CHAPTER 3

The Bourgain–Gamburd machine

The Bourgain–Gamburd machine, as it is now commonly called, is a general principle, devised by J. Bourgain and A. Gamburd [1] in 2008 to prove that certain families of Cayley graphs are expanders. At present it is perhaps the most versatile and effective way of doing that, being also, if this makes any sense, more elementary when compared to the previously known general approaches, which were based on rather deep results involving e.g. Kazhdan property (T), or the weaker Lubotzky property (τ) , Selberg Trace formula, etc.¹

In this chapter we prove Bourgain and Gamburd Theorem, which we will apply in the next chapter to families of Cayley graphs in groups $SL_2(q)$.

3.1. The Balog-Szemerédi-Gowers lemma

The Balog-Szemeredi-Gowers Lemma is a fundamental tool in additive combinatorics, later extended by Tao to the non-commutative setting. As such, it entirely belongs to the matter of Chapter 1. We postponed it here because its proof (for which we follow Tao's approach in [**35**]) proceeds through a graph-theoretical version, which is somehow easier to visualize and manipulate, and because, for our purposes, the Balog-Szemerédi-Gowers Lemma, which has lots of other important applications in combinatorics, is instrumental in the proof of the Bourgain-Gamburd Theorem.

Before starting, let us fix the following convention, which will be in force for the rest of these notes. If $\alpha(I)$, $\beta(I)$ are positive real numbers, that depend on a certain set of variables I, we write

$$\alpha(I) \ll \beta(I)$$

if there is an absolute constant C > 0 (that is, C does not depend on any of the variables) such that $\alpha(I) \leq C\beta(I)$ for every assignation of the variables I. Of course, $\alpha(I) \gg \beta(I)$ means $\beta(I) \ll \alpha(I)$.

As said, we begin with graphs. When saying that $\Gamma = (A \cup B, E)$ is a bipartite graph, we tacitly mean that $A \cup B$ is a partition of the vertex set by non-empty A and B, such that every edge intersects both non-trivially. Also, for every vertex

¹These methods are at any rate still very important, and deserve attention; however they are out of the scope of this course.

 $x \in V$ of a graph $\Gamma = (V, E)$, we denote by $N(x) = N_{\Gamma}(x)$ its neighborhood, that is $N(x) = \{y \in V \mid x \sim y\}$; clearly $|N(x)| = d_{\Gamma}(x)$.

LEMMA 3.1 (Balog-Szemerédi-Gowers: two-step walks). Let $\Gamma = (A \cup B, E)$ be a finite bipartite graph and $k \ge 1$ such that $|E| \ge |A||B|/k$. Let $\varepsilon > 0$. Then there exists a subset $X \subseteq A$ such that

•
$$|X| \ge \frac{|A|}{k\sqrt{2}}$$

• $\left| \left\{ (a,a') \in X \times X \mid |N(a) \cap N(a')| > \frac{\varepsilon}{2k^2} |B| \right\} \right| \ge (1-\varepsilon)|X|^2.$

(the second item says that at least $(1 - \varepsilon)|X|^2$ of the pairs $(a, a') \in X^2$ are such that there are more than $\frac{\varepsilon}{2k^2}|B|$ walks of length 2 from a to a').

PROOF. The idea is to search for such $X \subseteq A$ among the neighborhoods N(b) of the elements of B.

We say that a pair $(a, a') \in A \times A$ is bad if $|N(a) \cap N(a')| \leq \frac{\varepsilon}{2k^2}|B|$, and for $Y \subseteq A$ we denote by $\beta(Y)$ the set of all bad pairs (a, a') with $a, a' \in Y$.

Suppose, by contradiction, that for no $b \in B$ the set X = N(b) satisfies the properties in the statement, and let

$$B_0 = \left\{ b \in B \left| |N(b)| < \frac{|A|}{\sqrt{2k}} \right\}.$$

Finally, let $\Omega = \{(a, a', b) \in A \times A \times B \mid (a, a') \in \beta(N(b))\}$. Then

$$|\Omega| = \sum_{(a,a')\in\beta(A)} |N(a)\cap N(a')| \le |\beta(A)|\frac{\varepsilon}{2k^2}|B| \le \frac{\varepsilon}{2k^2}|A|^2|B|.$$

On the other hand

$$|\Omega| \ge \sum_{b \in B \setminus B_0} |\beta(N(b))| > \sum_{b \in B \setminus B_0} \varepsilon |N(b)|^2 = \varepsilon \sum_{b \in B \setminus B_0} |N(b)|^2.$$

Therefore

(3.1)
$$\sum_{b \in B \setminus B_0} |N(b)|^2 < \frac{|A|^2 |B|}{2k^2}.$$

Now, by the Cauchy-Schwarz inequality,

$$\sum_{b \in B} |N(b)|^2 \ge \frac{1}{|B|} \left(\sum_{b \in B} |N(b)|\right)^2 = \frac{|E|^2}{|B|} \ge \frac{|A|^2|B|}{k^2}$$

Hence,

(3.2)
$$\sum_{b \in B \setminus B_0} |N(b)|^2 \ge \frac{|A|^2 |B|}{k^2} - \sum_{b \in B_0} |N(b)|^2 > \frac{|A|^2 |B|}{k^2} - \frac{|A|^2}{2k^2} |B_0|.$$

Comparing with (3.1), we get

$$\frac{|A|^2|B|}{2k^2} > \frac{|A|^2|B|}{k^2} - \frac{|A|^2|B_0|}{2k^2},$$

whence the contradiction $|B_0| > |B|$.

LEMMA 3.2 (Balog-Szemerédi-Gowers: three-step walks). Let $\Gamma = (A \cup B, E)$ be a finite bipartite graph and $k \ge 1$ such that $|E| \ge |A||B|/k$. Then there exist subsets $A' \subseteq A, B' \subseteq B$ such that

- $|A'| \ge c|A|/k^3$, $|B'| \ge c|B|/k^3$;
- for every $a \in A'$, $b \in B'$, there are at least $d|A||B|/k^7$ walks of length three from a to b;

where c and d are absolute constants..

PROOF. Let A_1 be the set of all elements in A whose degree is at least |B|/2k, and let E_1 be the set of all edges that have an extreme in A_1 . Then

(3.3)
$$|A_1||B| \ge |E_1| \ge |E| - \frac{|B|}{2k}(|A| - |A_1|) \ge \frac{|A||B|}{k} - \frac{|A||B|}{2k}$$

hence $|A_1||B| \ge |E_1| \ge |A||B|/2k$, and so

$$|A_1| \ge \frac{|A|}{2k}$$
 and $|E_1| \ge \frac{|A_1||B|}{2k}$

Applying Lemma 3.1, for $\varepsilon > 0$ to be later specified, to the subgraph induced by $A_1 \cup B$, we find $A_2 \subseteq A_1$ with

$$|A_2| \ge \frac{|A_1|}{2\sqrt{2}k} \ge \frac{|A|}{4\sqrt{2}k^2},$$

such that if \mathcal{W} is the set of all pairs $(a, a') \in A_2 \times A_2$ that are not connected by more than $\frac{\varepsilon}{8k^2}|B|$ walks of length 2, then $|\mathcal{W}| < \varepsilon |A_2|^2$.

Let A' be the set of all $a \in A_2$ such that the number of vertices $a' \in A_2$ with $(a, a') \in \mathcal{W}$ is at most $\sqrt{\varepsilon}|A_2|$; then,

$$(|A_2| - |A'|)\sqrt{\varepsilon}|A_2| \le |\mathcal{W}| \le \varepsilon |A_2|^2$$

and so

$$(3.4) |A'| \ge (1 - \sqrt{\varepsilon})|A_2|.$$

Let $E_2 = \partial A_2$ be the set of all edges joining A_2 to B. Since every element of A_2 has degree at least |B|/2k, we have

(3.5)
$$|E_2| \ge \frac{|A_2||B|}{2k} \ge \frac{|A||B|}{8\sqrt{2k^3}}$$

Finally, let B' the subset of all vertices in B that are adjacent to at least $|A_2|/4k$ elements of A_2 ; by (3.5), arguing as for (3.3), we have

$$(3.6) |B'| \ge \frac{|B|}{4k}.$$

Let $a \in A'$ and $b \in B'$. Then, denoting by $\gamma(b)$ is the number of elements a' of A_2 such that a' is adjacent to b and $(a, a') \notin \mathcal{W}$, we have that the number of distinct walks of length 3 connecting b to a is at least $\frac{\varepsilon}{8k^2}|B|\gamma(b)$. Since, by construction,

$$\gamma(b) \ge \frac{|A_2|}{4k} - \sqrt{\varepsilon}|A_2|,$$

we have that the number of distinct walks of length 3 connecting b to a is at least

$$\frac{\varepsilon}{8k^2} \left(\frac{1}{4k} - \sqrt{\varepsilon}\right) |A_2| |B| \ge \left(\frac{1}{4k} - \sqrt{\varepsilon}\right) \frac{\varepsilon}{64k^4} |A| |B|.$$

Choosing, for instance $\varepsilon = (8k)^{-2}$ we obtain the desired bound for this number. The bounds for |A'| and |B'| follow at once from (3.4) and(3.6).

We come now to the group theoretical version of the Balog-Szemerédi-Gowers Lemma. Then, it is time for the following definition.

DEFINITION 3.3. Let A, B be finite non-empty subsets of a group. The *(multiplica-tive)* energy of the pair (A, B) is

$$E(A,B) = \left| \{ (a,b,a',b') \in A \times B \times A \times B \mid ab = a'b' \} \right|.$$

It is convenient to observe immediately that for any pair A, B of subsets of a finite group G one has

(3.7)
$$E(A,B) = ||\mathbf{1}_A * \mathbf{1}_B||^2.$$

Energy thus counts the number of "coincidences" in the product-set AB; it is then intuitive that large energy should correspond to small product. This is true only in one sense: let us prove the following elementary fact.

LEMMA 3.4. Let A, B be finite non-empty subsets of a group. Then

$$\frac{|A|^2|B|^2}{|AB|} \le E(A,B) \le |A|^{3/2}|B|^{3/2}.$$

PROOF. For $x \in AB$, write $r(x) = |\{(a, b) \in A \times B \mid ab = x\}|$. Then, by the Cauchy-Schwarz inequality,

$$E(A,B) = \sum_{x \in AB} r(x)^2 \ge \frac{1}{|AB|} \left(\sum_{x \in AB} r(x)\right)^2 = \frac{|A \times B|^2}{|AB|} = \frac{|A|^2|B|^2}{|AB|}.$$

For the upper bound, observe that we have, for any $x \in AB$,

(3.8)
$$r(x) \le \min\{|A|, |B|\} \le \sqrt{|A||B|},$$

hence

$$E(A,B) = \sum_{x \in AB} r(x)^2 \le \sqrt{|A||B|} \sum_{x \in AB} r(x) = \sqrt{|A||B|} |A \times B| = |A|^{3/2} |B|^{3/2}.$$

Therefore (letting A = B for easy), if $|AA| \leq k|A|$ is small we have, by the left inequality, that $E(A, A) \geq |A|^3/k$ is large. The converse is not quite true. Consider, for instance, a set which is the union of two parts of comparable sizes, one of which in charge of producing most of a big set-product, the other with small double and large enough energy; as an explicit example, you may take $P \cup Q = A \subseteq \mathbb{Z}$, with $P = \{0, 1, \ldots, n-1\}$ and $Q = \{n, n^2, \ldots, n^n\}$; then |A + A| and E(A, A) are both large (with respect to |A| = 2n, or to $|A \times A| = 4n^2$ for the energy) being of size, respectively, n^2 and n^3 . The Balog-Szemerédi-Gowers Lemma shows that what is going on in this example is in fact a general feature: if the energy E(A, B) is large then there are big portions of A and of B whose set-product is small.

Before coming to the proof of the group-theoretical version of Lemma 3.2, we observe a few elementary facts about convolution of characteristic functions. The first we leave as an exercise.

EXERCISE 29. Let A be a finite subset of a group. Prove that

$$E(A, A^{-1}) = E(A^{-1}, A).$$

Now, if A, B are two non-empty subsets of a finite group G then the function $\mathbf{1}_A * \mathbf{1}_B$ takes non-zero values only on the elements of AB, and for every $a \in A, b \in B$,

$$(\mathbf{1}_A * \mathbf{1}_B)(ab)$$

is a positive integer that counts the number of distinct solutions $(x, y) \in A \times B$ of xy = ab (i.e. the quantity r(ab) in the proof of Lemma 3.4). In particular, it follows that for every $g \in G$ we have

$$(\mathbf{1}_A * \mathbf{1}_B)(g)^2 = \sum_{(a,b)\in A\times B, ab=g} (\mathbf{1}_A * \mathbf{1}_B)(g),$$

and so

$$(3.9) E(A,B) = ||\mathbf{1}_A * \mathbf{1}_B||^2 = \sum_{g \in G} (\mathbf{1}_A * \mathbf{1}_B)(g)^2 = \sum_{(a,b) \in A \times B} (\mathbf{1}_A * \mathbf{1}_B)(ab).$$

Also, as observed in (3.8), for every $(a, b) \in A \times B$ one has

(3.10)
$$(\mathbf{1}_A * \mathbf{1}_B)(ab) \le |A|^{1/2} |B|^{1/2}.$$

We are now ready for the proof.

THEOREM 3.5 (Balog-Szemerédi-Gowers lemma: energy form). Let A, B be finite subsets of a group G such that, for some $k \ge 1$,

$$E(A, B) \ge |A|^{3/2} |B|^{3/2} / k.$$

Then there exist subsets $A' \subseteq A$, $B' \subseteq B$, with $|A'| \ge c|A|/k^3$, $|B'| \ge c|B|/k^3$, such that

$$|A'B'| \leq Ck^{10}|A|^{1/2}|B|^{1/2}$$

where c and C are absolute constants. Moreover, $|A'A'^{-1}| \leq C^2 k^{20} |A|$.

PROOF. Let E be the set of all pairs $(a, b) \in A \times B$ such that

$$(\mathbf{1}_A * \mathbf{1}_B)(ab) > |A|^{1/2} |B|^{1/2} / 2k.$$

Now, by hypothesis, using (3.9) and (3.10),

$$|E||A|^{1/2}|B|^{1/2} + |A \times B|\frac{|A|^{1/2}|B|^{1/2}}{2k} \ge E(A,B) \ge \frac{|A|^{3/2}|B|^{3/2}}{k},$$

whence

$$|E| \ge |A||B|/2k.$$

We now consider the bipartite graph $\Gamma = (A \cup B, E)$, where A and B are abstractly viewed and their union is taken disjoint, while the pairs from E are considered as indirected.

By applying Lemma 3.2 to Γ , we find subsets A' of A and B' of B such that

$$A'| \ge \frac{c|A|}{k^3}, \quad |B'| \ge \frac{c|B|}{k^3},$$

with the property that for every $a \in A'$ and $b \in B'$, the number of walks of length three connecting a to b is at least $d|A||B|/k^7$. This property, interpreted in the graph Γ means that there exist at least $d|A||B|/k^7$ distinct pairs $(x, y) \in A \times B$ such that

$$(a, y), (x, y), (x, b) \in E$$

Now, $(\mathbf{1}_A * \mathbf{1}_B)(xy) = (\mathbf{1}_{B^{-1}} * \mathbf{1}_{A^{-1}})((xy)^{-1})$. Since $ab = (ay)(xy)^{-1}(xb)$, the above condition implies that, for every $(a, b) \in A' \times B'$, there are at least

$$\frac{d|A||B|}{k^7} \Big(\frac{|A|^{1/2}|B|^{1/2}}{2k}\Big)^3 = \frac{d|A|^{5/2}|B|^{5/2}}{8k^{10}}$$

distinct 6-tuples $(a_1, b_1, a_2, b_2, a_3, b_3) \in (A \times B)^3$ such that $ab = a_1b_1(a_2b_2)^{-1}a_3b_3$. Therefore, setting $C = 8d^{-1}$,

$$|A'B'| \le |A|^3 |B|^3 \frac{8k^{10}}{d|A|^{5/2}|B|^{5/2}} = Ck^{10} |A|^{1/2} |B|^{1/2}.$$

The last claim follows from Ruzsa triangle inequality (Lemma 1.10). In fact, by that inequality,

$$|B||A'A'^{-1}| \le |A'B'||B'^{-1}A'^{-1}| = |A'B'|^2 \le C^2 k^{20}|A||B|,$$

and the proof is completed.

Another variation finally connects large energy to approximate groups.

COROLLARY 3.6 (B.S.G.; approximate subgroup form). Let A be a finite subset of a group G, and suppose that, for some $k \ge 2$,

$$E(A, A) \ge |A|^3/k.$$

Then there exist a k^d -approximate subgroup Q of G with

$$k^{-d}|A| \le |Q| \le k^d|A|,$$

and an element $g \in G$ such that

$$|A \cap Qg| \gg k^D |A|,$$

where d and D are absolute constants.

58

PROOF. Let A be as in the assumptions. By Theorem 3.5, there exists $A_1 \subseteq A$ such that $|A_i| \ge ck^{-3}|A|$ and $|A_1A_1^{-1}| \le c^2k^{20}|A|$. Then, by exercise 29 and Lemma 3.4,

$$E(A_1^{-1}, A_1) = E(A_1, A_1^{-1}) \ge \frac{|A|^3}{ck^{20}}$$

By another application of Theorem 3.5, we therefore find a subset X of A_1 with $|X| \ge c^{-2}k^{O(1)}|A|$, such that

$$|X^{-1}X| \ll k^{O(1)}|A_1| \ll k^{O(1)}|A|.$$

Now, observe that, since $X \subseteq A_1$

$$|XX^{-1}| \le |A_1A_1^{-1}| \le c^2 k^{20} |A|.$$

We may then apply Proposition 1.42 to deduce that there exist an absolute constant d and a k^d -approximate subgroup Q of G, with

$$k^{-d}|A| \le |Q| \le k^d|A|,$$

such that X is contained in at most $O(k^d)$ translates of Q (we have here exploited the fact that, since $k \ge 2$, $k^{\log_2 C} \ge C$ for any absolute constant C). In particular, there exists $y \in G$, and an absolute constant D such that

$$|X \cap Qy| \gg k^{-d}|X| \gg k^D|A|,$$

thus finishing the proof.

EXERCISE 30. [Tao [35]] Let A be a non-empty finite symmetric subsets of a group G, and $k \ge 1$. Suppose that there exist a k-approximate subgroup H of G with $|H| \le k|A|$, and a $g \in H$ such that $|A \cap gH| \ge |A|/k$. Prove that $E(A, A) \ge k^{-6}|A|^3$.

EXERCISE 31. [Commensurable sets] Let $t \ge 1$. Two finite subsets A, B of a group G, are said to be *t*-commensurable if $\max\{|A|, |B|\} \le t|A \cap B|$.

(i) Let A, B, C be finite subsets of a group. Prove that if A, B are t_1 -commensurable and B, C are t_2 -commensurable, then A, C are t_1t_2 -commensurable.

(ii) Consider the following properties of a finite subset A of a group.

- (1) A is a k-approximate subgroup;
- (2) $|A^2| \le k|A|;$
- (3) $E(A, A) \ge |A|^3/k$.

Prove that these properties are *roughly equivalent* in the following sense. For any two, (\star) and (\dagger) , of these three properties, if a finite subset A of a group G satisfies (\star) with parameter k, then there exits a finite subset B of G, satisfying (\dagger) with respect to a parameter f(k), such that A is t(k)-commensurable with a By for some $y \in G$, where t(k), f(k) are polynomial functions of k only.

3. THE BOURGAIN-GAMBURD MACHINE

3.2. Representations

In this section we are going to use some basic results of the representation theory of finite groups over the field of complex numbers. We do not have time to give any proofs of these fundamental but reasonably standard facts, so those who are not familiar with representations are invited to take them for granted.

Let $1 \leq d \in \mathbb{N}$. A *d*-dimensional \mathbb{C} -representation of a group *G* is a homomorphism $\pi : G \to GL(V)$, where *V* is a *d*-dimensional \mathbb{C} -vector space, and GL(V) is the group of all invertible linear maps of *V* in itself (in matrix form $\pi : G \to GL_d(\mathbb{C})$). The representation π is *trivial* if $\pi(g)$ is the identity map for every $g \in G$, while, on the opposite side, π is *faithful* if ker $\pi = \{1\}$. The principal representation of *G* is the trivial representation on the one-dimensional space \mathbb{C} .

A representation $\pi: G \to GL(V)$ is *irreducible* if $\{0\}$ and V are the only subspaces left invariant by $\pi(G)$. If G is finite, it can be proved that every \mathbb{C} -representation $\pi: G \to GL(V)$ may be decomposed as a sum of *irreducible representations*. This mean that there is a decomposition $V = V_1 \oplus \ldots \oplus V_n$ of V into the sum of subspaces V_i that are invariant by $\pi(G)$, such that the restrictions

$$\pi_i: G \to GL(V_i)$$
$$x \mapsto \pi(x)_{|V_i|}$$

are irreducible representations of G. In this case we write $\pi = \pi_1 \oplus \ldots \oplus \pi_n$. There is a natural definition of *equivalence* of \mathbb{C} -representations of given group, that we do not bother to recall; if G is finite, the number of distinct irreducible representations of G up to equivalence coincides with the number of distinct conjugacy classes of G.

If G is finite, one has the regular representation $\rho = \rho_G$, which is the permutation representation (i.e. a representation in which the map associated to any $x \in G$ is a permutation of the vectors of a fixed base) arising from right multiplication in G, and that in our setting may be conveniently described by letting $V = \ell^2(G)$. Thus, for every $f \in \ell^2(G)$ and $g, x \in G$,

(3.11)
$$(\rho(g)(f))(x) = f(xg^{-1}).$$

Let S be a symmetrical subset of G, and write $A_{\rho} = \sum_{x \in S} \rho(x)$. Then, if $A = A(\Gamma)$ is the adjacency operator of the Cayley graph $\Gamma = \Gamma[G; S]$ and $f \in \ell^2(G)$,

(3.12)
$$Af(g) = \sum_{y \sim g} f(y) = \sum_{x \in S} f(gx) = \sum_{x \in S} (\rho(x^{-1})f)(g) = A_{\rho}f(g),$$

for every $g \in G$. Hence $A = A_{\rho}$.

Now, a fundamental result in the representation theory of a finite group says that the regular representation ρ_G is the sum of all irreducible representations of G(up to equivalence) each appearing with multiplicity equal to its dimension. In our notation

(3.13)
$$\rho = d_0 \pi_0 \oplus d_1 \pi_1 \oplus \ldots \oplus d_c \pi_c$$

where $\pi_0, \pi_1, \ldots, \pi_c$ is a set of distinct representatives of the equivalence classes of irreducible representations of G and, for every $i = 0, \ldots, c, d_i \pi_i = \pi_i \oplus \ldots \oplus \pi_i$ $(d_i = \dim(\pi_i) \text{ summands})$. In connection with (3.12) we then in particular have a proof of the following observation.

PROPOSITION 3.7. Let S be a symmetrical subset of a finite group G. Then the eigenvalues of the adjacency operator $A(\Gamma)$ on $\ell^2(G)$ of the Cayley graph $\Gamma = \Gamma[G; S]$ are the eigenvalues of the operators

$$A_{\pi}(S) = \sum_{x \in S} \pi(x)$$

when π varies in the set of all irreducible representations of G. Moreover, if μ is an eigenvalue of A_{π} , then μ occurs in the spectrum of $A(\Gamma)$ with multiplicity at least dim (π) .

If in (3.13) we agree, as customary, that π_0 is the principal representation, then $A_{\pi_0}(S)$ is the multiplication (on the 1-dimensional space \mathbb{C}) by k = |S|, which in fact is the first eigenvalue μ_0 of $A(\Gamma)$.

NOTE. (Unitary representations) A \mathbb{C} -representation $\pi : G \to GL(V)$ of a group G is unitary if $\pi(G)$ is contained in the subgroup U(V) of all unitary transformations of V. This means that for every $g \in G$ and all $v \in V$,

$$\langle \pi(g)(v), \pi(g)(v) \rangle = \langle v, v \rangle.$$

This concept turns out to be central for infinite (locally compact) groups, and in fact it plays a fundamental role in the treatment of expansion of Cayley graphs via Kazhdan property (T) (see section 2.1 in Tao's book [**35**]); for finite groups, which will always be the case in our approach, it however does not make much difference: in fact, any representation of a finite group on a \mathbb{C} -space V may be "made" unitary (see exercise below).

EXERCISE 32. Let G be a finite group.

- (i) Prove that the regular representation ρ_G is unitary.
- (ii) Let $\pi : G \to GL(\mathbb{C}^n)$ be any \mathbb{C} -representation of G; prove that setting, for every $u, v \in \mathbb{C}^n$

$$\langle u, v \rangle_G = \sum_{x \in G} \langle \pi(x)u, \pi(x)v \rangle$$

defines a hermitian product $\langle \cdot, \cdot \rangle_G$ on \mathbb{C}^n with respect to which π is unitary.

3. THE BOURGAIN-GAMBURD MACHINE

3.3. The Bourgain-Gamburd machine

We begin by a fundamental application of the B.S.G. Lemma connecting behavior of approximate subgroups of a group to that of convolution power of a probability distribution on G.

Remember that a probability measure ν on a group G is said to be symmetrical if $\nu(g^{-1}) = \nu(g)$ for all $g \in G$.

LEMMA 3.8 (Bourgain–Gamburd "flattening" Lemma). There exists a constant R > 0 such that for any $k \ge 2$, a finite group G, and a symmetric probability measure ν on G, one of the following cases occur

- (i) either $||\nu * \nu|| \le k^{-1} ||\nu||$, or
- (ii) there exist a k^R -approximate subgroup Q of G, with

$$k^{-R}/||\nu||^2 \le |Q| \le k^R/||\nu||^2$$

and an element $x \in G$ such that $\nu(xQ) \ge k^{-R}$.

PROOF. We first assume that ν is the probability measure uniformly centered in a symmetric subset A of G, that is $\nu = |A|^{-1} \mathbf{1}_A$, with $\emptyset \neq A = A^{-1} \subseteq G$. Then, $||\nu||^2 = \sum_{x \in A} |A|^{-2} = |A|^{-1}$ and

$$||\nu * \nu||^2 = |A|^{-4}||1_A * 1_A||^2 = |A|^{-4}E(A, A).$$

Suppose that (i) is not the case. We then have

$$|A|^{-4}E(A,A) = ||\nu * \nu||^2 > k^{-2}||\nu||^2 = k^{-2}|A|^{-1},$$

whence

$$E(A, A) \ge k^{-2}|A|^3,$$

and by Corollary 3.6 we find a Q as in (ii).

We now treat the case of a general ν . The idea is to throw away those points $x \in G$ in which ν takes extremal values, and consider the uniform distribution on the remaining set.

Thus, let $m = ||\nu||^2$, and consider the subsets of G,

$$B = \{ x \in G \mid \nu(x) \ge 5^2 k^2 m \}$$
$$C = \{ x \in G \mid \nu(x) \le m/5^2 k^2 \}.$$

We then split $\nu = \nu_+ + \nu_- + \nu_0$, where

$$\begin{split} \nu_+ &= \nu \cdot \mathbf{1}_B \\ \nu_- &= \nu \cdot \mathbf{1}_C \\ \nu_o &= \nu - (\nu_+ + \nu_-), \end{split}$$

We have

(3.14)
$$||\nu_{-}||^{2} = \sum_{x \in C} \nu(x)^{2} \le \frac{m}{5^{2}k^{2}} \sum_{x \in G} \nu(x) = \frac{m}{5^{2}k^{2}}.$$

Also, $1 \ge \sum_{x \in B} \nu(x) \ge |B| 5^2 k^2 m$; hence $|B| \le 1/5^2 k^2 m$, and by the Cauchy-Schwarz inequality:

(3.15)
$$||\nu_{+}||_{1}^{2} \leq |B| \cdot ||\nu_{+}||^{2} \leq |B|m \leq \frac{1}{5^{2}k^{2}}$$

By (3.14) and Young inequality (Lemma 2.26) we deduce

$$||\nu * \nu_{-}||^{2} \le ||\nu||_{1}^{2}||\nu_{-}||^{2} = ||\nu_{-}||^{2} \le \frac{m}{5^{2}k^{2}},$$

and the same for $||\nu_{-} * \nu||$. Similarly, from (3.15),

$$||\nu * \nu_+||^2 \le ||\nu_+||_1^2 ||\nu||^2 = ||\nu_+||_1^2 m \le \frac{m}{5^2 k^2},$$

and the same for $||\nu_+ * \nu||$.

By the triangle inequality, $||f+g|| \leq ||f|| + ||g||,$ we obtain,

$$||\nu * \nu_0|| \ge ||\nu * \nu|| - ||\nu * \nu_-|| - ||\nu * \nu_+|| \ge ||\nu * \nu|| - 2\frac{\sqrt{m}}{5k},$$

and again, on the other side,

(3.16)
$$||\nu_0 * \nu_0|| \ge ||\nu * \nu|| - 4\frac{\sqrt{m}}{5k}.$$

Suppose now $||\nu * \nu|| > \sqrt{m}/k$. Then from (3.16) we have

(3.17)
$$||\nu_0 * \nu_0|| > \frac{\sqrt{m}}{5k}$$

Now, let $A = G \setminus (B \cup C)$. Observe that $A = A^{-1}$ because ν is symmetrical, and consider the uniform distribution $\nu_A = |A|^{-1} \mathbf{1}_A$. The support of ν_0 is contained in A and so, for every $x \in G$,

$$\nu_A(x) \ge \frac{|A|^{-1}}{||\nu_0||_{\infty}} \nu_0(x)$$

where $||\nu_0||_{\infty} = \max_{x \in G} \nu_0(x) < 5^2 k^2 m$. Therefore, by (3.17),

$$(3.18) ||\nu_A * \nu_A|| \ge \frac{|A|^{-2}}{||\nu_0||_{\infty}^2} ||\nu_0 * \nu_0|| > \frac{|A|^{-2}}{||\nu_0||_{\infty}^2} \frac{\sqrt{m}}{5k} \ge \frac{|A|^{-2}\sqrt{m}}{5^5k^5m^2}.$$

Now, $1 \geq \sum_{x \in A} \nu(x) \geq |A| m/5^2 k^2,$ hence

$$|A|^{-1} \ge \frac{m}{5^2 k^2}.$$

Since $||\nu_A|| = \sqrt{|A|^{-1}}$, from (3.18) we obtain

(3.19)
$$||\nu_A * \nu_A|| > \frac{1}{5^8 k^8} ||\nu_A||.$$

We are now in the case treated at the beginning of the proof, with $k_1 = 5^8 k^8$ instead of k. Since k_1 is bounded by a polynomial in k, this concludes the proof.

LEMMA 3.9. Let ν be a symmetrical probability measure on G and $m \ge 1$. Write $M = \sup\{\nu^{(2m)}(H) \mid H < G\}$; then for every $n \ge m$, (i) $||\nu^{(n)}|| \le M^{1/4}$;

(ii)
$$\nu^{(n)}(Hg) \leq M^{1/2}$$
, for every proper subgroup $H < G$ and all $g \in G$.

PROOF. Fix $g \in G$ and a proper subgroup H of G. Let $x \in H$, then

$$\nu^{(2m)}(x) = (\nu^{(m)} * \nu^{(m)})(x) \ge \sum_{yg \in Hg} \nu^{(m)}(xyg)\nu^{(m)}(g^{-1}y^{-1}),$$

and so

$$\nu^{(2m)}(H) \ge \sum_{x,y \in H} \nu^{(m)}(xg)\nu^{(m)}(g^{-1}y) = \nu^{(m)}(Hg)\nu^{(m)}(g^{-1}H)$$

Now, symmetry of ν implies $\nu^{(m)}(g^{-1}H) = \nu^{(m)}(Hg)$, whence

$$\nu^{(m)}(Hg) < M^{1/2}$$

Let $n \ge m$; then $\nu^{(n)} = \nu^{(m)} * \nu^{(n-m)}$, and

$$\nu^{(n)}(Hg) \le \max_{y \in G} \{\nu^{(m)}(Hy)\} ||\nu^{(n-m)}||_1 \le M^{1/2},$$

thus establishing (ii). To prove (i), we apply (ii) specialized at H = 1;

$$||\nu^{(n)}||^2 = \sum_{g \in G} \nu^{(n)}(g)^2 \le M^{1/2} ||\nu^{(n)}||_1 = M^{1/2}.$$

The Bourgain-Gamburd machine. Before coming to the proof of the Theorem of Bourgain and Gamburd, let us remind a few elementary facts from sections 2.4 and 2.5.

Let S be a finite symmetric subset of a group G, with |S| = k, and A the adjacency operator associated to the Cayley graph $\Gamma[G, S]$; then for every $n \ge 1$ (Proposition 2.12) $\omega_n := A_{1,1}^n$ coincides with the number of distinct closed walks of length n starting at 1. If $\nu = \nu_S = |S|^{-1} \mathbf{1}_S$ is the probability measure uniformly concentrated at S, and $\widehat{A} = k^{-1}A$, then by what said in section 2.5,

$$\nu^{(n)}(1) = \widehat{A}_{1,1}^n = \frac{\omega_n}{k^n}$$

From this and the symmetrical property of ν , we get a simple consequence: for every $n \ge 1$,

(3.20)
$$||\nu^{(n)}||^2 = \sum_{x \in G} \nu^{(n)}(x)^2 = \sum_{x \in G} \nu^{(n)}(x^{-1})\nu^{(n)}(x) = \nu^{(2n)}(1) = \frac{\omega_{2n}}{k^{2n}}.$$

Since all diagonal entries of a power of A are equal, we finally have

(3.21)
$$||\nu^{(n)}||^2 = |G|^{-1} \frac{\operatorname{trace}(A^{2n})}{k^{2n}}$$

Now for the main result of this chapter. In the form of a general statement, it was formulated soon after the Bourgain and Gamburd original paper [1], were it appeared implicitly. We follow the version in Tao [35].

64

THEOREM 3.10 (Bourgain-Gamburd). Let S be a symmetric set of generators of the finite group G, with |S| = k. Suppose there exist constants $0 < \kappa, \lambda < 1 < \Lambda$ satisfying the following properties.

- (1) The degree of any non-trivial \mathbb{C} -representation of G is at least $|G|^{\lambda+\kappa}$.
- (2) For every $\delta > 0$ there exists $c(\delta)$ such that for every $|G|^{c(\delta)}$ -approximate subgroup Q of G, if $|G|^{\delta} \leq |Q| \leq |G|^{1-\delta}$ then $\langle Q \rangle$ is a proper subgroup of G.
- (3) There exists an even integer $n \leq \Lambda \log |G|$ such that

$$\sup_{H < G} \nu_S^{(n)}(H) < |G|^{-\lambda}.$$

Then the Cayley graph $\Gamma[G, S]$ is a two-sided α -expander, where $\alpha > 0$ depends only on $k, \lambda, \kappa, \Lambda$ and the function $c(\delta)$.

PROOF. Let G, S, λ, Λ be as in the assumptions, and write $\nu = \nu_S$. By (3) there exists $m \leq \frac{1}{2}\Lambda \log |G|$ such that

$$\nu^{(2m)}(H) < |G|^{-\lambda}$$

for every proper subgroup H of G. Lemma 3.9 then yields the following fact.

(•) For every $n \ge \frac{1}{2}\Lambda \log |G|$ one has

(3.22)
$$||\nu^{(n)}|| \le |G|^{-\lambda/4}$$

Moreover, for every proper subgroup H < G and every $g \in G$,

(3.23)
$$\nu^{(n)}(Hg) \le |G|^{-\lambda/2}$$

Now, let $0 < \varepsilon \leq \lambda R^{-1}/8$, where R is the constant in Lemma 3.8. For the next step, we suppose that for some $n \geq \frac{1}{2}\Lambda \log |G|$ we have

(3.24)
$$||\nu^{(n)}||^2 \ge |G|^{\lambda-1},$$

while, at the same time,

$$|\nu^{(2n)}|| > |G|^{-\varepsilon} ||\nu^{(n)}||.$$

Then, by Lemma 3.8 there exists a $|G|^{\varepsilon R}$ -approximate subgroup Q with

$$|G|^{-R\varepsilon}/||\nu^{(n)}|| \le |Q| \le |G|^{R\varepsilon}/||\nu^{(n)}||$$

and an element $x \in G$ such that

(3.25)
$$\nu^{(n)}(Qx) \ge |G|^{-\varepsilon R}.$$

Consequently, by (3.22) and (3.24) and the choice of ε ,

$$|G|^{\lambda/8} \le |Q| \le |G|^{\varepsilon R + \frac{1}{2} - \lambda/2} \le |G|^{1 - \lambda/8}$$

If we choose ε such that $\varepsilon R \leq c(\lambda/8)$, we then have by assumption (2) that $H = \langle Q \rangle$ is a proper subgroup of G. Therefore, by comparing (3.25) with (3.23),

 $|G|^{-\lambda/2} \ge \nu^{(n)}(Hx) \ge \nu^{(n)}(Qx) \ge |G|^{-\varepsilon R} \ge |G|^{-\lambda/8}$

which is a contradiction. We have thus proved the following fact.

(••) There exists $\varepsilon > 0$, depending only on λ and $c(\delta)$, with the prperty that, for every $n \ge \frac{1}{2}\Lambda \log |G|$, if

(3.26)
$$||\nu^{(n)}||^2 \ge |G|^{\lambda - 1}$$

then $||\nu^{(2n)}|| \le |G|^{-\varepsilon}||\nu^{(n)}||.$

Now, given $m \geq \frac{1}{2} \Lambda \log |G|$, suppose that $\nu^{(m)}, \nu^{(2m)}, \ldots, \nu^{(2^t m)}$ satisfy (3.26); then

(3.27)
$$|||\nu^{(2^t m)}|| \le |G|^{-\varepsilon t} ||\nu^{(m)}||$$

This tells us that there exists $n \ge \Lambda \log |G|$ such that

(3.28)
$$||\nu^{(n)}||^2 < |G|^{\lambda-1}.$$

Observe that the magnifying factor 2^t in (3.27) that leads to inequality (3.28) depends only on ε and λ . It follows that a smallest *n* satisfying (3.28) may be found so that

$$(3.29) n \le C \log|G|$$

where the constant C depends only on the parameters Λ , λ and $c(\delta)$.

We are now going to use the remark which precedes the statement. Let A be the adjacency operator of $\Gamma[G, S]$ and μ the largest absolute value of an eigenvalue $\neq k$ of A. If \mathfrak{m} is the multiplicity of μ , then by Proposition 3.7 and assumption (1),

(3.30)
$$\mathfrak{m} \ge |G|^{\lambda + \kappa}$$

Let n as in (3.28); by (3.21) we have

$$|G|^{\lambda-1} > ||\nu^{(n)}||^2 = |G|^{-1} \frac{\operatorname{tr}(A^{2n})}{k^{2n}} \ge |G|^{-1} \frac{\mathfrak{m}\mu^{2n}}{k^{2n}} \ge |G|^{\lambda+\kappa-1} \frac{\mu^{2n}}{k^{2n}};$$

therefore

$$\left(\frac{\boldsymbol{\mu}}{k}\right)^{2n} < |G|^{-\kappa}.$$

Hence, by (3.29) (and since $\mu/k < 1$),

$$\left(\frac{\boldsymbol{\mu}}{k}\right)^{C\log|G|} < |G|^{-\kappa},$$

which implies

$$\mu < k e^{-\kappa/C}.$$

Therefore $k - \mu > k(1 - e^{-\kappa/C}) > 0$, and the proof is complete.

The three assumptions in the statement of Theorem 3.10 are called, respectively:

- (1) Quasirandomness;
- (2) Product theorem;
- (3) Non-concentration.

As it has been figuratively observed (e.g. by Gowers and others), these properties conduct, in the reverse order, three stages in the evolution of the distributions $\nu^{(n)}$. Quoting T. Tao [**35**]: "In the early stage $n = o(\log |G|)$ the non-concentration hypotheses creates some initial spreading of this random walk, in particular ensuring that the walk "escapes" from cosets of proper subgroups. In the middle stage $n \sim \log |G|$, the product theorem steadily flattens the distribution of the random walk, until it is very roughly comparable to the uniform distribution. Finally, in the late stage $n \gg \log |G|$, the quasirandomness property can smooth out the random walk almost completely to obtain the mixing necessary for expansion."

The term 'flattening', in Lemma 3.8 and in Tao's words, expresses the fact that the decreasing in norm, $||\nu^{(2n)}|| \leq k^{-1}||\nu^{(n)}||$, means that the evolving distribution takes non-zero, but progressively smaller, values on larger subsets, thus approaching a uniform spreading.

3.4. Quasirandom groups

The term *quasirandomness*, to denominate the first assumption in 3.10, deserves a bit of attention. When applied to groups this terminology was introduced by Gowers [13]. We give the most direct definition for finite groups.

DEFINITION 3.11. Let $d \ge 1$. A finite group G is said to be *d*-quasirandom if every non-trivial irreducible complex representation of G has degree at least d.

Of course, once again, the quantitative bound d is most important in the definition, and the property is interesting (that is, it makes some difference) when d is large with respect of |G|. In particular we like d to be around $|G|^{\alpha}$ for a fixed $\alpha \leq 1$ while the groups G belong to some infinite family.

In this perspective, condition (1) in Theorem 3.10 may be restated by asking that G is $|G|^{\lambda}$ -quasirandom.

Next lemma (usually called *Gowers' trick*), besides being very handy, shows at an initial level the utility that a quasirandomness assumption may have. Although the Lemma works for any subset S, we assume for simplicity that S is symmetric.

LEMMA 3.12 (Gowers [13], Nikolov and Pyber [26]). Let $d \ge 1$, G a finite dquasirandom group and S a symmetric subset of G. If $d^{1/3} \ge |G|/|S|$, then $S^3 = G$.

PROOF. Write k = |S|, let A be the adjacency operator in the Cayley graph $\Gamma[G, S]$ (there is no harm in assuming $1 \notin S$), and μ the largest absolute value

of an eigenvalue $\neq k$ of A. Then, by Proposition 3.7 and the d-quasirandomness assumption,

$$k|G| = tr(A^2) \ge d\mu^2 \ge \frac{|G|^3}{k^3}\mu^2,$$

whence

$$\mu^2 \le \frac{k^4}{|G|^2}.$$

We complete the indicator function $\mathbf{1}_S$ to a zero-sum function g, that is we let

$$g = \mathbf{1}_S - \frac{k}{|G|} \mathbf{1}_G.$$

Recalling (2.18), we have

$$Ag = g * \mathbf{1}_S = \mathbf{1}_S * \mathbf{1}_S - \frac{k^2}{|G|} \mathbf{1}_G.$$

Since $g \in \mathbb{Z}^{\perp}$ (the space of zero-sum elements of $\ell^2(G)$), we have by the Rayleigh bound:

(3.32)
$$||g * \mathbf{1}_{S}|| = ||Ag|| = \sqrt{\langle Ag, Ag \rangle} = \sqrt{\langle A^{2}g, g \rangle} \le \sqrt{\mu^{2}||g||^{2}} \le \frac{k^{2}}{|G|}||g||.$$

Now, as g and $(k/|G|)\mathbf{1}_S$ are orthogonal,

$$||g||^2 = ||\mathbf{1}_S||^2 - ||\frac{k}{|G|}\mathbf{1}_G||^2 = k - \frac{k^2}{|G|} < k,$$

hence,

(3.33)
$$||g * \mathbf{1}_S|| < \frac{k^{5/2}}{|G|}.$$

Let $x \in G$, then by Cauchy-Schwarz,

$$(g * \mathbf{1}_S * \mathbf{1}_S)(x)^2 = \left(\sum_{y \in S} g * \mathbf{1}_S(xy^{-1})\right)^2 \le k \sum_{y \in S} |g * \mathbf{1}_S(xy^{-1})|^2 \le k ||g * \mathbf{1}_S||^2,$$

and so, from (3.33),

(3.34)
$$|(g * \mathbf{1}_S * \mathbf{1}_S)(x)| < \frac{k^3}{|G|}$$

On the other hand,

$$g * \mathbf{1}_S * \mathbf{1}_S = \mathbf{1}_S^{(3)} - \frac{k^3}{|G|}$$

Now, (3.34) tells us that this function never takes, in absolute value, the value $k^3/|G|$, that is $\mathbf{1}_S^{(3)}(x) \neq 0$ for every $x \in G$. This finishes the proof because the support of $\mathbf{1}_S^{(3)}$ is contained in S^3 .

68

Representations of $SL_2(p)$. We now prove a theorem of Frobenius which shows quasirandomness of groups of type $SL_2(p)$, with p a prime; in fact, it implies that given $\varepsilon > 0$, then for all sufficiently large primes p the group $G = SL_2(p)$ is $|G|^{1/3+\varepsilon}$ -quasirandom.

In the proof, we use the well known fact (see Proposition 4.3) that the only proper non-trivial normal subgroup of $SL_2(q)$, for $q \ge 4$ a prime power, is the centre $Z = \{1, -1\}$ (where 1 is the identity matrix). In fact, Frobenius Theorem holds for $SL_2(q)$ and any prime-power $q \ge 4$; the prime case is however enough to our purposes.

THEOREM 3.13 (Frobenius). Let $p \ge 5$ be a prime. Then every non-trivial irreducible \mathbb{C} -representation of $SL_2(p)$ has degree at least (p-1)/2.

PROOF. Let $G = SL_2(p)$; let V be a C-vector space, with dim $V \ge 1$, and $\pi : G \to GL(V)$ a non-trivial representation of G on V. Since π is not trivial, $K = \ker(\pi)$ is a proper normal subgroup of G, and so, by what we reminded above, K = 1 or K = Z. For $g \in G$ and $v \in V$ we write $g \cdot v$ for $\pi(g)(v)$. Let

$$y = \left(\begin{array}{cc} 1 & 1\\ 0 & 1 \end{array}\right),$$

then

$$\langle y \rangle = U = \left\{ \left(\begin{array}{cc} 1 & b \\ 0 & 1 \end{array} \right) \middle| b \in \mathbb{Z}/p\mathbb{Z} \right\},$$

(the standard unipotent subgroup - see section 4.2) is a cyclic group of order p. Since $Z \cap U = 1$, we have $\ker(\pi) \cap U = 1$, thus the restriction of π to U is injective, and so it is an embedding of the group U in the group GL(V). Let $\zeta \neq 1$ be an eigenvalue for $\pi(y)$ on V; then $\zeta \in \mathbb{C}$ is a primitive p-th root of unity. Let $0 \neq v \in V$ with $y \cdot v = \zeta v$.

For $1 \le a \le p-1$, let $x(a) = \begin{pmatrix} \bar{a} & 0 \\ 0 & \bar{a}^{-1} \end{pmatrix}$, then x(a) normalizes U in $SL_2(p)$, and in fact (see exercise 35),

$$y^{x(a)} = \left(\begin{array}{cc} 1 & \bar{a}^2 \\ 0 & 1 \end{array}\right) = y^{a^2}.$$

Write $v_a = x(a) \cdot v$; then we have

$$y \cdot (v_a) = yx(a) \cdot v = x(a) \cdot (y^{x(a)} \cdot v) = x(a) \cdot (y^{a^2} \cdot v) = x(a) \cdot (\zeta^{a^2} v) = \zeta^{a^2} v_a.$$

Since ζ is a primitive *p*-th root of unity and there are (p-1)/2 distinct squares modulo *p*, we conclude that *y* (that is $\pi(y)$) has at least (p-1)/2 distinct eigenvalues, whence in particular dim $V \ge (p-1)/2$.

COROLLARY 3.14. Let $p \ge 5$ be a prime, and let S be a symmetric subset of $G = SL_2(p)$ with $|S| \ge 2|G|^{8/9}$; then $S^3 = G$.

PROOF. Let G and S be as in the assumptions. Then, since $|G| = p(p^2 - 1)$,

$$\frac{|G|^3}{|S|^3} \le \frac{|G|^{1/3}}{2^3} \le \frac{q+1}{2^3} \le \frac{p-1}{2}$$

and so, by Theorem 3.13 and Lemma 3.12, $S^3 = G$.

In general, it has been proved (by Landazuri and Seitz) that if G is a finite simple group of Lie-type then G is $|G|^{\lambda(r)}$ -quasirandom, where $\lambda(r)$ depends only on the Lie rank r of the group G.

On the other hand, for every $n \geq 5$ the alternating group $A_n = Alt(n)$ admits a non-trivial representation, induced by the natural permutation representation, of degree n; indeed, it not difficult to see that A_n admits an irreducible (non-trivial) representation of degree n - 1. Hence the largest d for which A_n is d-quasirandom is n - 1 (this is in fact the exact bound for $n \geq 6$, while A_5 is only 3-quasirandom). Therefore, for the class of alternating groups there exists no uniform λ such that $G = A_n$ is $|G|^{\lambda}$ -quasirandom.

Note. The treatment of the Balog-Szemeredi-Gowers Lemma in the first section essentially follows that in the text of Tao [35]. Similarly, I have arranged from [35] the statement and the proof of the Bourgain-Gamburd Theorem (a similar, but not exactly the same, rendition you may find, for instance, in Breuillard [4]), although substantial help I received from the consultation of Bourgain-Gamburd original paper [1], and of Helfgott's [19] and Breuillard's [3] surveys.

Tao's book also includes a chapter that explains expansion check by means of Kazhdan property (T), and another one about the number theoretical interplay between the two methods.