11.1.14 Let G be a subgroup of a finite additive group Z, and let $f : Z \to \mathbf{C}$ and $d \geq 1$. For each coset $y + G$ of G, define $\|f\|_{U^d(y+G)}$ in the obvious translation-invariant manner. Show that

$$\|f\|_{U^d(Z)} \leq \left(\mathbf{E}_{y\in Z}\|f\|_{U^d(y+G)}^{2^d/(d+1)}\right)^{(d+1)/2^d}$$

thus generalizing the previous exercise and demonstrating that "local uniformity norms" control "global uniformity norms".

11.1.15 Let $f : Z \to \mathbf{C}$ be a function on a finite additive group Z. Establish the Parseval-type identity $\|f\|_{U^3(Z)} = |Z|^{1/2}\|\hat{f}\|_{U^3(Z)}$, which shows that the Fourier transform does not simplify the U^3 norm. (This phenomenon is related to the fact that the Fourier transform of a Gaussian is again a Gaussian.) Deduce a similar Plancherel-type identity for the U^3 inner product. For the higher U^d norms, $d \geq 4$, the situation is even worse; the Fourier representation is more complicated than the spatial representation.

11.1.16 Let A be a subset of a finite additive group Z, and let $d \geq 1$. Using (11.7), show that there are at least $\mathbf{P}_Z(A)^{2^d}|Z|^{d+1}$ $d+1$-tuples $(x, h_1, \ldots, h_d) \in Z^{d+1}$ such that the cube $x + [0,1]^d \cdot (h_1, \ldots, h_d)$ is contained in A. Use this to give another proof of Lemma 10.49.

11.1.17 Let $f : Z \to \mathbf{C}$ be a random function such that the random variables $f(x)$ for $x \in Z$ are jointly independent, have mean zero, and are bounded by 1. Show that $\mathbf{E}\|f\|_{U^d(Z)}^{2^d} = O_d(1/|Z|)$ for all $d \geq 1$. Thus random balanced functions tend to be Gowers uniform of very high order.

11.2 Hard obstructions to uniformity

In this section we consider the following *inverse problem*: suppose $f : Z \to \mathbf{C}$ is a function bounded in magnitude by one which fails to be Gowers uniform of some order $k-2$, say $\|f\|_{U^{k-2}(Z)} \geq \delta$ for some $0 < \delta \leq 1$. What structural information can one then conclude about f? As it turns out, a sufficiently strong answer to this question will lead to a proof of Szemerédi's theorem for progressions of length k. This is the strategy employed by Gowers [137], [138] in his proof of Szemerédi's theorem, with the focus on obtaining as strong an inverse theorem for the $U^{k-2}(Z)$ norm as possible. In Section 11.4 we describe a slightly different approach in which one obtains a much weaker (and easier to prove) inverse theorem, but one which is still sufficient to obtain Szemerédi's theorem (but with much worse quantitative bounds).

A good model case is provided by the case $k = 3$. From (11.9) we see that if $\|f\|_{U^2(Z)} \geq \delta$, then $\|f\|_{u^2(Z)} \geq \delta^2$, and hence there exists a linear phase function $g(x) := e(\xi \cdot x)$ which has a large inner product with f: $|\langle f, g\rangle_{L^2(Z)}| \geq \delta^2$. This fact, combined with (11.8), can be used to give a variant of Proposition 10.10 or Proposition 10.11, which in turn can be employed in either a density increment argument or energy increment argument to prove the $k = 3$ case of Szemerédi's theorem, as was done in Sections 10.2, 10.3 and Section 10.5 respectively. One can view these linear phase functions as being the *obstructions* to Gowers uniformity of order 1; we have just seen that failure of Gowers uniformity of order 1 implies correlation with one of these linear phase functions, and conversely the other inequality in (11.9) implies that correlation with a linear phase function implies lack of Gowers uniformity of order 1.

This model case, combined with the observations in Exercise 11.1.12, suggest that, more generally, lack of Gowers uniformity of order $k - 2$ should be tied to correlation with a phase function which is somehow polynomial of degree $k - 2$. This can be made precise as follows.

Definition 11.5 (Polynomial bias) Let Z be a finite additive group, and let $\phi : Z \to \mathbf{R}/\mathbf{Z}$ be a phase function. Given any $h \in Z$, we define the *difference operator* $(h \cdot \nabla)$ applied to ϕ as

$$(h \cdot \nabla)\phi(x) := \phi(x + h) - \phi(x).$$

We will sometimes subscript ∇ by ∇_x to emphasize the variable being differenced over (in case ϕ also depends on some other variables). If $d \geq 1$, we say that ϕ is a *phase polynomial of degree less than d* if we have

$$(h_1 \cdot \nabla_x) \cdots (h_d \cdot \nabla_x)\phi(x) = 0 \text{ for all } x, h_1, \ldots, h_d \in Z.$$

Phase polynomials of degree less than 2 will be referred to as *linear*, phase polynomials of degree less than 3 will be referred to as *quadratic*, and so forth. If $f : Z \to \mathbf{C}$ is a function, then we define the *polynomial bias* of f of degree d to be the quantity

$$\|f\|_{u^d(Z)} := \sup_{\phi} \left|\langle f, e(\phi)\rangle_{L^2(Z)}\right| = \sup_{\phi} |\mathbf{E}_{x\in Z} f(x)e(-\phi(x))|$$

where ϕ ranges over all phase polynomials of degree less than d.

More generally, if $B \subset Z$ is non-empty, we say that $\phi : B \to \mathbf{R}/\mathbf{Z}$ is a *locally polynomial phase function of degree less than d* if we have

$$(h_1 \cdot \nabla_x) \ldots (h_d \cdot \nabla_x)\phi(x) = 0 \text{ whenever } x + [0, 1]^d \cdot (h_1, \ldots, h_d) \subseteq B,$$

and then define

$$\|f\|_{u^d(B)} := \sup_{\phi} \left|\langle f, e(\phi)\rangle_{L^2(B)}\right| = \sup_{\phi} |\mathbf{E}_{x\in B} f(x)e(-\phi(x))|$$

where ϕ ranges over all phase functions which are locally polynomial on B of degree less than d.

To illustrate this definition, first observe that the only phase polynomials of degree less than 1 are the constants $\phi(x) = c$, and hence

$$\|f\|_{u^1(Z)} = |\mathbf{E}(f)| = \|f\|_{U^1(Z)}. \tag{11.10}$$

Thus $\|\cdot\|_{u^1(Z)}$ is a seminorm. For $d > 1$, one easily verifies that $\|\cdot\|_{u^d Z}$ is a genuine norm. For instance, from Exercise 4.1.4, we see that the only phase polynomials of degree less than 2 are the linear phases $\phi(x) = \xi \cdot x + c$, and thus the definition of the u^2 norm matches the one given in (10.5). In particular we still have the relation (11.9).

More generally, the $u^d(Z)$ and $U^d(Z)$ norms are quite related, enjoying the same symmetries. For instance, if ϕ is a phase polynomial of degree less than d, then one can easily verify that

$$\|fe(-\phi)\|_{u^d(Z)} = \|f\|_{u^d(Z)}; \quad \|fe(-\phi)\|_{U^d(Z)} = \|f\|_{U^d(Z)} \tag{11.11}$$

In particular, from (11.7) we have

$$|\mathbf{E}_{x\in Z} f(x)e(-\phi(x))| = \|fe(-\phi)\|_{U^1(Z)} \leq \|fe(-\phi)\|_{U^d(Z)} = \|f\|_{U^d(Z)}$$

and hence on taking suprema we have

$$\|f\|_{u^d(Z)} \leq \|f\|_{U^d(Z)}.$$

Thus correlation with a phase function of degree less than d implies lack of Gowers uniformity of order $d - 1$. In light of (11.10), (11.9), one may hope that a converse statement is true, namely that lack of Gowers uniformity of order $d - 1$ implies a correlation with a phase of degree less than d. One hopeful sign in this direction is the identity

$$\|e(\phi)\|_{U^d(Z)}^{2^d} = \mathbf{E}_{x,h_1,\ldots,h_d\in Z} e((h_1 \cdot \nabla_x) \cdots (h_d \cdot \nabla_x)\phi(x)) \tag{11.12}$$

whose verification we leave as an exercise. This suggests, though does not quite prove, that a function has large $U^d(Z)$ norm if and only if its phase is approximately polynomial of degree less than d. The above statement would then be an assertion that a phase which is approximately polynomial of degree d, in fact correlates with a genuine polynomial of degree d. Such an assertion should remind one of the Balog–Szemerédi–Gowers theorem, Theorem 2.29, and in fact that theorem plays a key role in establishing facts such as these.

In the case when Z is a vector space over a finite field of small order and $k = 4$, we can formalize these conjectures affirmatively as follows.

Theorem 11.6 (Inverse theorem for $U^3(F^n)$) *[137], [160] Let Z be a vector space over a finite field F, let $f : Z \to \mathbf{C}$ have magnitude bounded by* 1, *such that $\|f\|_{U^3(Z)} \geq \eta$ for some $0 < \eta \leq 1$. Then there exists a subspace $W \subset Z$*

with

$$\dim_F(W) \geq \dim_F(Z) - O\big(\eta^{-O(1)}\big) \tag{11.13}$$

such that

$$\mathbf{E}_{y\in Z}\|f\|_{u^3(y+W)} = \Omega\big(\eta^{O(1)}\big). \tag{11.14}$$

In particular, there exists $y \in Z$ such that $\|f\|_{u^3(y+W)} \geq \Omega(\eta^{O(1)})$.

The proof of this inverse theorem is quite lengthy, using techniques from previous chapters as well as a heavy reliance on Fourier-analytic methods and the van der Corput lemma (Lemma 11.3), and will be deferred to the next section. We remark that the case when F has characteristic 2 was not quite dealt with in the above-cited papers, but requires an additional observation of Samorodnitsky (private communication). Assuming it for now, we can now prove Szemerédi's theorem for vector spaces Z and in the case $k = 4$. In fact the inverse theorem allows us to give both a density increment proof and an energy increment proof. The density increment proof is based on the following proposition, analogous to (though somewhat weaker than in some respects) Lemma 10.15.

Proposition 11.7 (Lack of uniformity implies density increment) *Let Z be a vector space over a finite field F of odd prime order, and let $f : Z \to \mathbf{C}$ have magnitude bounded by 1 be such that $\mathbf{E}_Z(f) = 0$ and $\|f\|_{U^3(Z)} \geq \eta$ for some $0 < \eta \leq 1$. Then there exists a subspace Z' of Z with*

$$\dim_F(Z') \geq \frac{1}{2}\dim(Z) - O\big(\eta^{-O(1)}\big)$$

and a point $x_0 \in Z$, such that

$$\mathbf{E}_{x\in x_0+Z'} f(x) \geq \Omega\big(\eta^{O(1)}\big).$$

Proof From Theorem 11.6 we can find a subspace W obeying the dimension bound (11.13) and the correlation bound (11.14). Also note that if $\mathbf{E}_{y+W}(f) \geq \Omega(\eta^{O(1)})$ for even a single $y \in Z$ then we will be done, so we may take $\mathbf{E}_{y+W}(f) \leq c\eta^C$ for any given absolute constants $c, C > 0$. Since we also have

$$\mathbf{E}_{y\in Z}\mathbf{E}_{y+W}(f) = \mathbf{E}_Z(f) = 0$$

we conclude that

$$\mathbf{E}_{y\in Z}|\mathbf{E}_{y+W}(f)| = 2\mathbf{E}_{y\in Z}\max(\mathbf{E}_{y+W}(f), 0) \leq 2c\eta^C$$

and so from (11.14) we have (choosing the constants c, C appropriately)

$$\mathbf{E}_{y\in Z}\|f\|_{u^3(y+W)} - 2|\mathbf{E}_{y+W}(f)| = \Omega\big(\eta^{O(1)}\big).$$

In particular we can find $y \in Z$ such that

$$\|f\|_{u^3(y+W)} \geq 2|\mathbf{E}_{y+W}(f)| + \Omega\big(\eta^{O(1)}\big).$$

By translating f by y if necessary we may take $y = 0$. By definition of the u^3 norm and Exercise 11.2.6, we can thus find a self-adjoint linear operator $M : W \to W$ and $\xi \in W$ such that

$$|\mathbf{E}_{x\in W} f(x)e(-Mx\cdot x)e(-\xi\cdot x)| \geq 2|\mathbf{E}_W(f)| + \Omega\big(\eta^{O(1)}\big).$$

Observe that the quantity $Mx\cdot x + \xi\cdot x$ only takes at most $|F|$ values. If we thus partition W into $|F|$ level sets $S_1, S_2, \ldots, S_{|F|}$, each of the form $\{x \in W : Mx\cdot x + \xi\cdot x = \text{const}\}$, then we have from the triangle inequality that

$$\sum_{j=1}^{|F|} |\mathbf{E}_{x\in W} 1_{S_j}(x)f(x)| \geq 2\left|\sum_{j=1}^{|F|} \mathbf{E}_W(1_{S_j}(x)f(x))\right| + \Omega\big(\eta^{O(1)}\big)$$

and hence, by the identity $\max(y, 0) = (|y| + y)/2$,

$$\sum_{j=1}^{|F|} \max(\mathbf{E}_{x\in W} 1_{S_j}(x)f(x), 0) > \Omega\big(\eta^{O(1)}\big)$$

and so by the pigeonhole principle we can find j such that

$$\mathbf{E}_{x\in W} 1_{S_j}(x)f(x) > \Omega\big(\eta^{O(1)}\big)\mathbf{P}_W(S_j).$$

Now we need to take the quadratic surface S_j and partition it into affine spaces. We first observe that there exists a subspace U of W with dimension

$$\dim_F(U) \geq \frac{1}{2}\dim_F(W) - \frac{3}{2} \geq \frac{1}{2}\dim_F(W) - O\big(\eta^{-O(1)}\big)$$

which is null with respect to M: see Exercise 4.3.16. Splitting S_j into cosets of U, we see from the pigeonhole principle that there exists a coset $x_1 + U$ such that

$$\mathbf{E}_{x\in x_1+U} 1_{S_j}(x)f(x) > \Omega\big(\eta^{O(1)}\big)\mathbf{P}_{x_1+U}(S_j),$$

so in particular $S_j \cap (x_1 + U)$ is non-empty and

$$\mathbf{E}_{x\in S_j\cap(x_1+U)} f(x) > \Omega\big(\eta^{O(1)}\big).$$

The point of working on a coset $x_1 + U$ of a null space is that the quantity $Mx\cdot x + \xi\cdot x$ becomes linear with respect to x. Thus the intersection of S_j with $x_1 + U$ is an affine subspace $x_0 + Z'$ of $x_1 + U$ of codimension at most 1. The claim follows. □

Iterating this proposition as in the proof of Roth's theorem, one can eventually deduce the bound

$$r_4(F^n) = O\left(\frac{|F|^n}{\log^c n}\right) \tag{11.15}$$

for all $n > 1$ and some absolute constant c. It is also possible to adapt the energy increment argument from Section 10.5, with the the concept of quasi-periodic being replaced with that of being determined by a bounded number of quadratic phase functions, however the bounds on $r_4(F^n)$ obtained this way are rather poor. One can do a bit better by adapting the argument in Theorem 10.27, obtaining the bound

$$r_4(F^n) = O\left(\frac{|F|^n}{n^c}\right);$$

see [161].

It is likely that the above inverse theory extends to higher values of k, but there are some technical difficulties in carrying this out, and this has not yet been achieved at this time of writing.

Given the success of the inverse U^3 approach to establish in the finite field case, one then is led to see whether a similar inverse theorem holds for other groups, such as cyclic groups $\mathbf{Z}_N$. Here one encounters an interesting phenomenon, which is that the quadratic phase functions on $\mathbf{Z}_N$ do *not* form a complete set of obstructions to Gowers uniformity of order 2. An example is given as follows.

Proposition 11.8 (Furstenberg–Weiss example) *Let N be a large integer, and let $M := \lfloor\sqrt{N}\rfloor$, and let α be an irrational number obeying the diophantine condition $\|n\alpha\|_{\mathbf{R}/\mathbf{Z}} = \Omega(n^{-C})$ for some constant $C > 0$. Define the function $f : \mathbf{Z}_N \to \mathbf{C}$ by $f(x) := e(\alpha\lfloor x/M\rfloor^2)$ when $x \in [0, M/10) + M \cdot [0, M/10)$, and $f(x) := 0$ otherwise. Then $\|f\|_{U^3(\mathbf{Z}_N)} = \Theta(1)$, but $\|f\|_{u^3(\mathbf{Z}_N)} = o_{N\to\infty;\alpha}(1)$.*

As the name implies, this example was essentially discovered by Furstenberg and Weiss[126], though in a substantially different language to that presented here (they constructed a characteristic factor for quadruple recurrence which was not given by quadratic eigenfunctions).

Proof (Sketch) We can write $f(x) = e(\phi(x))1_P(x)$, where P is the progression $P := [0, M/10) + M \cdot [0, M/10)$ and ϕ is the phase function $\phi(x) := \alpha\lfloor x/M\rfloor^2$. One can easily verify that ϕ is locally quadratic on ϕ, and hence by (11.7)

$$\|f\|_{U^3(\mathbf{Z}_N)} = \|1_P\|_{U^3(\mathbf{Z}_N)} \geq \|1_P\|_{U^1(\mathbf{Z}_N)} = \mathbf{P}_{\mathbf{Z}_N}(P) = \Theta(1).$$

On the other hand, since f is bounded by 1, we have $\|f\|_{U^3(\mathbf{Z}_N)} \leq 1$. Thus $\|f\|_{U^3(\mathbf{Z}_N)} = \Theta(1)$ as claimed.

To prove the second claim, we see from (11.2.2) that it suffices to show that

$$\mathbf{E}_{x\in\mathbf{Z}_N} e(\phi(x) + (c_2x^2 + c_1x + c_0)/N)1_P(x) = o_{N\to\infty;\alpha}(1)$$

for all integers c_0, c_1, c_2. Writing $x = yM + z$ for $y, z \in [0, M/10)$, it suffices to show that

$$\mathbf{E}_{y,z\in[0,M/10)} e(\alpha y^2 + (c_2(yM + z)^2 + c_1(yM + z) + c_0)/N) = o_{N\to\infty;\alpha}(1).$$

To estimate this sum one has two choices. Either one can apply van der Corput's lemma (Exercise 11.2.9) twice in the y variable (with $H_1 = M^{1-\varepsilon}$ and $H_2 = M^{1-2\varepsilon}$ for some small ε), and reduce to showing that

$$\mathbf{E}_{h_1 \in [1,H_1], h_2 \in [1,H_2]} e(2(\alpha + c_2 M^2) h_1 h_2) = o_{N \to \infty; \alpha}(1);$$

or one can apply van der Corput's lemma once in the y variable and once in the z variable to reduce to showing that

$$\mathbf{E}_{h_1 \in [1,H_1], h_2 \in [1,H_2]} e(2 c_2 M h_1 h_2) = o_{N \to \infty; \alpha}(1).$$

While neither of these two bounds holds uniformly in c_2, it turns out that one of the two bounds is always true, the latter in the "minor arc" case when $c_2 M$ is not within $M^{-2+O(\varepsilon)}$ to being a rational with denominator at most $M^{O(\varepsilon)}$, and the former in the complementary "major arc" case. The exact verification of the bounds requires some basic machinery from Diophantine approximation, but we omit it as it is somewhat messy. □

This example shows that in addition to the *globally quadratic* phase obstructions that appeared in the finite field case, we now must consider *locally quadratic* phase obstructions, which are only defined on a suitable progression in the group such as $[0, M/10) + M \cdot [0, M/10)$. One can alternatively replace progressions with Bohr sets, which are of course closely related (cf. Section 4.4). A typical inverse theorem in this setting is as follows.

Theorem 11.9 (Inverse theorem for $U^3(Z)$) *[160] Let Z be an finite additive group of odd order, let $f : Z \to \mathbf{C}$ be a function bounded in magnitude by 1, such that $\|f\|_{U^3(Z)} \geq \eta$. Then there exists a regular Bohr set $B := B(S, \rho)$ in G with $|S| \leq O(\eta^{-O(1)})$ and $\rho = \Omega(\eta^{O(1)})$ such that*

$$\mathbf{E}_{y \in Z} \|f\|_{u^3(y+B)} = \Omega\big(\eta^{O(1)}\big). \tag{11.16}$$

In particular, there exists $y \in Z$ such that $\|f\|_{u^3(y+B)} = \Omega(\eta^{O(1)})$.

The proof of this theorem is similar to that of Theorem 11.6 which we give below, but is somewhat more complicated as we must work with (regular) Bohr sets instead of subspaces (which ultimately arises from the application of a version of Chang's theorem, Theorem 4.42, for arbitrary groups). It can then be used to prove

Proposition 11.10 (Lack of uniformity implies density increment) *[137], [138] Let $Z = \mathbf{Z}_N$ be a cyclic group of odd prime order, and let $f : Z \to \mathbf{C}$ have magnitude bounded by 1 be such that $\mathbf{E}_Z(f) = 0$ and $\|f\|_{U^3(Z)} \geq \eta$ for some $0 < \eta \leq 1$. If $N \geq \exp(O(\eta^{-O(1)}))$, then there exists a proper arithmetic* *progression P in Z of length $|P| = \Omega(N^c)$ for some absolute constant* $0 < c < 1$ *such that*

$$\mathbf{E}_{x \in P} f(x) \geq \Omega\big(\eta^{O(1)}\big).$$

This result was first established by Gowers[1] [137], [138] without directly proving an inverse theorem. However, the method of proof of Theorem 11.9 in [160] is based almost entirely the techniques used in [137] to establish Proposition 11.10. By the usual iteration arguments, this proposition can be used to establish the bound

$$r_4(\mathbf{Z}_N) = O\left(\frac{N}{(\log\log N)^c}\right)$$

for some absolute constant $0 < c < 1$ and all large N; this is the best bound on $r_4(\mathbf{Z}_N)$ known to date. See [137], [138], [160] for further discussion. In a similar spirit, Theorem 11.9 can eventually be used to establish the more general result

$$r_4(Z) = O\left(\frac{|Z|}{(\log\log |Z|)^c}\right)$$

for any large finite additive group Z; see [160]. It seems likely that this bound can be improved to $O(\frac{|Z|}{\log^c |Z|})$ by using the arguments in Theorem 10.27 or Theorem 10.30 but this will probably be quite messy.

Exercises

11.2.1 Prove (11.11).

11.2.2 Let $\mathbf{Z}_N$ be a cyclic group (and thus also a ring), and let $\phi : \mathbf{Z}_N \to \mathbf{R}/\mathbf{Z}$ be a phase polynomial of degree less than d. Show that there exist $c_0, c_1, \ldots, c_{d-1} \in \mathbf{Z}_N$ such that $\phi(x) = (c_{d-1}x^{d-1} + \cdots + c_1 x + c_0)/N$ for all $x \in \mathbf{Z}_N$, where the map $x \mapsto x/N$ is defined from $\mathbf{Z}_N$ to $\mathbf{R}/\mathbf{Z}$ in the obvious manner. Conversely, every function of this form is a phase polynomial of degree less than d. Thus in the cyclic case, the concept of a phase polynomial collapses to the usual definition of a polynomial.

11.2.3 Prove (11.12). (You may need to reflect some of the variables or take conjugates to eliminate a $(-1)^d$ factor.)

11.2.4 Let $f : Z \to \mathbf{C}$ be a function bounded in magnitude by 1, and let $d \geq 1$. Show that $\|f\|_{u^d(Z)}, \|f\|_{U^d(Z)} \leq 1$, and that $\|f\|_{u^d(Z)} = 1$ if and only if $\|f\|_{U^d(Z)} = 1$.

---

[1] The original argument in [137] had a exponential dependence on η rather than a polynomial one for $\mathbf{E}_{x \in P} f(x)$, leading ultimately to the weaker bound of $O(\frac{N}{(\log\log\log N)^c})$ for $r_4(\mathbf{Z}_N)$. This is due to a reliance on Freiman's theorem instead of a Chang–Bogulybov type theorem; the problem being that the Freiman theorem employed (essentially Theorem 5.32) suffers an exponential loss in an unfavorable location.