}\frac{|UW^{-1}|}{|U|}.$$

First we prove a special case of Lemma 7.6.1.

Lemma 7.6.3. *Let S be a finite subset of a group G. Suppose that K is a positive number such that $|SS^g| \leq K|S|$ for each g in G. Then $|S_1S_2S_3| \leq K^{14}|S|$, where each of S_1, S_2 and S_3 is any conjugate of either S or S^{-1}.*

Proof Choose elements a and b of G. We can apply Lemma 7.2.1 with $J = K$ to obtain

$$|S^3| \leq K^8|S|.$$

Using this inequality and Lemma 7.6.2 (with $U = S^{-1}$, $V = SS$ and $W = S$) we obtain

$$\frac{|SSS^{-1}|}{|S|} \leq \frac{|S^{-1}S^{-1}S^{-1}|}{|S|}\frac{|S^{-1}S^{-1}|}{|S|} = \frac{|S^3|}{|S|}\frac{|S^2|}{|S|} \leq K^9.$$

Using this inequality and Lemma 7.6.2 (with $U = S$, $V = S^{-1}$ and $W = SS^{-1}$) we obtain

$$\frac{|S^{-1}SS^{-1}|}{|S|} \leq \frac{|SS|}{|S|}\frac{|SSS^{-1}|}{|S|} \leq K^{10}.$$

Using this inequality and Lemma 7.6.2 (with $U = S^{-1}$, $V = SS^{-1}$ and $W = Sa$) we obtain

$$\frac{|SS^{-1}a^{-1}S^{-1}|}{|S|} \leq \frac{|S^{-1}SS^{-1}|}{|S|}\frac{|S^{-1}a^{-1}S^{-1}|}{|S|} \leq K^{11}.$$

Using this inequality and Lemma 7.6.2 (with $U = S$, $V = SaS$ and $W = S^{-1}b^{-1}$) we obtain

$$\frac{|SaSbS|}{|S|} \leq \frac{|SS^{-1}a^{-1}S^{-1}|}{|S|}\frac{|SbS|}{|S|} \leq K^{12}. \tag{7.6.1}$$

Using this inequality and Lemma 7.6.2 (with $U = S$, $V = S^{-1}$ and $W = S^{-1}b^{-1}S^{-1}a^{-1}$) we obtain

$$\frac{|S^{-1}aSbS|}{|S|} \leq \frac{|SS|}{|S|}\frac{|SaSbS|}{|S|} \leq K^{13}. \tag{7.6.2}$$

Finally, using this inequality and Lemma 7.6.2 (with $U = S^{-1}$, $V = S^{-1}aSb$ and $W = S$) we obtain

$$\frac{|S^{-1}aSbS^{-1}|}{|S|} \leq \frac{|S^{-1}b^{-1}S^{-1}a^{-1}S|}{|S^{-1}|}\frac{|S^{-1}S^{-1}|}{|S^{-1}|} = \frac{|S^{-1}aSbS|}{|S|}\frac{|SS|}{|S|} \leq K^{14}. \tag{7.6.3}$$

Equations (7.6.1), (7.6.2) and (7.6.3) imply that, given any conjugates S_1, S_2 and S_3 of either S or S^{-1}, we have $|S_1S_2S_3|/|S| \leq K^{14}$, as required.

We need the following proposition.

Proposition 7.6.4. *If A and B are finite subsets of a group G such that, for some positive number K, $|BB^g| \leq K|B|$ for every conjugate B^g of B, then*

$$|AB_1B_2| \leq K^{14}|AB_3|,$$

where each of B_1, B_2 and B_3 is any conjugate of B or B^{-1}.

Proof By Lemma 7.6.3 we have

$$\frac{|B_3^{-1}B_1B_2|}{|B_3|} \leq K^{14},$$

where each of B_1, B_2 and B_3 is any conjugate of B or B^{-1}. Applying Lemma 7.6.2 with $U = B_3^{-1}$, $V = A$ and $W = B_2^{-1}B_1^{-1}$ we obtain

$$\frac{|AB_1B_2|}{|AB_3|} = \frac{|AB_1B_2|}{|B_3^{-1}A^{-1}|} \leq \frac{|B_3^{-1}B_1B_2|}{|B_3|} \leq K^{14},$$

as required.

We can finally prove Lemma 7.6.1.

Proof [Proof of the Skew doubling lemma] The result holds trivially when $m = 1$ and $m = 2$. Suppose that $m \geq 3$. Apply Proposition 7.6.4 with $B = S$, $A = S_1 \cdots S_{n-2}$, $B_1 = B_3 = S_{n-1}$ and $B_2 = S_n$ to see that

$$\frac{|S_1 \cdots S_n|}{|S_1 \cdots S_{n-1}|} \leq K^{14}$$

for $n = 3, 4, \ldots, m$. It follows that

$$\begin{aligned}\frac{|S_1 \cdots S_m|}{|S|} &= \left(\frac{|S_1 \cdots S_m|}{|S_1 \cdots S_{m-1}|}\right)\left(\frac{|S_1 \cdots S_{m-1}|}{|S_1 \cdots S_{m-2}|}\right) \cdots \left(\frac{|S_1S_2S_3|}{|S_1S_2|}\right)\left(\frac{|S_1S_2|}{|S_1|}\right) \\ &\leq (K^{14})^{m-2}K \\ &\leq K^{14(m-1)},\end{aligned}$$

as required.

Using the Skew doubling lemma we can derive Conjecture 7.1.6 from Conjecture 7.1.1. The proof is similar to the proof of Theorem 7.1.5.

Proof [Proof that Conjecture 7.1.1 implies Conjecture 7.1.6] Let c be the absolute constant from Conjecture 7.1.1. We define b to be a positive integer greater than $2c$, and $\varepsilon = 1/(28c)$. Suppose first that $|S| \geq \sqrt{|G|}$. Then Conjecture 7.1.1 implies that $G = S_1 \cdots S_N$, for conjugates $S_1, \ldots, S_N$ of S, where

$$N \leq \frac{c \log |G|}{\log |S|} \leq 2c < b,$$

and hence G is certainly the product of b conjugates of S.

Now suppose that $|S| \leq \sqrt{|G|}$. Then

$$\frac{\log |G| - \log |S|}{c \log |G| - \log |S|} \geq \frac{\log |G| - \log |S|}{c \log |G|} \geq \frac{1}{2c}.$$

In particular observe that

$$c \log |G| - \log |S| \leq 2c(\log |G| - \log |S|) = 2c \log(|G|/|S|).$$

Conjecture 7.1.1 implies, once again, that for some $N \leq \frac{c \log |G|}{\log |S|}$ we have $G = S_1 \cdots S_N$, for conjugates $S_1, \ldots, S_N$ of S. Using the Skew doubling lemma, Lemma 7.6.1, we see that there is an element g of G for which

$$\frac{|SS^g|}{|S|} \geq \left(\frac{|S_1 \ldots S_N|}{|S|}\right)^{\frac{1}{14(N-1)}} \geq \left(\frac{|G|}{|S|}\right)^{\frac{\log |S|}{14(c \log |G| - \log |S|)}} \geq \left(\frac{|G|}{|S|}\right)^{\frac{\log |S|}{28c(\log(|G|/|S|))}} \geq |S|^{\frac{1}{28c}},$$

and this completes the proof.

dc_650_12

Index

dc_650_12

List of Symbols

$\langle A \rangle$	the subgroup generated by the subset A, page 65
$\langle \mathcal{A}, \oplus \rangle$	an Abelian topological group, page 146
α	in Chapter 2, an ordered finite subset of the affine space $\overline{\mathbb{F}}^m$, page 60
α	usually a finite subset in a group, page 5
$\overline{\alpha}$	in Chapter 5, the graph of the function α, page 148
$\langle \alpha \rangle$	the subgroup generated by the subset α, page 65
α^K	for permutation groups, the K-orbit of the vertex α, page 162
α^n	in a group, the set of all n-term products formed from the elements of α, page 5
$\tilde{\alpha}$	the pojection of a subset α of a group into a quotient group, page 117
$\overline{A}$	the closure of the subset A, page 55
$\mathrm{Alt}(n)$	alternating group on n elements, page 168
$\mathrm{Alt}(X)$	alternating group on the finite set X, page 168
$\mathrm{Aut}(\Gamma)$	the automorphism group of the graph Γ, page 161
$\mathrm{Aut}(L)$	the automorphism group of the group L, page 108
b	a fixed constant used in Definition 1.2.1, Definition 1.2.12 and in the Convention of Section 1.4, page 25
$\overline{\beta}$	in Chapter 5, the graph of the function β, page 148
β^K	for permutation groups, the K-orbit of the vertex β, page 162
$\overline{B}$	the closure of the subset B, page 55
$\mathrm{BCP}(r)$	the class of finite groups G which have no section H/K isomorphic to the alternating group $\mathrm{Alt}(r+1)$, page 161
$\mathrm{bd}()$	the boundary of a set, page 134
$\mathrm{cdim}_b(P, Q)$	combinatorial dimension of a geometric configuration, page 26
$\mathrm{cdim}_b(S, T)$	combinatorial dimension of bipartate graph, page 25
$\mathcal{C}_G(A)$	the centraliser of the subset A in the group G, page 65
$\mathcal{C}_G(A)^0$	the connected centraliser of A in G, it is just the unit component of the centraliser $\mathcal{C}_G(A)$, page 65
Δ	in Chapter 2, upper bound on degrees, page 54
$\deg_{\mathbb{C}}(G)$	the minimum degree of a non-trivial complex representation of the group G, page 117
$\deg(F)$	degree of the multi-valued function F, page 12
$\deg(f)$	degree of a morphism f, page 56
$\deg(X)$	the degree of the algebraic set X, page 55
$\mathrm{diam}(X)$	diameter of a graph X, page 49
$\dim(X)$	dimension of X, page 31
$\mathcal{E}$	in Chapter 4, $\mathcal{E}$ denotes a partial envelope of a family of curves, page 134

η usually a small positive number in the exponent, like in $n^{2-\eta}$, page 24

ε in Chapter 2, the error-margin we allow in the exponents, page 55

(ε, M, δ) a symbol used as in "(ε, M, δ)-spreading", page 70

$(F : A \to B)$ a generalised multi–function F from A to B, page 35

$(F : A \to B)$ a multi–function F from A to B, page 34

$(F_\gamma : V \to V, \gamma \in \Gamma)$ standard family of multi–functions corresponding to the group Γ acting on the variety V, page 36

$(F_t : A \to B, t \in T)$ family of multi–functions from A to B parametrised by the irreducible variety T, page 34

$\mathbb{F}$ an arbitrary field, page 6

$\overline{\mathbb{F}}$ the algebraic closure of the field $\mathbb{F}$, page 6

$\overline{\mathbb{F}}^m$ affine space of dimension m over the algebraically closed field $\overline{\mathbb{F}}$, page 55

$\mathbb{F}_p$ the field with p elements, for some prime p, i.e. the ring of remainder-classes modulo p, page 6

$\overline{\mathbb{F}_p}$ the algebraic closure of the field $\mathbb{F}_p$, page 6

$\mathbb{F}_q$ the field with q elements, for some prime power p, page 6

$Frob_q$ Frobenius morphism, the q-th power map, page 87

Γ often denotes a family of continuous curves in the plane, page 127

G^0 unit component of the algebraic group G, page 65

G_α the stabiliser of the permutation group G at the point α, page 161

$\overline{G}$ the closure of the set G, page 134

$G(\mathbb{F})$ the subgroup in G of those elements whose matrix entries belong to the field $\mathbb{F}$, page 87

$G(\mathbb{F}_q)$ the subgroup in G of those elements whose matrix entries belong to the field $\mathbb{F}_q$, page 87

$[G, G]$ commutator subgroup of the group G, page 65

Γ in Chapter 6, Γ is a graph, and the group G acts on Γ, page 161

Γ often denotes a subgroup with nice properties, i.e., a virtually soluble subgroup, page 52

Γ_f the graph of the function f, page 56

γ^n in a group, the set of all n-term products formed from the elements of γ, page 102

$\overline{\gamma}$ in Chapter 5, the graph of the function γ, page 148

$\Gamma(G, S)$ the Cayley graph of the group G corresponding to the generating set S, page 49

Γ_K for permutation group acting on Γ, the fixpoint set of K, page 162

$\gamma^{(t)}$ a member of a family of plane curves, page 130

$GL(n, \mathbb{F})$ the group of invertible $n \times n$ matrices, whith entries taken from an arbitrary field $\mathbb{F}$, page 6

$GL(n, p)$ the group of invertible $n \times n$ matrices, whith entries taken from the field $\mathbb{F}_p$, for some prime p, page 6

$GL(n, q)$ the group of invertible $n \times n$ matrices, whith entries taken from the field $\mathbb{F}_q$, for some prime power q, page 6

$GL(n, R)$ the group of $n \times n$ invertible $n \times n$ matrices, whith entries taken from an arbitrary ring R, for example, $R = \mathbb{Z}$ or $R = BZ/m\mathbb{Z}$, page 6

$\mathcal{N}_G(A)$ the normaliser of the subset A in the group G , page 65
$\mathcal{O}(\dots)$ usual big–Oh expression, page 25
$O_p(G)$ the maximal normal p-subgroup of a finite group G, page 98
$\mathcal{P}$ often denotes the parameter space of a family, page 31
P often denotes a perfect group, page 99
$\prod^m \alpha$ m-fold direct product of the subset α with itself, page 65
$\prod^m G$ m-fold direct product of the group G with itself, page 65
$PSL(n,p)$ the quotiont group of $SL(n,p)$ by its centre, q is a prime, page 115
$PSL(n,q)$ the quotiont group of $SL(n,q)$ by its centre, q is a prime power, page 115
q_σ parameter used to calculate the number of elements in finite simple groups of Lie type, page 87
$\mathrm{Reg}(\Gamma)$ the set of regular points of the algebraic curve Γ, page 147
S_F surface in $\mathbb{R}^3$ defined by the three-variate polynomial equation $F = 0$, page 130
$SL(n,\mathbb{F})$ the group of $n \times n$ matrices of determinant 1, whith entries taken from an arbitrary field $\mathbb{F}$, page 6
$SL(n,p)$ the group of $n \times n$ matrices of determinant 1, whith entries taken from the field $\mathbb{F}_p$, for some prime p, page 6
$SL(n,q)$ the group of $n \times n$ matrices of determinant 1, whith entries taken from the field $\mathbb{F}_q$, for some prime power q, page 6
$SL(n,R)$ the group of $n \times n$ matrices of determinant 1, whith entries taken from an arbitrary ring R, for example, $R = \mathbb{Z}$ or $R = BZ/m\mathbb{Z}$, page 6
$Sol(G)$ the soluble radical of the group G, page 98
$\tau_{\underline{g}}$ a morphism from $\prod^m G$ to G defined as the product of certain conjugates, page 65
$\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n)$ the maximum number of triple points of $n+n+n$ members from the three families $\Gamma_1, \Gamma_2, \Gamma_3$ of plane curves, page 127
$\mathcal{T}_\Gamma(n)$ the maximum number of triple points of n members from the family Γ of plane curves, page 127
$\mathrm{tr}(\alpha, \Sigma)$ in groups, the trace of the subset α in the section Σ, page 121
$V\Gamma$ the vertex set of the graph Γ, page 161
$\overline{X}$ the closure of the subset X, page 55
X^{gen} the set of all CC-generators in $\prod^{\dim(G)} X$, page 77
X^{nongen} the complementer set of X^{gen}, page 77
$\overline{Y}$ the closure of the subset Y, page 55
$\mathcal{Z}(G)$ centre of the group G, page 65

Bibliography

1. D. Aldous, *On the Markov chain simulation method for uniform combinatorial distributions and simulated annealing*, Probab. Eng. Inform. Sci **1** (1987), 33–46.
2. M Aschbacher, *Finite group theory*, Cambridge Univ. Press, 1986.
3. L. Babai, *Local expansion of vertex-transitive graphs and random generation in finite groups*, Proc. 23rd ACM Symp. on Theoretical Computing (STOC), ACM, New York, 1991, pp. 164–174.
4. L. Babai, P. J. Cameron, and P. P. Pálfy, *On the orders of primitive groups with restricted nonabelian composition factors*, J. Algebra **79** (1982), 161–168.
5. L. Babai, W. M. Kantor, and A. Lubotzky, *Small-diameter Cayley graphs for finite simple groups*, Europ. J. Combinatorics **10** (1989), 507–522.
6. L. Babai, N. Nikolov, and L. Pyber, *Product growth and mixing in finite groups*, Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms (New York), ACM, 2008, pp. 248–257.
7. L. Babai and Á. Seress, *On the diameter of permutation groups*, European J. Comb. **13** (1992), 231–243.
8. G. M. Bergman and H. W. Lenstra, Jr., *Subgroups close to normal subgroups*, J. Algebra **127** (1989), no. 1, 80–97.
9. W. Bosma, J. Cannon, and C. Playoust, *The Magma algebra system. I. The user language*, J. Symbolic Comput. **24** (1997), 235–265.
10. J. Bourgain, *A modular Szemeredi-Trotter theorem for hyperbolas*, preprint: arXiv:1208.4008, 2012.
11. J. Bourgain and A. Gamburd, *Expansion and random walks in $SL_d(\mathbb{Z}/p^n\mathbb{Z})$ I*, J. European Math. Soc. **10** (2008), 987–1011.
12. ———, *Uniform expansion bounds for Cayley graphs of $SL_2(F_p)$*, Annals of Math. **167** (2008), no. 625–642, 625–642.
13. ———, *Expansion and random walks in $SL_d(\mathbb{Z}/p^n\mathbb{Z})$ II, with an appendix by J. Bourgain*, J. European Math. Soc. **11** (2009), 1057–1103.
14. J. Bourgain, A. Gamburd, and P. Sarnak, *Affine linear sieve, expanders, and sum-product*, Invent. Math. **179** (2010), no. 3, 559–644.
15. J. Bourgain, N. Katz, and T. Tao, *A sum-product estimate in finite fields, and applications*, Geometric And Functional Analysis **14** (2004), no. 1, 27–57.
16. ———, *A sum-product estimate in finite fields, and applications*, Geom. Funct. Anal. **14** (2004), 27–57.
17. J. Bourgain and P. P. Varjú, *Expansion in $SL_d(\mathbb{Z}/q\mathbb{Z})$, q arbitrary*, Inventiones **188** (2012), no. 1, 151–173.
18. P. Brass and C. Knauer, *On counting point-hyperplane incidences*, Comput. Geom. Theory Appl. **25** (2003), no. 1–2, 13–20.
19. P. Brass, W. Moser, and J. Pach, *Research problems in discrete geometry*, Springer, New York, 2005.
20. E. Breuillard, *Mini-course on approximate groups*, http://www.math.u-psud.fr/∼breuilla/Breuillard_MSRI.pdf.
21. E. Breuillard and B. Green, *Approximate groups, II: the solvable linear case*, Quarterly Journal of Math **62** (2011), no. 3, 513–521.
22. E. Breuillard, B. Green, R. Guralnick, and T. Tao, *Expansion in finite simple groups of Lie type*, in preparation.
23. E. Breuillard, B. Green, and T. Tao, *Small doubling in groups*, arXiv:1301.7718.

24. ———, *Linear approximate groups*, Electronic Research Announcements in Mathematical Sciences **17** (2010), 5767.
25. ———, *Approximate subgroups of linear groups*, Geometric And Functional Analysis **21** (2011), no. 4, 774–819.
26. ———, *Suzuki groups as expanders*, Groups, Geom. Dyn. **5** (2011), no. 2, 281–299, in volume in honour of Fritz Grunewald.
27. ———, *The structure of approximate groups*, Publ. Math. IHES **116** (2012), no. 1, 115–221, arXiv:1110.5008.
28. P. J. Cameron, C. E. Praeger, J. Saxl, and G. M. Seitz, *On the Sims conjecture and distance transitive graphs*, Bull. Lond. Math. Soc. **15** (1983), 499–506.
29. R. W. Carter, *Simple groups of Lie type*, Pure and Applied Mathematics, vol. 28, John Wiley & Sons, London-New York-Sidney, 1972.
30. ———, *Finite groups of Lie type, conjugacy classes and complex characters*, Wiley, New York, 1985.
31. B. Chazelle, H. Edelsbrunner, L. Guibas, and M. Sharir, *A singly-exponential stratification scheme for real semi-algebraic varieties and its applications*, Theoretical Computer Science **84** (1991), 77–105.
32. A. M. Cohen and G. M. Seitz, *The r-rank of the groups of exceptional Lie type*, Nederl. Akad. Wetensch. Indag. Math. **49** (1987), 251–259.
33. M. J. Collins, *Modular analogues of Jordan's theorem for finite linear groups*, J. Reine Angew. Math. **624** (2008), 143–171.
34. M. D. Conder, C. H. Li, and C. E. Praeger, *On the Weiss conjecture for finite locally primitive graphs*, Proc. Edinburgh Math. Soc. **43** (2000), 129–138.
35. P. de la Harpe, *Topics in geometric group theory*, Chicago Lectures in Math., The University of Chicago Press, 2000.
36. O. Dinai, *Expansion properties of finite simple groups*, Ph.D. thesis, Hebrew University, 2009, arXiv:1001.5069.
37. J. D. Dixon, *The structure of linear groups*, Van Nostrand Reinhold Co., 1971.
38. Gy. Elekes, *n points in the plane can determine $n^{3/2}$ unit circles*, Combinatorica **4** (1984), no. 2–3, 131.
39. ———, *Circle grids and bipartite graphs of distances*, Combinatorica **15** (1995), 167–174.
40. ———, *On the number of sums and products*, Acta Arithmetica **LXXXI** (1997), no. 4, 365–367.
41. ———, *On linear combinatorics III*, Combinatorica **19** (1999), no. 1, 43–53.
42. ———, *Sums versus products in number theory, algebra and Erdős geometry — a survey*, Paul Erdős and his Mathematics II, Bolyai Math. Soc. Stud., vol. 11, Bolyai Math. Soc., Budapest, 2002, pp. 241–290.
43. Gy. Elekes and Z. Király, *On the combinatorics of projective mappings*, Journal of Algebraic Combinatorics **14** (2001), no. 3, 183–197.
44. Gy. Elekes, M. B. Nathanson, and I. Z. Ruzsa, *Convexity and sumsets*, Journal of Number Theory **83** (1999), 194–201.
45. Gy. Elekes and L. Rónyai, *A combinatorial problem on polynomials and rational functions*, Journal of Combinatorial Theory, series A **89** (2000), 1–20.
46. Gy. Elekes, M. Simonovits, and E. Szabó, *A combinatorial distinction between unit circles and straight lines: How many coincidences can they have?*, Combinatorics, Probability and Computing **18** (2009), no. 5, 691–705.
47. Gy. Elekes and E. Szabó, *On triple lines and cubic curves — the orchard problem revisited*, preprint.
48. ———, *How to find groups? (and how to use them in Erdős geometry?)*, Combinatorica **32** (2012), no. 5, 537–571.
49. J. Ellenberg, C. Hall, and E. Kowalski, *Expander graphs, gonality and variation of Galois representations*, preprint, arXiv:1008.3675.
50. P. Erdős, *Some applications of graph theory and combinatorial methods to number theory and geometry*, Algebraic methods in Graph Theory, Coll. Math. Soc. J. Bolyai, vol. 25, Bolyai János Math. Soc., 1981, pp. 137–148.

51. P. Erdős, L. Lovász, and K. Vesztergombi, *On graphs of large distances*, Discrete and Computational Geometry **4** (1989), 541–549.
52. P. Erdős and G. Purdy, *Some extremal problems in geometry IV*, Proc. 7th Southeastern Conference Combinatorics, Graph Th. and Comp., 1976, pp. 307–322.
53. P. Erdős and E. Szemerédi, *On sums and products of integers*, To the memory of Paul Turán (P. Erdős, L. Alpár, and G. Halász, eds.), Studies in Pure Mathematics, Akademiai Kiadó - Birkhauser Verlag, 1983, pp. 213–218.
54. W. Feit and J. Tits, *Projective representations of minimum degree of group extensions*, Canad. J. Math. **30** (1978), no. 5, 1092–1102.
55. G Freiman, *Groups and the inverse problems of additive number theory (in Russian)*, Number-theoretic studies in the Markov spectrum and in the structural theory of set addition (in Russian), Kalinin. Gos. Univ., Moscow, 1973, pp. 175–183.
56. W. Fulton, *Algebraic curves*, W.A. Benjamin Inc., New York – Amsterdam, 1969.
57. ———, *Intersection theory*, second ed., Springer-Verlag, New York, 1998.
58. Z. Füredi and I. Palásti, *Arrangements of lines with a large number of triangles*, Proc. AMS **92** (1984), 561–566.
59. N. Gill and H. A. Helfgott, *Growth in solvable subgroups of $GL_r(\mathbb{Z}/p\mathbb{Z})$*, preprint, arXiv:1008.5264, 2010.
60. ———, *Growth of small generating sets in $SL_n(\mathbb{Z}/p\mathbb{Z})$*, Int Math Res Notices **18** (2011), 4226–4251, arXiv:1002.1605.
61. N. Gill, L. Pyber, I. Short, and E. Szabó, *On the product decomposition conjecture for finite simple groups*, accepted in Groups, Geometry, and Dynamics. arXiv:1111.3497, 2012.
62. A. S. Golsefidy and P. P. Varjú, *Expansion in perfect groups*, preprint: arXiv:1108.4900.
63. W. T. Gowers, *Quasirandom groups*, Comb. Probab. Comp. **17** (2008), 363–387.
64. ———, *Quasirandom groups*, Combin. Probab. Comput. **17** (2008), no. 3, 363–387.
65. B. Green and I. Ruzsa, *Freiman's theorem in an arbitrary abelian group*, Jour. London Math. Soc. **75** (2007), no. 1, 163–175.
66. B. Green and T. Tao, *On sets defining few ordinary lines*, preprint: arXiv:1208.4714.
67. M. Gromov, *Groups of polynomial growth and expanding maps*, Publ. Math., Inst. Hautes Étud. Sci. **53** (1981), 53–78.
68. R. M. Guralnick and W. M. Kantor, *Probabilistic generation of finite simple groups*, J. Algebra **234** (2000), no. 2, 743–792, Special issue in honor of Helmut Wielandt.
69. J. Harris, *Algebraic geometry: A first course*, Springer-Verlag, New York, 1992.
70. B. Hartley, *Subgroups of finite index in profinite groups*, Math. Z. **168** (1979), 71–76.
71. R. Hartshorne, *Algebraic geometry*, Graduate texts in mathematics, vol. 52, Springer-Verlag, New York, 1977.
72. H. A. Helfgott, *Growth and generation in $SL_2(\mathbb{Z}/p\mathbb{Z})$*, Annals of Math. **167** (2008), 601–623.
73. ———, *Growth in $SL_3(\mathbb{Z}/p\mathbb{Z})$*, J. European Math. Soc. **13** (2011), no. 3, 761–851.
74. F. Hirzebruch, *Singularities of algebraic surfaces and characteristic numbers*, Contemp. Math. **58** (1986), 141–155.
75. E. Hrushovski, *Contributions to stable model theory*, Ph.D. thesis, University of California, Berkeley, 1986.
76. ———, *The elementary theory of the Frobenius automorphisms*, preprint, arXiv:math.LO/0406514, 2004.
77. ———, *Stable group theory and approximate subgroups*, J. American Math. Soc. **25** (2012), no. 1, 189243, arXiv:0909.2190.
78. E. Hrushovski and A. Pillay, *Definable subgroups of algebraic groups over finite fields*, J. Reine Angew. Math. **462** (1995), 69–91.
79. J. E. Humphreys, *Linear algebraic groups*, Springer Verlag, 1975.
80. ———, *Conjugacy classes in semisimple algebraic groups*, Math. Surveys Monographs, vol. 43, Amer. Math. Soc., Providence, RI, 1995.
81. I. M. Isaacs, *Character theory of finite groups*, AMS Chelsea Publishing, Providence, RI, 2006.

82. A. A. Ivanov and S. V. Shpectorov, *Amalgams determined by locally projective actions*, Nagoya Math. J. **2004** (176), 19–98.
83. J. Jackson, *Rational amusements for winter evenings*, Longman Hurst Rees Orme and Brown, London, 1821.
84. N. Jacobson, *Lectures in abstract algebra*, University Series in Higher Mathematics, vol. III. Theory of fields and Galois theory, Van Nostrand Reinhold Co., 1964.
85. R. E. Jamison, *Planar configurations which determine few slopes*, Geometriae Dedicata **16** (1984), 17–34.
86. W. M. Kantor and E. M. Luks, *Computing in quotient groups*, STOC '90 Proceedings of the twenty-second ACM symposium on Theory of computing, 1990, pp. 524–534.
87. M. Kassabov, A. Lubotzky, and N. Nikolov, *Finite simple groups as expanders*, Proc. Natl. Acad. Sci. USA **103** (2006), no. 16, 6116–6119 (electronic).
88. E. I. Khukhro, A. A. Klyachko, N. Yu. Makarenko, and Y. B. Melnikova, *Automorphism invariance and identities*, Bull. London Math. Soc. **41** (2009), no. 5, 804–816.
89. P. Kleidman and M. Liebeck, *The subgroup structure of the finite classical groups*, London Mathematical Society Lecture Note Series, vol. 129, Cambridge University Press, Cambridge, 1990.
90. J. Kollár, *Rational curves on algebraic varieties*, Springer, Berlin, 1996.
91. ———, *Rational curves on algebraic varieties*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Springer-Verlag, Berlin, 1996.
92. V. Landazuri and G. M. Seitz, *On the minimal degrees of projective representations of the finite Chevalley groups*, J. Algebra **32** (1974), 418–443.
93. ———, *On the minimal degrees of projective representations of the finite Chevalley groups*, J. Algebra **32** (1974), 418–443.
94. M. J. Larsen, *P-adic Nori theory*, preprint, arXiv:0905.2149.
95. M. J. Larsen and R. Pink, *Finite subgroups of algebraic groups*, Journal of the AMS **24** (2011), no. 4, 11051158.
96. R. Lawther and M. W. Liebeck, *On the diameter of a Cayley graph of a simple group of Lie type based on a conjugacy class*, J. Combin. Theory Ser. A **83** (1998), no. 1, 118–137.
97. M. W. Liebeck, N. Nikolov, and A. Shalev, *A conjecture on product decompositions in simple groups*, Groups Geom. Dyn. **4** (2010), no. 4, 799–812.
98. ———, *Groups of Lie type as products of* SL_2 *subgroups*, J. Algebra **326** (2011), 201–207.
99. ———, *Product decompositions in finite simple groups*, Bulletin of the LMS **44** (2012), no. 3, 469–472.
100. M. W. Liebeck, C. E. Praeger, and J. Saxl, *The maximal factorizations of the finite simple groups and their automorphism groups*, Mem. Amer. Math. Soc. **86** (1990), no. 432, iv+151.
101. M. W. Liebeck and L. Pyber, *Upper bounds for the number of conjugacy classes of a finite group*, J. Algebra **198** (1997), no. 2, 538–562.
102. ———, *Finite linear groups and bounded generation*, Duke Math. J. **107** (2001), no. 1, 159–171.
103. M. W. Liebeck and G.M. Seitz, *On the subgroup structure of exceptional groups of Lie type*, Trans. Amer. Math. Soc. **350** (1998), 3409–3482.
104. M. W. Liebeck and A. Shalev, *Diameters of finite simple groups: sharp bounds and applications*, Ann. of Math. (2) **154** (2001), no. 2, 383–406.
105. A. Lubotzky, *Cayley graphs: eigenvalues, expanders and random walks*, vol. 218, ch. Surveys in Combinatorics, pp. 155–189, Cambridge Univ. Press, 1995.
106. ———, *Finite simple groups of Lie type as expanders*, J. Eur. Math. Soc. (JEMS) **13** (2011), no. 5, 1331–1341.
107. A. Lubotzky and D. Segal, *Subgroup growth*, Birkhäuser, 2003.
108. J. Matoušek, *Lectures on discrete geometry*, Springer–Verlag, Berlin, Heidelberg, New York, 2002.
109. C. Matthews, L. Vaserstein, and B. Weisfeiler, *Congruence properties of Zariski-dense subgroups*, Proc. LMS **48** (1984), 514–532.

110. G. Megyesi and E. Szabó, *On the tacnodes of configurations of conics in the projective plane*, Mathematische Annalen **305** (1996), 693–703.
111. D. Segal N. Nikolov, *On finitely generated profinite groups, I: strong completeness and uniform bounds*, Annals of Math. **165** (2007), 171–238.
112. N. Nikolov, *A product of decomposition for the classical quasisimple groups*, J. Group Theory **10** (2007), no. 1, 43–53.
113. N. Nikolov and L. Pyber, *Product decompositions of quasirandom groups and a Jordan-type theorem*, J. European Math. Soc. **13** (2011), 1063–1077.
114. ———, *Product decompositions of quasirandom groups and a Jordan-type theorem*, J. Eur. Math. Soc. **13** (2011), no. 4, 1063–1077.
115. M. V. Nori, *On subgroups of $GL_n(F_p)$*, Invent. Math. **88** (1987), 257–275.
116. J. E. Olson, *On the sum of two sets in a group*, J. Number Theory **18** (1984), 110–120.
117. A. L. Onishchik and E. B. Vinberg, *Lie groups and algebraic groups*, Springer, Berlin, 1990.
118. J. Pach and P. K. Agarwal, *Combinatorial geometry*, J. Wiley and Sons, New York, 1995.
119. J. Pach and M. Sharir, *Repeated angles in the plane and related problems*, Journal of Combinatorial Theory, series A **59** (1990), 12–22.
120. ———, *On the number of incidences between points and curves*, Combinatorics, Probability and Computing **7** (1998), 121–127.
121. G. Petridis, *New proofs of Plünnecke-type estimates for product sets in groups*, 2011, Preprint available on the Math arXiv: `http://arxiv.org/abs/1101.3507`.
122. A. Pillay, *Geometric stability theory*, Oxford Logic Guides, vol. 32, Clarendon Press, Oxford, 1996.
123. H. Plünnecke, *Eigenschaften und Abschätzungen von Wirkungsfunktionen*, BMwF-GMD-22, Gesellschaft für Mathematik und Datenverarbeitung, Bonn, 1969.
124. ———, *Eine zahlentheoretische Anwendung der Graphentheorie*, J. Reine Angew. Math. **243** (1970), 171–183.
125. C. Praeger, L. Pyber, P. Spiga, and E. Szabó, *Graphs with automorphism groups admitting composition factors of bounded rank*, Proc. of the AMS. **140** (2012), no. 7, 2307–2318.
126. C. E. Praeger, *Imprimitive symmetric graphs*, Ars Combinatoria **19A** (1985), 149–163.
127. ———, *Finite quasiprimitive graphs*, Surveys in combinatorics, London Mathematical Society Lecture Note Series, vol. 24, London Mathematical Society, 1997, pp. 65–85.
128. ———, *Finite transitive permutation groups and bipartite vertex-transitive graphs*, Illinois Journal of Mathematics **47** (2003), 461–475.
129. C. E. Praeger, P. Spiga, and G. Verret, *Bounding the size of the vertex-stabiliser in vertex-transitive graphs*, preprint: arXiv:1102.1543, 2011.
130. L. Pyber, *Asymptotic results for permutation groups*, Groups and computation, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 11, Amer. Math. Soc., Providence, RI, 1993, pp. 197–219.
131. L. Pyber and E. Szabó, *Helfgott's conjecture, soluble version*, preprint.
132. ———, *Growth in finite simple groups of Lie type*, announcement, arXiv:1001.4556, 2010.
133. ———, *Growth in finite simple groups of Lie type of bounded rank*, preprint, arXiv:1005.1858, 2010.
134. M. Reid, *Undergraduate algebraic geometry*, Cambridge University Press, 1988.
135. A. H. Rhemtulla, *Commutators of certain finitely generated soluble groups*, Canad. J. Math. **21** (1969), 1160–1164.
136. I. Ruzsa, *Towards a noncommutative Plünnecke-type inequality*, An Irregular Mind, Bolyai Society Mathematical Studies, vol. 21, Bolyai Society, 2010, pp. 591–605.
137. I. Ruzsa and S. Turjányi, *Note on additive bases of integers*, Publ. Math. (Debrecen) **32** (1985), 101–104.
138. I. Z. Ruzsa, *Generalized arithmetical progressions and sumsets*, Acta Math. Hung. **65** (1994), no. 4, 379–388.

139. ———, *Sums of finite sets*, Number theory (New York, 1991–1995), Springer, New York, 1996, pp. 281–293.
140. ———, *Sumsets and structure*, Combinatorial number theory and additive group theory, Adv. Courses Math. CRM Barcelona, Birkhäuser Verlag, Basel, 2009, pp. 87–210.
141. G. Sabidussi, *Vertex-transitive graphs*, Monatsh. Math. **68** (1964), 426–438.
142. A. Shalev, *Word maps, conjugacy classes, and a noncommutative Waring-type theorem*, Ann. of Math. (2) **170** (2009), no. 2, 1383–1416.
143. C. C. Sims, *Graphs and finite permutation groups*, Math. Z. **95** (1967), 76–86.
144. J. Solymosi, *On the number of sums and products*, Bull. London Math. Soc. **37** (2005), no. 4, 491494.
145. J. Solymosi and T. Tao, *An incidence theorem in higher dimensions*, Discrete & Computational Geometry **48** (2012), no. 2, 255–280.
146. J. Spencer, E. Szemerédi, and W. T. Trotter, Jr., *Unit distances in the euclidean plane*, Graph theory and combinatorics, Academic Press, London, 1984, pp. 293–303.
147. P. Spiga, *Two local conditions on the vertex stabiliser of arc-transitive graphs and their effect on the Sylow subgroups*, preprint: arXiv:1102.4421, 2011.
148. A. Stein, $1\frac{1}{2}$*-generation of finite simple groups*, Beiträge Algebra Geom. **39** (1998), no. 2, 349–358.
149. R. Steinberg, *Endomorphisms of linear algebraic groups*, Memoirs Amer. Math. Soc., vol. 80, Amer. Math. Soc., 1968.
150. M. Suzuki, *Group theory I*, Grundlehren der Mathematischen Wissenschaften, vol. 247, Springer-Verlag, Berlin-New York, 1982.
151. ———, *Group theory. II*, Grundlehren der Mathematische Wissenschaften, vol. 248, Springer-Verlag, New York, Berlin, Heidelberg, and Tokyo, 1986.
152. J. J. Sylvester, *Problem 2473*, Math. Questions from the Educational Times **8** (1867), 106–107.
153. L. A Székely, *Crossing numbers and hard Erdős problems in discrete geometry*, Combinatorics, Probability and Computing **6** (1997), no. 3, 353–358.
154. E. Szemerédi and W. T. Trotter, Jr., *Extremal problems in discrete geometry*, Combinatorica **3** (1983), no. 3–4, 381–392.
155. T. Tao, *The Bourgain-Gamburd expansion machine*, blog: http://terrytao.wordpress.com/2012/01/13/254b-notes-4-the-bourgain-gamburd-expansion-machine.
156. ———, *Expanding polynomials over finite fields of large characteristic, and a regularity lemma for definable sets*, preprint: arXiv:1211.2894.
157. ———, *Expanding polynomials over finite fields of large characteristic, and a regularity lemma for definable sets*, blog: http://terrytao.wordpress.com/2012/11/14/expanding-polynomials-over-finite-fields-of-large-characteristic-and-a-regularity-lemma-for-definable-sets.
158. ———, *Product set estimates for non-commutative groups*, Combinatorica **28** (2008), no. 5, 547–594.
159. ———, *The sum-product phenomenon in arbitrary rings*, Contrib. Discrete Math. **4** (2009), no. 2, 59–82.
160. T. Tao and V. Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, vol. 105, Cambridge University Press, Cambridge, 2006.
161. J. G. Thompson, *Bounds for the orders of maximal subgroups*, J. Algebra **14** (1970), 135–138.
162. D. Cs. Tóth, *The Szemeredi-Trotter theorem in the complex plane*, preprint: arXiv:math/0305283, 2003.
163. V. I. Trofimov, *Graphs with projective suborbits. exceptional cases of characteristic 2. IV*, Izv. Ross. Akad. Nauk Ser. Mat. **67** (2003), 193–222, translation: *Izv. Math.* **67** (2003), 1267–1294.
164. V. I. Trofimov and R. M. Weiss, *The group $E_6(q)$ and graphs with a locally linear group of automorphisms*, Math. Proc. Cambridge Philos. Soc. **148** (2010), 1–32.
165. P. Varjú, *Expansion in $SL_d(O_K/I)$, I square-free*, J. Eur. Math. Soc. (JEMS) **14** (2012), no. 1, 273–305.

166. B. A. F. Wehrfritz, *Infinite linear groups*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 76, Springer-Verlag, 1973.
167. A. Weil, *On algebraic groups and homogeneous spaces*, American Journal of Mathematics **77** (1955), 493–512.
168. ———, *On algebraic groups of transformations*, American Journal of Mathematics **77** (1955), 355–391.
169. B. Weisfeiler, *Post-classification version of Jordan's theorem on finite linear groups*, Proc. Nat. Acad. Sci. U.S.A. **81** (1984), no. 16, 5278–5279.
170. R. Weiss, *s-transitive graphs*, Colloq. Math. Soc. János Bolyai, vol. 25, Math. Soc. János Bolyai, 1978, pp. 827–847.
171. ———, *Symmetric graphs with projective subconstituents*, Proc. Amer. Math. Soc. **72** (1978), 213–217.
172. ———, *Symmetrische Graphen mit auflösbaren Stabilisatoren*, J. Algebra **53** (1978), 412–415.
173. ———, *An application of p-factorization methods to symmetric graphs*, Math. Proc. Cambridge Philos. Soc. **85** (1979), 43–48.
174. F. Wettl, *On the nuclei of a point set of a finite projective plane*, Journal of Geometry **30** (1987), no. 2, 157–163.