almost periodic. You may need the fact that Volterra integral operators are compact.)

11.5.9 (Koopman–von Neumann theorem) Let $(X, \mathcal{B}, \mathbf{P}, T)$ be measure-preserving system, and let $f \in L^2(X)$. Show that there is a unique decomposition $f = f_{U^\perp} + f_U$, where $\|f_U\|_{U^2(X)} = 0$ and $f_{U^\perp}$ is the limit in $L^2(X)$ of almost periodic functions.

11.6 The hypergraph approach

In Section 10.6 we saw that the Szemerédi regularity lemma led to a result in graph theory, namely the *triangle removal lemma*, which in turn implied Roth's theorem (as well as a generalization to right-angled triangles). It is then natural to ask whether a similar approach can prove Szemerédi's theorem for more general k. This turns out to be the case, but requires one work with *hypergraphs* (also known as *set systems*) instead of graphs. We need some notation. If A is a finite set and $k \geq 0$, we let $\binom{A}{k}$ denote the collection of all the k-element subsets of A. Define a *k-uniform hypergraph* $H = H(V, E)$ to be any pair (V, E), where V is a finite set (the *vertex set*), and E is a subset of $\binom{V}{k}$ (the *edge set*). Thus a 2-uniform hypergraph is the same as an ordinary graph.

The triangle removal lemma, Lemma 10.46, can be generalized as follows. If $H = H(V, E)$ is a k-uniform hypergraph, we define a *k-simplex* in H to be any set $S = \{v_1, \ldots, v_{k+1}\} \subset V$ of $k+1$ vertices such that $\binom{S}{k} \subset e$, i.e. the $k+1$ edges $S\backslash\{v_1\}, \ldots, S\backslash v_{k+1}$ all lie in E. Note that a 2-simplex is the same as a triangle.

Theorem 11.27 (Simplex removal lemma) *[283],[284],[140] Let $k \geq 2$, and let $H = H(V, E)$ be a k-uniform hypergraph which contains at most $\delta|V|^{k+1}$ k-simplices. Then it is possible to remove $o_{\delta\to 0;k}(|V|^2)$ edges from H to obtain a hypergraph which is simplex-free (it contains no k-simplices whatsoever).*

This result was conjectured by Erdős, Frankl and Rödl [87] in 1986, but not proven in full until much later. The $k = 2$ case dates back of course to [304] in 1978, but even the $k = 3$ case did not appear until 2002 [110] (though unpublished versions of this result existed much earlier, see for instance [109]); see also [139]. The full result was proven independently and simultaneously by Rödl and Skokan [283], [284] (see also [282], [254]) and Gowers [140]. A slight strengthening of this result was later established in [360] for the purposes of establishing arbitrary constellations in the Gaussian primes. For a recent survey of developments, see [281].

Just as Lemma 10.46 implies Proposition 10.47, Theorem 11.27 implies the following higher-dimensional analog:

Proposition 11.28 *Let Z be a finite additive group, let $k \geq 2$, and let $A \subset Z^k$ be such that A contains no "right-angled simplices" $x, x + re_1, \ldots, x + re_k$ with $x \in Z^k$ and $r \in Z \backslash 0$, where (by slight abuse of notation) we write re_i for $(0, \ldots, 0, r, \ldots, 0)$ with the ith position being the only non-zero one. Then $|A| = o_{|Z| \to \infty; k}(|Z|^k)$.*

This in turn can be used to deduce Theorem 11.24 as well as Szemerédi's theorem. In fact it yields Szemerédi's theorem for an arbitrary group: more precisely we have $r_k(Z) = o_{|Z| \to \infty; k}(|Z|)$ for any finite additive group Z with $|Z|$ coprime to $(k - 1)!$.

Just as the triangle removal lemma can be proven by the Szemerédi regularity lemma (indeed, this is currently its only proof known), the simplex removal lemma can be proven by a hypergraph regularity lemma. It turns out, however, that unlike the situation with the regularity lemma, where there is essentially one formulation (up to equivalences), there are several choices of hypergraph regularity lemma to choose from. The first such regularity lemma, introduced by Chung and Graham [58], regularizes the k-edge set E in terms of a partition of the vertex sets V, and is proven very similarly to the regularity lemma for graphs. Unfortunately, this lemma seems to be too weak to easily deduce the simplex removal lemma, the problem being that the regularity properties conferred by this lemma are not sufficient to obtain an accurate count for the number of simplices in the hypergraph, even in the 3-uniform case. The situation is intriguingly similar to the phenomenon noted in earlier sections that Fourier uniformity is insufficient to count progressions of length 4 or greater, even though Fourier analysis does not make an appearance in the regularity lemma. The solution (again in the 3-uniform case for simplicity) is to regularize the 3-edges by a partition of the 2-*edge set* $\binom{V}{2}$, and then regularize the 2-edge partition further (essentially using the ordinary regularity lemma) using a partition of the vertex set V (or equivalently $\binom{V}{1}$). This however leads to some new issues not present in the ordinary graph case. First, it is possible for the secondary partition to somehow disrupt the regularity obtained by the primary partition. Second, one has to decide on the relative strength of regularity between the primary partition and the secondary partition; this is particularly important since there is an expensive (tower-exponential) trade-off between the amount of regularity conferred by a partition, and the number of cells needed in the partition, and one may need the regularity in one partition to dominate the number of cells in another. Finally, even after all the appropriate regularity has been attained, one still needs to accurately count the number of simplices in the hypergraph.

These problems can all be solved, but require a certain amount of technicality. We will not give the general details here, but we will do the $k = 2$ case (i.e. the usual regularity lemma) in detail, in a way which allows for a relatively easy extension to the higher k case. Our treatment here follows [360], [359] (see also [7], [282] for some closely related arguments). Namely, we will view the regularity lemma as being akin to the Koopman–von Neumann theorems of previous sections, decomposing a function into a "compact" component and a "uniform" component. Indeed we have the following analog of Proposition 11.18. Let us say that a function $f : V_1 \times \cdots \times V_d \to \mathbf{R}^+$ is *K-constant* if for each $1 \leq i \leq d$ there exists a partition of V_i into K cells $V_{i,1}, \ldots, V_{i,K}$ (some of the cells may be empty or otherwise unequal in size) such that f is constant on each of the products $V_{1,i} \times \cdots \times V_{d,j}$. This will be analogous to the concept of almost periodicity used in previous sections.

Proposition 11.29 (Preliminary regularity lemma) *[359] Let $V_1, \ldots, V_d$ be arbitrary finite non-empty sets, let $f : V_1 \times \cdots \times V_d \to \mathbf{R}^+$ be such that $0 \leq f(x_1, \ldots, x_d) \leq 1$, let $\sigma > 0$, and let $F : \mathbf{R}^+ \times \mathbf{R}^+ \to \mathbf{R}^+$ be an arbitrary function. Then there exists a quantity $K = O_{\sigma,F}(1)$ and a decomposition $f = f_{U^\perp} + f_U$ with the following properties:*

- *the "anti-uniform" component $f_{U^\perp}$ obeys the bounds $0 \leq f_{U^\perp} \leq 1$. Furthermore there exists a K-constant approximation f_C to $f_{U^\perp}$ with $0 \leq f_C \leq 1$ and $\|f_{U^\perp} - f_C\|_{L^2(V_1 \times \cdots \times V_d)} \leq \sigma$;*
- *the "uniform" component f_U obeys the regularity estimate*

$$|\mathbf{E}_{x_1 \in V_1, \ldots, x_d \in V_d} 1_{A_1 \times \cdots \times A_d} f_U(x_1, \ldots, x_d)| \leq \frac{1}{F(\sigma, K)}$$

for all $A_1 \subseteq V_1, \ldots, A_d \subseteq V_d$.

Remark 11.30 For application to graph regularity one only needs $d = 2$. For more general d, the above lemma is closely related to the hypergraph regularity lemma of Chung and Graham [58].

Let us assume this proposition for the moment, and establish the regularity lemma in the more traditional formulation of Lemma 10.42.

Proof of Lemma 10.42 assuming Proposition 11.29 Let ϵ, m, and $G = G(V, E)$ be as in the lemma. We set $V_1 = V_2 = V$, and let $f(x_1, x_2)$ be the incidence matrix of G, i.e. $f(x_1, x_2) = \mathbf{I}(\{x_1, x_2\} \in E)$. Let $\sigma > 0$ be a small number to be chosen later (it will eventually be a small multiple of ε^4), and let $F : \mathbf{R}^+ \times \mathbf{R}^+ \to \mathbf{R}^+$ be a growth function (depending on m) to be chosen later. We apply Proposition 11.29 with $d = 2$ to obtain a decomposition $f = f_{U^\perp} + f_U$, a K-constant approximation f_C to $f_{U^\perp}$ for some $K = O_{\sigma,F}(1)$, and partitions $V_{1,1}, \ldots, V_{1,K}$ and $V_{2,1}, \ldots, V_{2,K}$ of V with respect to which f_C is constant. We can take the common refinement $V'_{K(i-1)+j} := V_{1,i} \cap V_{2,j}$ for $i, j \in [1, K]$ to obtain a unified partition $V'_1, \ldots, V'_{K^2}$ such that f_C is constant on each product $V'_i \times V'_j$. Next, we let N be the largest integer less than $\sigma|V|/100mK^2$ (we will assume V large enough depending on m, K, σ that $N \geq 1$), and partition each cell V'_j arbitrarily into disjoint sub-cells of size N, plus an error of size at most N. Note that all the errors, when put together, will give an error set V_* of size at most $K^2 N \leq \sigma|V|/100m$, while the remaining sub-cells yield a partition $V''_1, \ldots, V''_k$ of the remaining set $V \backslash V_*$, with each V''_j having cardinality exactly N. Thus we have

$$|V| - \frac{\sigma|V|}{100m} \leq kN \leq |V|,$$

thus $k \geq m$ and $k = \Theta(|V|/N) = \Theta(mK^2/\sigma)$. The partition $V''_1, \ldots, V''_k, V_*$ is not quite near-uniform in the sense of Section 10.6, because of the exceptional set V_*. But we can break up V_* arbitrarily into k near-uniform pieces and distribute them to the V''_j, replacing each set V''_j by a slightly larger set V'''_j such that the $V'''_1, \ldots, V'''_k$ are a near-uniform partition of V and

$$|V'''_j \backslash V''_j| = O\left(\frac{|V_*|}{k} + 1\right) = O\left(\frac{K^2 N}{mK^2/\sigma} + 1\right) = O(\sigma N)$$

where we again assume V to be suitably large depending on m, K, σ. Thus V'''_j is only larger than V''_j by a factor of $1 + O(\sigma)$.

Let us now investigate the ϵ-regularity of a pair V'''_i, V'''_j. Let $X \subset V'''_i, Y \subset V'''_j$ be such that $|X| \geq \epsilon|V'''_i|$ and $|Y| \geq \epsilon|V'''_j|$, in particular $|X|, |Y| \geq \epsilon N$. We wish to see if

$$|d(X', Y') - d(V'''_i, V'''_j)| \leq \epsilon;$$

this will follow if we can show

$$\left|\mathbf{E}_{x_1 \in V'''_i, x_2 \in V'''_j} f(x_1, x_2)\left(1_X(x_1)1_Y(y_1) - \frac{|X|}{|V'''_i|}\frac{|Y|}{|V'''_j|}\right)\right| \leq \varepsilon^3. \tag{11.33}$$

for arbitrary subsets $X \subset V'''_i, Y \subset V'''_j$ (with no lower bound on cardinality). We make some preliminary reductions. Observe that we may restrict X to V''_i and Y to Y''_i, and replace the range of x_1 and x_2 to V''_i and V''_j respectively, and only incur an error of $O(\sigma)$ on the left-hand side. Similarly we may replace $|X|/|V'''_i|$ and $|Y|/|V'''_j|$ by $|X|/N$ and $|Y|/N$ and again only accept an error of $O(\sigma)$, thus estimating the left-hand side of (11.33) by

$$\left|\mathbf{E}_{x_1 \in V''_i, x_2 \in V''_j} f(x_1, x_2)\left(1_X(x_1)1_Y(y_1) - \frac{|X|}{N}\frac{|Y|}{N}\right)\right| + O(\sigma).$$

Now since f_C is constant on $V''_i \times V''_j$ we have

$$\mathbf{E}_{x_1 \in V''_i, x_2 \in V''_j} f_C(x_1, x_2)\left(1_X(x_1)1_Y(y_1) - \frac{|X|}{N}\frac{|Y|}{N}\right) = 0.$$

Next, by the uniformity of f_U we have

$$\left|\mathbf{E}_{x_1 \in V_i'', x_2 \in V_j''} f_U(x_1, x_2) \left(1_X(x_1)1_Y(y_1) - \frac{|X|}{N}\frac{|Y|}{N}\right)\right| \leq \frac{|V|^2}{N^2}\frac{2}{F(\sigma, K)}$$
$$= O\left(\frac{mK^2}{\sigma F(\sigma, K)}\right).$$

By choosing F suitably (e.g. $F(\sigma, K) = mK^2/\sigma^2$) we can ensure that the right-hand side is $O(\sigma)$. Putting this all together, we can bound the left-hand side of (11.33) by

$$\left|\mathbf{E}_{x_1 \in V_i'', x_2 \in V_j''} (f_{U^\perp} - f_C)(x_1, x_2) \left(1_X(x_1)1_Y(y_1) - \frac{|X|}{N}\frac{|Y|}{N}\right)\right| + O(\sigma).$$

By the triangle inequality, this is less than

$$2\mathbf{E}_{x_1 \in V_i'', x_2 \in V_j''} |(f_{U^\perp} - f_C)(x_1, x_2)| + O(\sigma).$$

Now from Proposition 11.29 and Cauchy–Schwarz we have

$$\mathbf{E}_{x_1 \in V, x_2 \in V} |(f_{U^\perp} - f_C)(x_1, x_2)| \leq \sigma$$

which after trimming away the exceptional set V_* gives

$$\mathbf{E}_{x_1 \in V_1'' \cup \dots \cup V_k'', x_2 \in V_1'' \cup \dots \cup V_k''} |(f_{U^\perp} - f_C)(x_1, x_2)| = O(\sigma).$$

By Markov's inequality (and the uniform sizes of the V_j'') we conclude that

$$\mathbf{E}_{x_1 \in V_i'', x_2 \in V_j''} |(f_{U^\perp} - f_C)(x_1, x_2)| = O(\sigma/\varepsilon)$$

for all but at most ϵ of the pairs (i, j). In such a case we obtain a bound of $O(\sigma/\varepsilon)$ for the left-hand side of (11.33), which will be acceptable by choosing σ equal to a small multiple of ε. Finally, the bound $K = O_{\sigma,F}(1)$ now implies a bound $k = O_{\epsilon,m}(1)$ as required. This establishes the lemma (with the partition $V_1'', \dots, V_k''$). □

Note that in the above proof only a very specific choice of function $F()$ was needed. However, the ability to set the function F *arbitrarily* becomes very important in the hypergraph theory, as it is the easiest way to reconcile the problem mentioned earlier of needing to have the regularity control given by one partition dominate the number of cells of another partition without totally losing control of all the error terms. Of course the price one pays for this is that the total number of cells at the end of the argument becomes extremely large.

We now begin the proof of Proposition 11.29. The reader may wish to focus on the $d = 2$ case for sake of familiarity, although the general d case is no different. We will re-interpret the partitions $V_{i,1}, \dots, V_{i,K}$ of V_i as σ-algebras $\mathcal{B}_i$ on V_i for

$1 \leq i \leq d$, which induces a further σ-algebra $\mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_d$ on $V_1 \times \cdots \times V_d$, formed by the Cartesian products $V_{1,i_1} \times \cdots \times V_{d,i_d}$. Note in particular that the function $f_C := \mathbf{E}(f|\mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_d)$ will be a K-constant function between 0 and 1. The decomposition $f = f_{U^\perp} + f_U$ will be given by $f_{U^\perp} := \mathbf{E}(f|\mathcal{B}'_1 \otimes \cdots \otimes \mathcal{B}'_d)$ and $f_U := f - f_{U^\perp}$, where the $\mathcal{B}'_i$ are somewhat finer σ-algebras than the $\mathcal{B}_i$. The exact choice of $\mathcal{B}_i$, and $\mathcal{B}'_i$ will be determined by an energy increment algorithm very similar to that used to prove Proposition 10.36.

We turn to the details. We fix $V_1, \ldots, V_d$ and the function $f : V_1 \times \cdots \times V_d \to \mathbf{R}^+$. Given any σ-algebras $\mathcal{B}_1, \ldots, \mathcal{B}_d$ of V_1, V_2, we define the *energy* $\mathcal{E}_f(\mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_d)$ by

$$\mathcal{E}_f(\mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_d) := \|\mathbf{E}(f|\mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_d)\|^2_{L^2(V_1 \times \cdots \times V_d)},$$

thus the energy ranges between 0 and 1 and finer σ-algebras have higher energy. Just as Proposition 10.36 relied on Lemma 10.40, Proposition 11.29 will rely on the following analog.

Lemma 11.31 (Lack of uniformity implies energy increment) *Let $\mu > 0$ and $K' \geq 1$, and for each $1 \leq i \leq d$ let $\mathcal{B}'_i$ be a σ-algebra on V_i with at most K' atoms each such that*

$$|\mathbf{E}_{x_1 \in V_1, \ldots, x_d \in V_d} 1_{A_1 \times \cdots \times A_d}(f - \mathbf{E}(f|\mathcal{B}'_1 \otimes \cdots \otimes \mathcal{B}'_d))(x_1, \ldots, x_d)| \geq \mu$$

for some $A_1 \subseteq V_1, \ldots A_d \subseteq V_d$. Then for each $1 \leq i \leq d$ there exists finer σ-algebras $\mathcal{B}''_i$ than $\mathcal{B}'_i$ with at most $2K'$ atoms each such that

$$\mathcal{E}_f(\mathcal{B}''_1 \otimes \cdots \otimes \mathcal{B}''_d) \geq \mathcal{E}_f(\mathcal{B}'_1 \otimes \cdots \otimes \mathcal{B}'_d) + \mu^2.$$

Proof For $1 \leq i \leq d$, let $\mathcal{B}''_i$ be the σ-algebra generated by $\mathcal{B}'_i$ and A_i. Observe that

$$\mathbf{E}_{x_1 \in V_1, \ldots, x_d \in V_d} 1_{A_1 \times \cdots \times A_d}(f - \mathbf{E}(f|\mathcal{B}''_1 \otimes \cdots \otimes \mathcal{B}''_d))(x_1, \ldots, x_d) = 0$$

since $A_1 \times \cdots \times A_d$ is the union of atoms in $\mathcal{B}''_1 \otimes \cdots \otimes \mathcal{B}''_d$, on each of which $f - \mathbf{E}(f|\mathcal{B}''_1 \otimes \cdots \otimes \mathcal{B}''_d)$ has mean zero. Subtracting this from the hypothesis we conclude

$$\begin{aligned}|\mathbf{E}_{x_1 \in V_1, \ldots, x_d \in V_d} 1_{A_1 \times \cdots \times A_d}(&\mathbf{E}(f|\mathcal{B}''_1 \otimes \cdots \otimes \mathcal{B}''_d) \\ &- \mathbf{E}(f|\mathcal{B}'_1 \otimes \cdots \otimes \mathcal{B}'_d))(x_1, \ldots, x_d)| \geq \mu\end{aligned}$$

and hence by Cauchy–Schwarz

$$\|\mathbf{E}(f|\mathcal{B}''_1 \otimes \cdots \otimes \mathcal{B}''_d) - \mathbf{E}(f|\mathcal{B}'_1 \otimes \cdots \otimes \mathcal{B}'_d)\|^2_{L^2(V_1 \times \cdots \times V_d)} \geq \mu^2.$$

The claim then follows from Pythagoras' theorem. □

Proof of Proposition 11.29 This will be almost identical to Proposition 10.36. We construct a nested pair of σ-algebras $\mathcal{B}_i \subset \mathcal{B}'_i$ on V_i for each $1 \leq i \leq d$ and an integer $K \geq 1$ by the following double-loop algorithm.

- Step 0. Initialize $\mathcal{B}_i = \{\emptyset, V_i\}$ for each i.
- Step 1. Let K be the smallest integer such that each of the $\mathcal{B}_i$ have at most K atoms. Set $\mathcal{B}'_i := \mathcal{B}_i$ for each i; thus we trivially have $\mathcal{E}_f(\mathcal{B}'_1 \otimes \cdots \otimes \mathcal{B}'_d) \leq \mathcal{E}_f(\mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_d) + \sigma^2$.
- Step 2. If

$$|\mathbf{E}_{x_1 \in V_1, \ldots, x_d \in V_d} 1_{A_1 \times \cdots \times A_d}(f - \mathbf{E}(f|\mathcal{B}'_1 \otimes \cdots \otimes \mathcal{B}'_d))(x_1, \ldots, x_d)| \leq \frac{1}{F(\sigma, K)}$$

  for all $A_1 \subseteq V_1, \ldots, A_d \subseteq V_d$, then we terminate the algorithm. If not, then we can apply Lemma 11.31 to obtain for each $1 \leq i \leq d$ a new σ-algebra $\mathcal{B}''_i$ with at most twice as many atoms as $\mathcal{B}'_i$ such that

$$\mathcal{E}_f(\mathcal{B}''_1 \otimes \cdots \otimes \mathcal{B}''_d) \geq \mathcal{E}_f(\mathcal{B}'_1 \otimes \cdots \otimes \mathcal{B}'_d) + \frac{1}{F(\sigma, K)^2}.$$

- Step 3. If we have

$$\mathcal{E}_f(\mathcal{B}''_1 \otimes \cdots \otimes \mathcal{B}''_d) \leq \mathcal{E}_f(\mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_d) + \sigma^2$$

  then we set $\mathcal{B}' := \mathcal{B}''$ and return to Step 2. If instead we have

$$\mathcal{E}_f(\mathcal{B}''_1 \otimes \cdots \otimes \mathcal{B}''_d) > \mathcal{E}_f(\mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_d) + \sigma^2$$

  then we set $\mathcal{B} = \mathcal{B}''$ and return to Step 1.

Once the algorithm terminates we set $f_{U^\perp} := \mathbf{E}(f|\mathcal{B}'_1 \otimes \cdots \otimes \mathcal{B}'_d)$, $f_C := \mathbf{E}(f|\mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_d)$, and $f_U := f - f_{U^\perp}$. The verification that the algorithm does indeed terminate in finite time and gives the desired properties is almost identical to the analogous arguments in Proposition 10.36 and is left to the reader as an exercise. □

Remark 11.32 A closer inspection of the above argument shows that the number of atoms K in the σ-algebra $\mathcal{B}$ can increase from K to as much as $K2^{F(\sigma,K)^2}$ whenever we return from Step 3 to Step 1. Since the latter step can occur as often as $1/\sigma^2$ times, we see that the final complexity will most likely be a tower exponential in K or worse (unless we restrict F to have logarithmic growth or so, see [239] for some discussion of this type of lemma). As mentioned in Section 10.6, this tower-exponential behavior is unavoidable, see [136].

Now we discuss the extension of the regularity lemma to hypergraphs. To simplify the exposition we shall consider only the 3-uniform case. First it turns out that a minor modification of the proof of Proposition 11.29 yields the following variant, in which we obtain much stronger uniformity control (with respect to arbitrary 2-edge sets rather than vertex sets), but with a weaker and more complex notion of K-constancy. More precisely, let us say that a function $f : V_1 \times V_2 \times V_3 \to \mathbf{R}^+$ is $(K, 2)$-*constant* if there exist partitions $V_i \times V_j = E_{ij,1} \cup \cdots , E_{ij,K}$ for $ij = 12, 23, 31$ such that f is constant on each set

$$\{(x_1, x_2, x_3) \in V_1 \times V_2 \times V_3 : (x_i, x_j) \in E_{ij,a_{ij}} \text{ for all } ij = 12, 23, 31\}$$

for all $a_{12}, a_{23}, a_{31} \in [1, K]$.

Proposition 11.33 (Preliminary hypergraph regularity lemma) *[360] Let $V_1, \ldots, V_3$ be arbitrary finite non-empty sets, let $f : V_1 \times \cdots \times V_3 \to \mathbf{R}^+$ be such that $0 \leq f(x_1, x_2, x_3) \leq 1$, let $\sigma > 0$, and let $F : \mathbf{R}^+ \times \mathbf{R}^+ \to \mathbf{R}^+$ be an arbitrary function. Then there exists a quantity $K = O_{\sigma,F}(1)$ and a decomposition $f = f_{U^\perp} + f_U$ with the following properties:*

- *the "anti-uniform" component $f_{U^\perp}$ obeys the bounds $0 \leq f_{U^\perp} \leq 1$. Furthermore there exists a $(K, 2)$-constant approximation f_C to $f_{U^\perp}$ with $0 \leq f_C \leq 1$ and $\|f_{U^\perp} - f_C\|_{L^2(V_1 \times V_2 \times V_d)} \leq \sigma$;*
- *the "uniform" component f_U obeys the regularity estimate*

$$|\mathbf{E}_{x_1 \in V_1, x_2 \in V_2, x_3 \in V_3} 1_{A_{12}}(x_1, x_2) 1_{A_{23}}(x_2, x_3) 1_{A_{31}}(x_3, x_1) f_U(x_1, \ldots, x_d)| \leq \frac{1}{F(\sigma, K)}$$

for all $A_{12} \subseteq V_1 \times V_2$, $A_{23} \subseteq V_2 \times V_3$, $A_{31} \subseteq V_3 \times V_1$.

The proof of Proposition 11.33 is almost identical to that of Proposition 11.29 but with a somewhat heavier notational burden. We leave it as an exercise.

One can then deduce a full-strength regularity lemma for 3-uniform hypergraphs. The exact formulation of this lemma is rather messy and too unenlightening to be given here (see [139], [283], [284], [282], [360]), but we will describe the formulation indirectly by informally outlining the *proof* of the lemma, following [360]. Given a function $f : V_1 \times V_2 \times V_3 \to \mathbf{R}^+$, an initial error tolerance σ, and a growth function $F : \mathbf{R}^+ \times \mathbf{R}^+ \to \mathbf{R}^+$, we then decide upon a much faster growth function $F^{\text{fast}} : \mathbf{R}^+ \times \mathbf{R}^+ \to \mathbf{R}^+$, the exact choice of which will be chosen later. Applying Proposition 11.33 with this much faster growth function F^{fast} we obtain a primary decomposition $f = f_{U^\perp} + f_U$, where f_U is extremely regular with respect to 2-edge partitions (enjoying the fast function F^{fast} in the denominator), and $f_{U^\perp}$ is approximable by a $(K, 2)$-constant function f_C, where K has some (rather lousy) upper bound. The K-constant function can be described using $O(K)$ edge sets $E_{ij,a}$ in $V_i \times V_j$ for $ij = 12, 23, 31$. We then apply Proposition 11.29 to the indicators $1_{E_{ij,a}} : V_i \times V_j \to \mathbf{R}^+$ each of these edge sets $E_{ij,a}$, using the original function F, and replacing the error tolerance σ by something smaller, e.g. $1/F(\sigma, K)$. Strictly speaking, we need a "multiple function" or "vector-valued" version of Proposition 11.29 in which we regularize multiple functions simultaneously using a single partition, but this is not hard to set up. This gives us a new parameter $K' := O_{K,F,\sigma}(1)$, such that we have secondary decompositions of each of the indicator functions $1_{E_{ij,a}}$ into a K'-constant main term and some manageable errors. Finally, we choose F^{fast} so that $F^{\text{fast}}(\sigma, K)$ dominates any expression that will arise from K'; this basically means that F^{fast} is a tower-iterated version of F, and will ensure that the error term f_U in the primary decomposition is also manageable.

Thus to summarize (and glossing over the delicate issues regarding the relative sizes of various parameters), we start with a function $f(x_1, x_2, x_3)$ of three variables, and approximate it by a combination of $O(K)$ functions $1_{E_{ij,a}}(x_i, x_j)$ of two variables, plus manageable errors; we then approximate each of the $1_{E_{ij,a}}(x_i, x_j)$ by a combination of $O(K')$ functions of one variable (i.e. the indicators of the vertex classes), again plus manageable errors. With carefully chosen relative sizes of parameters as given above, this regularization of the original function f is suitable for such tasks as accurately counting the number of 3-simplices in a 3-uniform hypergraph, in a manner similar in spirit to (but somewhat lengthier than) the proof of Lemma 10.46. This in turn eventually leads to a proof of Theorem 11.27, which in turn implies Szemerédi's theorem and a number of other consequences.

Exercises

11.6.1 Deduce Proposition 11.28 from Lemma 10.46. (Hint: the vertex set V for the k-uniform hypergraph should consist of coordinate hyperplanes such as $\{(x_1, \dots, x_k) : x_i = \text{const}\}$, as well as the diagonal hyperplanes $\{(x_1, \dots, x_k) : x_1 + \cdots + x_k = \text{const}\}$.)

11.6.2 Use Proposition 11.28 to deduce Theorem 11.24.

11.6.3 Use Proposition 11.28 to deduce the claim $r_k(Z) = o_{|Z|\to\infty;k}(|Z|)$ whenever $|Z|$ is coprime to $(k-1)!$.

11.6.4 [139] Let V, W be disjoint sets of n vertices each. Let us color the vertices in V red or blue randomly and independently, with equal probability of each. Suppose we also color the *edges* in W (i.e. the elements of $\binom{W}{2}$) red or blue randomly and independently. Let $H = H(V \cup W, E)$ be the 3-uniform hypergraph with edge set consisting of all triples $\{v, w, w'\}$, where $v \in V$ and $\{w, w'\} \in \binom{W}{2}$ have the same color, together with all triples of the form $\{v, v', w\}$ with $\{v, v'\} \in \binom{V}{2}$ and $w \in w$. Let us also define the competing 3-uniform hypergraph $H' = H'(V \cup W, E')$, where E' consists of all the triples of the form $\{v, v', w\}$ with $\{v, v'\} \in \binom{V}{2}$ and $w \in W$, and with each triple of the form $\{v, w, w'\}$ with $v \in V$ and $\{w, w'\} in \binom{W}{2}$ belonging to E' with independent probability 1/2. Show