# REGULARITY LEMMA AND APPLICATIONS

MATHIAS SCHACHT

Abstract. We present Szemerédi's regularity lemma and a few standard applications, including the removal lemma for cliques, Roth's theorem on arithmetic progressions, and the Ramsey-Turán theorem for $K_4$.

## §1. The regularity lemma

Let $G = (V, E)$ be a graph $\varepsilon > 0$ and $d \geqslant 0$. We say a pair $(X, Y)$ of disjoint subsets of $V$ is $(\varepsilon, d)$-*regular*, if for all subsets $X' \subseteq X$ and $Y' \subseteq Y$ we have

$$\big|e_G(X', Y') - d|X'||Y'|\big| \leqslant \varepsilon|X||Y|\,.$$

Moreover, a pair $(X, Y)$ is $\varepsilon$-*regular*, if it is $(\varepsilon, d)$-regular for $d = d(X, Y) := \frac{e_G(X,Y)}{|X||Y|}$, where we use the convention $d(X, Y) = 0$ when $|X||Y| = 0$. We remark that this definition slightly differs from the original formulation of Szemerédi [45], where the error on the right-hand side is of the form $\varepsilon|X'||Y'|$ and one requires $|X'| \geqslant \varepsilon|X|$ and $|Y'| \geqslant \varepsilon|Y|$. However, both versions are equivalent up to the order of $\varepsilon$.

It is easy to see that any $(\varepsilon, d)$-regular pair $(X, Y)$ is *approximately degree regular*, in the sense that

$$\sum_{x\in X} \big||N(x) \cap Y| - d|Y|\big| \leqslant 3\varepsilon|X||Y| \quad \text{and} \quad \sum_{y\in Y} \big||N(y) \cap X| - d|X|\big| \leqslant 3\varepsilon|X||Y|\,,$$

i.e., at most $2\sqrt{\varepsilon}|X|$ vertices in $X$ have $(d \pm \sqrt{\varepsilon})|Y|$ neighbours in $Y$ and at most $2\sqrt{\varepsilon}|Y|$ vertices in $Y$ have $(d \pm \sqrt{\varepsilon})|X|$ neighbours in $X$. On the other hand, the uniform edge distribution imposed by $\varepsilon$-regularity is a much stronger property, as it is easy to come up with vertex degree regular graphs that are not $\varepsilon$-regular. Due to this Szemerédi's regularity lemma is sometimes referred to as *uniformity lemma*.

**Theorem 1.1** (Szemerédi's regularity lemma)**.** *For every $\varepsilon > 0$ and $t_0 \in \mathbb{N}$ there is some $T_0 = T_0(\varepsilon, t_0)$ such that every graph $G = (V, E)$ with $|V| = n \geqslant T_0$ admits a vertex partition $V_0 \cup V_1 \cup \ldots \cup V_t = V$ satisfying the following properties:*

(*i*) $|V_0| \leqslant \varepsilon n$ *and* $|V_1| = \cdots = |V_t|$,
(*ii*) $t_0 \leqslant t \leqslant T_0$, *and*
(*iii*) *all but at most $\varepsilon t^2$ pairs $(V_i, V_j)$ with $1 \leqslant i < j \leqslant t$ are $\varepsilon$-regular.*

The proof of Theorem 1.1 makes use of the *index* of a partition. For a partition $\mathcal{P} = (V_1, \ldots, V_t)$ of $V$ we define its index by

$$\operatorname{ind}(\mathcal{P}) = \frac{1}{|V|^2} \sum_{1 \leqslant i < j \leqslant t} d^2(V_i, V_j)|V_i||V_j|\,.$$

It follows from the definition of the index that

$$0 \leqslant \operatorname{ind}(\mathcal{P}) \leqslant \frac{1}{2} \tag{1.1}$$

for any partition $\mathcal{P}$ of $V$.

The following two lemmas are simple consequences of the Cauchy-Schwarz inequality and the and key observations for the proof of Theorem 1.1. The first lemma implies that the index is monotone under refinements of partition.

**Lemma 1.2.** *Let $G = (V, E)$ be a graph. For disjoint sets $U$, $W \subseteq V$ with partitions $U_1 \mathbin{\dot\cup} \ldots \mathbin{\dot\cup} U_s = U$ and $W_1 \mathbin{\dot\cup} \ldots \mathbin{\dot\cup} W_t = W$ we have*

$$\sum_{i \in [s]} \sum_{j \in [t]} d^2(U_i, W_j)|U_i||W_j| \geqslant d^2(U, W)|U||W|\,.$$

*In particular, if $\mathcal{Q}$ and $\mathcal{P}$ are partitions of $V$ and $\mathcal{Q}$ refines $\mathcal{P}$, then* $\operatorname{ind}(\mathcal{Q}) \geqslant \operatorname{ind}(\mathcal{P})$*.*

*Proof.* If $U$ or $W$ is empty, then the inequality is trivial. Otherwise we obtain from the Cauchy-Schwarz inequality

$$\sum_{i \in [s]} \sum_{j \in [t]} \Big(d(U_i, W_j)\sqrt{|U_i||W_j|}\Big)^2 \sum_{i \in [s]} \sum_{j \in [t]} \Big(\sqrt{|U_i||W_j|}\Big)^2 \geqslant \bigg(\sum_{i \in [s]} \sum_{j \in [t]} d(U_i, W_j)|U_i||W_j|\bigg)^2\,.$$

Consequently, since

$$\sum_{i \in [s]} \sum_{j \in [t]} d(U_i, W_j)|U_i||W_j| = e(U, W) = d(U, W)|U||W|$$

and $\sum_{i \in [s]} \sum_{j \in [t]} |U_i||W_j| = |U||W|$ we infer

$$\sum_{i \in [s]} \sum_{j \in [t]} d^2(U_i, W_j)|U_i||W_j| \geqslant \frac{d^2(U, W)|U|^2|W|^2}{|U||W|} = d^2(U, W)|U||W|\,,$$

as claimed. □

The next lemma shows that the index increases, if we split a pair along a "witness of irregularity."

**Lemma 1.3.** *Let $G = (V, E)$ be a graph. For disjoint sets $U$, $W \subseteq V$ with $U' \mathbin{\dot\cup} U'' = U$ and $W' \mathbin{\dot\cup} W'' = W$ satisfying*

$$e(U'', W'') = d(U, W)|U''||W''| + \eta|U||W| \tag{1.2}$$

*for some $\eta \in \mathbb{R}$, we have*

$$d^2(U',W')|U'||W'|+d^2(U',W'')|U'||W''|+d^2(U'',W')|U''||W'| \qquad (1.3)$$
$$+d^2(U'',W'')|U''||W''| \geqslant \big(d^2(U,W)+4\eta^2\big)|U||W|\,.$$

*Proof.* The lemma is trivial if $|U||W| = 0$ and for $\eta = 0$ it follows from Lemma 1.2. Hence, in view of (1.2) we may also assume $\eta > 0$ and $|U||W| > |U''||W''| > 0$.

Starting with the left-hand side we apply the Cauchy-Schwarz inequality as in the proof of Lemma 1.2, but this time only to the first three terms of the sum, and obtain

$$\text{left-hand side of (1.3)} \geqslant \frac{\big(e(U,W)-e(U'',W'')\big)^2}{|U||W|-|U''||W''|} + d^2(U'',W'')|U''||W''|\,.$$

Using $e(U,W) = d(U,W)|U||W|$, $e(U'',W'') = d(U'',W'')|U''||W''|$, and substituting (1.2) yields

$$\begin{aligned}
&\frac{\big(e(U,W)-e(U'',W'')\big)^2}{|U||W|-|U''||W''|} + d^2(U'',W'')|U''||W''| \\
&\qquad = d^2(U,W)|U||W| + \eta^2\frac{|U|^3|W|^3}{(|U||W|-|U''||W''|)|U''||W''|} \\
&\qquad \geqslant d^2(U,W)|U||W| + 4\eta^2|U||W|\,,
\end{aligned}$$

where we used the inequality of arithmetic and geometric means. □

After these preparations we establish Theorem 1.1.

*Proof of Theorem 1.1.* Let $\varepsilon > 0$ and $t_0 \geqslant 1$ be given. Starting with $t_0$ we define a sequence of integers $(t_i)_{i\in\mathbb{N}}$ recursively through

$$t_i = \left\lceil \frac{t_{i-1}2^{i+t_{i-1}}}{\varepsilon} \right\rceil \qquad (1.4)$$

and we set

$$T_0 = t_{\lfloor 1/\varepsilon^3 \rfloor}\,,$$

i.e., $T_0$ is given by a tower-type function of height poly$(1/\varepsilon)$. Given a graph $G = (V,E)$ with $n = |V| \geqslant T_0$ we prove the existence of a partition $\mathcal{P}$ satisfying properties *(i)*–*(iii)* of Theorem 1.1.

Starting with an arbitrary partition $\mathcal{P}^0 = (V_0^0, V_1^0, \dots, V_{t_0}^0)$ with $|V_s^0| = \lfloor n/t_0 \rfloor$ for $s \in [t_0]$ and

$$|V_0^0| < t_0 \overset{(1.4)}{\leqslant} \frac{\varepsilon}{2}t_1 \leqslant \frac{\varepsilon}{2}T_0 \leqslant \frac{\varepsilon}{2}n \qquad (1.5)$$

we shall consider a sequence of partitions $\mathcal{P}^i = (V_0^i, V_1^i \dots, V_{s_i}^i)$ of $V$ all of which satisfy properties *(i)* and *(ii)* with $T_0$ replaced by $t_i$. Moreover, the partition $\mathcal{P}^i$ will be *almost* a refinement of $\mathcal{P}^{i-1}$, with the exception that $V_0^i$ might be a superset of $V_0^{i-1}$. In order to work with refinements we shall consider partitions $\mathcal{P}_0^i$, which are obtained from $\mathcal{P}^i$ by splitting the exceptional class $V_0^i$ into singletons. This way we will obtain a sequence of refinements

$$\mathcal{P}_0^0 \succcurlyeq \mathcal{P}_0^1 \succcurlyeq \dots \succcurlyeq \mathcal{P}_0^i$$

with non-decreasing index (see Lemma 1.2). Furthermore, working under the assumption that $\mathcal{P}^{i-1}$ fails to satisfy property *(iii)* of Theorem 1.1, will enable us to show via Lemma 1.3, that in addition $\operatorname{ind}(\mathcal{P}_0^i) \geqslant \operatorname{ind}(\mathcal{P}_0^{i-1}) + \varepsilon^3$. In view of (1.1) this can happen not more than $2/\varepsilon^3$ times, which means for some $i \leqslant 1/\varepsilon^3$ we arrive at a partition $\mathcal{P}^i$ satisfying all three properties *(i)*–*(iii)* of Theorem 1.1. Below we give the details of the described approach.

Suppose for some $i \geqslant 1$ we are given a partition $\mathcal{P}^{i-1} = (V_0^{i-1}, V_1^{i-1}, \dots, V_{s_{i-1}}^{i-1})$ of $V$ satisfying

$$|V_0^{i-1}| \leqslant \big(1 - 2^{-i}\big)\varepsilon n\,, \qquad |V_1^{i-1}| = \dots = |V_{s_{i-1}}^{i-1}|\,, \qquad \text{and} \qquad s_{i-1} \leqslant t_{i-1}\,, \tag{1.6}$$

but failing to satisfy property *(iii)* of Theorem 1.1. We shall construct a partition $\mathcal{P}^i = (V_0^i, V_1^i, \dots, V_{s_i}^i)$ of $V$ such that

$$|V_0^i| \leqslant \big(1 - 2^{-(i+1)}\big)\varepsilon n\,, \qquad |V_1^i| = \dots = |V_{s_i}^i|\,, \qquad \text{and} \qquad s_i \leqslant t_i\,, \tag{1.7}$$

and, in addition, $\mathcal{P}_0^{i-1} \succcurlyeq \mathcal{P}_0^i$ and

$$\operatorname{ind}(\mathcal{P}_0^i) \geqslant \operatorname{ind}(\mathcal{P}_0^{i-1}) + \varepsilon^3\,. \tag{1.8}$$

The initial partition $\mathcal{P}^0$ satisfies (1.6) for $i = 1$ and $s_0 = t_0$ (see (1.5)) and (1.7) establishes (1.6) for $i+1$, which allows us to proceed by induction. Since (1.8) can hold for at most $2/\varepsilon^3$ indices $i$, this procedure must eventually end with a partition $\mathcal{P}^j$ satisfying properties *(i)*–*(iii)* of Theorem 1.1.

It is left to construct $\mathcal{P}^i$ from $\mathcal{P}^{i-1}$ such that (1.7) and (1.8) hold. Given $\mathcal{P}^{i-1}$ let $I$ be the set of all pairs $\{a,b\} \in [s_{i-1}]^{(2)}$ of indices such that $(V_a^{i-1}, V_b^{i-1})$ is not $\varepsilon$-regular. Assuming that $\mathcal{P}^{i-1}$ fails to satisfy property *(iii)* of Theorem 1.1 implies

$$|I| > \varepsilon s_{i-1}^2\,. \tag{1.9}$$

In particular, for every $\{a,b\} \in I$ there are sets $U_a^b \subseteq V_a^{i-1}$ and $U_b^a \subseteq V_b^{i-1}$ such that

$$e(U_a^b, U_b^a) = d(V_a^{i-1}, V_b^{i-1})|U_a^b||U_b^a| + \eta_{ab}|V_a^{i-1}||V_b^{i-1}|\,,$$

for some $\eta_{ab} \in \mathbb{R}$ with $|\eta_{ab}| > \varepsilon$.

We consider the auxiliary partition $\mathcal{Q}$ given by the coarsest common refinement of $\mathcal{P}^{i-1}$, that also refines all partitions $\big(U_a^b, V \smallsetminus U_a^b\big)$ and $\big(U_b^a, V \smallsetminus U_b^a\big)$ for all $\{a,b\} \in I$. In particular, $V_0^{i-1}$ is a class in $\mathcal{Q}$ and every partition class $Q$ from $\mathcal{Q}$ with $Q \subseteq V_a^{i-1}$ is either contained in $U_a^b$ or it is contained in $V^{i-1} \smallsetminus U_a^b$ whenever $\{a,b\} \in I$. Since every vertex class $V_a^{i-1}$ from $\mathcal{P}^{i-1}$ can be involved in at most $s_{i-1}-1$ pairs that are not $\varepsilon$-regular, we know that $\mathcal{Q}$ has besides the exceptional class $V_0^{i-1}$ at most

$$s_{i-1}2^{s_{i-1}-1} \leqslant \frac{1}{2}t_{i-1}2^{t_{i-1}} \tag{1.10}$$

other classes. Moreover, by definition the partition $\mathcal{Q}_0$, which splits the class $V_0^{i-1}$ from $\mathcal{Q}$ into singletons, refines $\mathcal{P}_0^{i-1}$. Applying Lemma 1.3 to every $(V_a^{i-1}, V_b^{i-1})$ with $1 \leqslant a < b \leqslant s_{i-1}$ and $\{a,b\} \in I$ and applying Lemma 1.2 to all pairs yields

$$\begin{aligned} \operatorname{ind}(\mathcal{Q}_0) &\geqslant \operatorname{ind}(\mathcal{P}_0^{i-1}) + \frac{1}{n^2}\sum_{\{a,b\}\in I} 4\eta_{ab}^2 |V_a^{i-1}||V_b^{i-1}| \\ &\overset{(1.9)}{\geqslant} \operatorname{ind}(\mathcal{P}_0^{i-1}) + \frac{1}{n^2}\cdot \varepsilon s_{i-1}^2 \cdot 4\varepsilon^2 \left(\frac{n-|V_0^{i-1}|}{s_{i-1}}\right)^2 \\ &\overset{(1.6)}{\geqslant} \operatorname{ind}(\mathcal{P}_0^{i-1}) + \varepsilon^3 \,. \end{aligned} \tag{1.11}$$

Finally, we derive $\mathcal{P}^i$ from $\mathcal{Q}$. For that we split every class $Q \neq V_0^{i-1}$ from $\mathcal{Q}$ into as many sets of size $\lceil n/t_i \rceil$ as possible and we add the remainders to $V_0^{i-1}$. Let $\mathcal{P}^i = (V_0^i, V_1^i, \dots, V_{s_i}^i)$ the resulting partition. Obviously, $|V_1^i| = \dots = |V_{s_i}^i|$ and $s_i \leqslant t_i$. Moreover,

$$\mathcal{P}_0^i \preccurlyeq \mathcal{Q}_0 \preccurlyeq \mathcal{P}_0^{i-1}\,,$$

and by Lemma 1.2 we have

$$\operatorname{ind}(\mathcal{P}_0^i) \geqslant \operatorname{ind}(\mathcal{Q}_0) \overset{(1.11)}{\geqslant} \operatorname{ind}(\mathcal{P}_0^{i-1}) + \varepsilon^3\,.$$

Finally, we observe

$$\left|V_0^i \smallsetminus V_0^{i-1}\right| \overset{(1.10)}{\leqslant} \frac{1}{2}t_{i-1}2^{t_{i-1}}\cdot\left(\lceil n/t_i\rceil - 1\right) \leqslant \frac{t_{i-1}2^{t_{i-1}}}{2t_i}n \overset{(1.4)}{\leqslant} \frac{\varepsilon}{2^{i+1}}n\,,$$

which combined with (1.6) implies $|V_0^i| \leqslant (1-2^{-(i+1)})\varepsilon n$. Consequently, we established (1.7) and (1.8) for the partition $\mathcal{P}^i$, which concludes the proof of Theorem 1.1. □

The proof of the regularity lemma shows that setting

$$T_0 = 2^{2^{2^{\cdot^{\cdot^{\cdot^{2^{t_0}}}}}}}$$

for a tower of twos of height poly$(1/\varepsilon)$ suffices. Somewhat surprisingly it turned out that this type of bound is "essentially" best possible. This was shown by Gowers [22] (see also [18, 30] for improved lower bound constructions).

## §2. The counting lemma

In many applications the regularity lemma is used in conjunction with some lemma that embeds a given graph in a suitable collection of $\varepsilon$-regular pairs. In fact, often we do not only find one copy, but many copies of the given graph, which is established by the *counting lemma*. For the special case of cliques $K_\ell$ it states that if all $\binom{\ell}{2}$ pairs of an $\ell$-partite graph $G$ are $\varepsilon$-regular, then the number of cliques in $G$ is close to the expected number of $K_\ell$ in a random $\ell$-partite graph on the same vertex partition and with the same edge densities.

**Proposition 2.1** (Counting lemma)**.** *Let $\varepsilon > 0$ and let $G = (V_1 \cup \ldots \cup V_\ell, E_G)$ be an $\ell$-partite graph. If every bipartite pair $(V_i, V_j)$ is $(\varepsilon, d_{ij})$-regular for some $d_{ij} \geqslant 0$, then*

$$\left| |\mathcal{K}_\ell(G)| - \prod_{1 \leqslant i < j \leqslant \ell} d_{ij} \cdot \prod_{i \in [\ell]} |V_i| \right| \leqslant \varepsilon \binom{\ell}{2} \prod_{i \in [\ell]} |V_i| \,,$$

*where $\mathcal{K}_\ell(G)$ is the set of copies of $K_\ell$ in $G$.*

The proof of Proposition 2.1 yields a more general result (see Proposition 2.2 below) and we introduce the necessary notation below.

For graphs $F$ and $G$ we denote by $\mathrm{Hom}(F, G)$ the set of *graph homomorphisms* $\varphi$ from $F$ to $G$, i.e., $\varphi\colon V(F) \longrightarrow V(G)$ and $\varphi(i)\varphi(j) \in E(G)$, whenever $ij \in E(F)$. For graph homomorphisms we simply write $\varphi\colon F \longrightarrow G$ and we denote by $\hom(F, G)$ the number of homomorphisms $|\mathrm{Hom}(F, G)|$. Note that injective homomorphisms correspond to labeled copies of $F$ in $G$.

Suppose $\varphi \in \mathrm{Hom}(F, R)$ for some graph $R$ with vertex set $V(R) = [t]$ and suppose $G = (V_1 \cup \ldots \cup V_t, E_G)$ is a $t$-partite graph. We denote by $\mathrm{Hom}_\varphi(F, G)$ the set of *$\varphi$-partite homomorphism*, i.e., $\mathrm{Hom}_\varphi(F, G)$ contains those $\psi \in \mathrm{Hom}(F, G)$ which in addition satisfy

$$\psi(w) \in V_{\varphi(w)}$$

for every $w \in V(F)$. Again we write $\hom_\varphi(F, G)$ for the number of $\varphi$-partite homomorphisms. Note that, in the special case when $\varphi$ is injective, then $\varphi$ yields a of $F$ in $G$ and, therefore, the following counting lemma is a generalisation of Proposition 2.1.

**Proposition 2.2** (Counting lemma)**.** *Let $\varphi \in \mathrm{Hom}(F, R)$ for graphs $F = (U, E_F)$ and $R = ([t], E_R)$ and let $\varepsilon > 0$. If $G = (V_1 \cup \ldots \cup V_t, E_G)$ is a $t$-partite graph such that for every edge $ij \in E_R$ the pair $(V_i, V_j)$ is $(\varepsilon, d_{ij})$-regular for some $d_{ij} \geqslant 0$, then*

$$\left| \hom_\varphi(F, G) - \prod_{uw \in E_F} d^\varphi_{uw} \cdot \prod_{u \in U} |V_{\varphi(u)}| \right| \leqslant \varepsilon |E_F| \prod_{u \in U} |V_{\varphi(u)}| \,,$$

*where $d^\varphi_{uw} = d_{\varphi(u)\varphi(w)}$.*

For example, in the case when $R$ contains some clique $K_\ell$ on vertices $i_1, \dots, i_\ell$, then Proposition 2.2 guarantees a copy of any $\ell$-chromatic graph $F = (U, E_F)$ in $G[V_{i_1} \cup \dots \cup V_{i_\ell}]$ with all $\binom{\ell}{2}$ pairs being $\varepsilon$-regular of density at least $d$ as long as

$$\varepsilon < \frac{d^{|E_F|}}{|E_F|} \quad \text{and} \quad |V_{i_1}| = \dots = |V_{i_\ell}| = m \text{ is sufficiently large.}$$

Indeed in this situation Proposition 2.2 yields at least $(d^{|E_F|} - \varepsilon|E_F|)m^{|U|} = \Omega(m^{|U|})$ homomorphism from $F$ to $G$. At most $O(m^{|U|-1}) = o(m^{|U|})$ of these homomorphism are not injective and, hence, for sufficiently large $m$ there is an (in fact, there are $\Omega(m^{|U|})$) injective homomorphism(s) in $\mathrm{Hom}(F, G)$, which gives rise to labeled copies of $F$ in $G$.

*Proof of Proposition 2.2.* The proposition is clearly true for graphs $F$ with at most one edge and we proceed by induction on $|E_F|$. Let graphs $F = (U, E_F)$, $R = ([t], E_R)$, and $G = (V_1 \cup \dots \cup V_t, E_G)$ and a homomorphism $\varphi \in \mathrm{Hom}(F, R)$ be given.

Fix some edge $ab \in E_F$ and consider the spanning subgraph $F' = F - ab$ of $F$ obtained by removing the edge $ab$ from $F$. We count the $\varphi$-partite homomorphisms from $F$ to $G$ by

$$\hom_\varphi(F, G) = \sum_{\psi \in \mathrm{Hom}_\varphi(F', G)} \mathbb{1}_{E_G}\big(\psi(a), \psi(b)\big)\,,$$

where $\mathbb{1}_{E_G}$ denotes the indicator function of the edge set of $G$, i.e., $\mathbb{1}_{E_G}(u, v) = 1$ if $uv \in E_G$ and 0 otherwise. In order to apply the induction assumption for $F'$, we rewrite the sum in the form

$$\begin{aligned}\hom_\varphi(F, G) &= \sum_{\psi \in \mathrm{Hom}_\varphi(F', G)} \Big(\mathbb{1}_{E_G}\big(\psi(a), \psi(b)\big) - d^\varphi_{ab} + d^\varphi_{ab}\Big)\\ &= \sum_{\psi \in \mathrm{Hom}_\varphi(F', G)} \Big(\mathbb{1}_{E_G}\big(\psi(a), \psi(b)\big) - d^\varphi_{ab}\Big) + d^\varphi_{ab} \cdot \hom_\varphi(F', G)\,.\end{aligned}$$

Owing to the induction assumption, we have

$$\begin{aligned}\Big|d^\varphi_{ab} \cdot \hom_\varphi(F', G) - \prod_{uw \in E_F} d^\varphi_{uw} \prod_{u \in U} |V_{\varphi(u)}|\Big| &= \Big|d^\varphi_{ab} \cdot \hom_\varphi(F', G) - d^\varphi_{ab} \cdot \prod_{uw \in E_{F'}} d^\varphi_{uw} \prod_{u \in U} |V_{\varphi(u)}|\Big|\\ &\leqslant d^\varphi_{ab} \cdot \varepsilon\,\big(|E_F| - 1\big) \prod_{u \in U} |V_{\varphi(u)}|\\ &\leqslant \varepsilon\,\big(|E_F| - 1\big) \prod_{u \in U} |V_{\varphi(u)}|\end{aligned}$$

and, therefore, proving

$$\Big|\sum_{\psi \in \mathrm{Hom}_\varphi(F', G)} \Big(\mathbb{1}_{E_G}\big(\psi(a), \psi(b)\big) - d^\varphi_{ab}\Big)\Big| \leqslant \varepsilon \prod_{u \in U} |V_{\varphi(u)}| \tag{2.1}$$

completes the inductive step. For the proof of (2.1) we consider the induced subgraph $F^\star$ obtained from $F$ by removing the vertices $a$ and $b$. Moreover, let $\varphi^\star$ be the homomorphism $\varphi\colon F \longrightarrow R$ restricted to $U^\star = U \smallsetminus \{a,b\}$. With this notation at hand we observe

$$\bigg|\sum_{\psi\in\mathrm{Hom}_\varphi(F',G)}\Big(\mathbb{1}_{E_G}\big(\psi(a),\psi(b)\big)-d^{\varphi}_{ab}\Big)\bigg| = \bigg|\sum_{\psi^\star\in\mathrm{Hom}_{\varphi^\star}(F^\star,G)}\ \sum_{\substack{\psi\in\mathrm{Hom}_\varphi(F',G)\\ \psi|_{U^\star}\equiv\psi^\star}}\Big(\mathbb{1}_{E_G}\big(\psi(a),\psi(b)\big)-d^{\varphi}_{ab}\Big)\bigg|$$
$$\leqslant \sum_{\psi^\star\in\mathrm{Hom}_{\varphi^\star}(F^\star,G)}\bigg|\sum_{\substack{\psi\in\mathrm{Hom}_\varphi(F',G)\\ \psi|_{U^\star}\equiv\psi^\star}}\Big(\mathbb{1}_{E_G}\big(\psi(a),\psi(b)\big)-d^{\varphi}_{ab}\Big)\bigg|.$$

The inner sum runs over all extension $\psi$ of a fixed partite homomorphism $\psi^\star$ of $F^\star$ to a homomorphism of $F'$. In particular, $\psi(a)$ must be in the neighbourhood of $\psi^\star(u)$ for every $u\in N_{F'}(a)$, i.e.,

$$\psi(a)\in W_a\,,\quad\text{where}\quad W_a = V_{\varphi(a)}\cap\bigcap_{u\in N_{F'}(a)} N_G\big(\psi^\star(u)\big)\,.$$

Similarly, we require $\psi(b)\in W_b = V_{\varphi(b)}\cap\bigcap_{u\in N_{F'}(b)} N_G\big(\psi^\star(u)\big)$, which leads to

$$\bigg|\sum_{\psi\in\mathrm{Hom}_\varphi(F',G)}\Big(\mathbb{1}_{E_G}\big(\psi(a),\psi(b)\big)-d^{\varphi}_{ab}\Big)\bigg| \leqslant \sum_{\psi^\star\in\mathrm{Hom}_{\varphi^\star}(F^\star,G)}\bigg|\sum_{\substack{w_a\in W_a\\ w_b\in W_b}}\big(\mathbb{1}_{E_G}(w_a,w_b)-d^{\varphi}_{ab}\big)\bigg|$$
$$= \sum_{\psi^\star\in\mathrm{Hom}_{\varphi^\star}(F^\star,G)}\Big|e_G\big(W_a,W_b\big)-d^{\varphi}_{ab}|W_a||W_b|\Big|$$
$$\leqslant \hom_{\varphi^\star}(F^\star,G)\cdot\varepsilon\big|V_{\varphi(a)}\big|\big|V_{\varphi(b)}\big|\,,$$

where we used the $\varepsilon$-regularity of $(V_{\varphi(a)},V_{\varphi(b)})$. Since

$$\hom_{\varphi^\star}(F^\star,G)\leqslant\prod_{u\in U^\star}\big|V_{\varphi^\star(u)}\big| = \prod_{u\in U\smallsetminus\{a,b\}}\big|V_{\varphi(u)}\big|$$

the estimate (2.1) follows, which concludes the proof of the proposition. □

## §3. The removal lemma

The removal lemma follows from a combined application of the regularity lemma and the counting lemma. The removal lemma asserts that for every graph $H$ the following is true, if the number of copies of $F$ in a large graph $G=(V,E)$ is at most $o(|V|^{|V(F)|})$ then one can remove $o(|V|^2)$ edges from $G$ in such a way that the resulting graph is $F$-free. The case $F=K_3$ was essentially proved by Ruzsa and Szemerédi [37], where a preliminary version of the regularity lemma from [44] was used. Erdős, Frankl, and Rödl [12] proved a very similar result for general $F$ and, in fact, the same proof yields the removal lemma as well.

The removal lemma in the form as stated below first appeared in the work Alon, Duke, Lefmann, Rödl, and Yuster [1] for cliques and in the work of Füredi [20] for general $F$.

**Theorem 3.1** (Removal lemma)**.** *For every graph $F$ and $\varrho > 0$ there exist $\eta > 0$ and $n_0$ such that the following holds.*

*If $G = (V, E)$ with $|V| = n \geqslant n_0$ contains at most $\eta n^{|V(F)|}$ labeled copies of $F$, then there exists a set $E_\star \subseteq E$ with $|E_\star| \leqslant \varrho n^2$ such that $G' = (V, E \smallsetminus E_\star)$ is $F$-free.*

*Proof.* Let $F = (U, E_F)$ and $\varrho > 0$ be given. The theorem is void for graphs $F$ with no edges and for $\varrho \geqslant 1/2$. Hence, we may assume $|E_F| > 0$ and $\varrho < 1/2$.

For an intended application of the regularity lemma we set

$$\varepsilon = \frac{\varrho^{|E_F|}}{8 \cdot |E_F|} \leqslant \frac{\varrho}{8} \qquad \text{and} \qquad t_0 = \left\lceil \frac{2}{\varrho} \right\rceil . \tag{3.1}$$

and Theorem 1.1 yields $T_0 = T_0(\varepsilon, t_0)$. We then fix the promised constants

$$\eta = \frac{\varrho^{|E_F|}}{3 \cdot (2T_0)^{|U|}} \qquad \text{and} \qquad n_0 = \max\left\{T_0, \left\lceil \frac{|U|^2}{\eta} \right\rceil\right\} . \tag{3.2}$$

Let $G = (V, E)$ be a graph with $|V| = n \geqslant n_0$ that contains at most $\eta n^{|U|}$ labeled copies of $F$. In other words, the number of injective homomorphisms in $\mathrm{Hom}(F, G)$ is at most $\eta n^{|U|}$. Since there are at most

$$|U|^2 n^{|U|-1} = \frac{|U|^2}{n} n^{|U|} \leqslant \frac{|U|^2}{n_0} n^{|U|} \leqslant \eta n^{|U|}$$

non-injective homomorphisms, we have

$$\hom(F, G) \leqslant 2\eta n^{|U|} . \tag{3.3}$$

We apply the regularity lemma with the chosen parameters $\varepsilon$ and $t_0$ to $G$ and obtain a partition $V_0 \mathbin{\dot\cup} V_1 \mathbin{\dot\cup} \dots \mathbin{\dot\cup} V_t = V$ with

$$\frac{2}{\varrho} \overset{(3.1)}{\leqslant} t_0 \leqslant t \leqslant T_0 \tag{3.4}$$

such that $|V_0| \leqslant \varepsilon n$, $|V_1| = \dots = |V_t|$, and all but at most $\varepsilon t^2$ pairs $(V_i, V_j)$ are $\varepsilon$-regular.

Next we select the edges for $E_\star$. We include an edge $xy$ of $G$ in $E_\star$ if at least one of the following statements holds

($a$) $x$ or $y$ is in $V_0$, or
($b$) $xy \in E_G(V_i)$ for some $i \in [t]$, or
($c$) $xy \in E_G(V_i, V_j)$ for $1 \leqslant i < j \leqslant t$ such that $(V_i, V_j)$ is not $\varepsilon$-regular, or
($d$) $xy \in E_G(V_i, V_j)$ for $1 \leqslant i < j \leqslant t$ such that $d(V_i, V_j) < \varrho$.

It is left to show that $E_\star$ has the desired properties, i.e.,

$$|E_\star| \leqslant \varrho n^2 \qquad \text{and} \qquad G' = (V, E \smallsetminus E_\star) \text{ is } F\text{-free.}$$

For the upper bound on $|E_\star|$ we observe that there are

- at most $|V_0| \cdot n \leqslant \varepsilon n^2$ edges satisfying (*a*),
- at most $t \cdot \binom{n/t}{2} < \frac{n^2}{2t}$ edges satisfying (*b*),
- at most $\varepsilon t^2 \cdot (n/t)^2 = \varepsilon n^2$ edges satisfying (*c*),
- less than $\binom{t}{2} \cdot \varrho \cdot (n/t)^2 < \varrho n^2/2$ edges satisfying (*d*).

Consequently, we derive the promised upper bound

$$|E_\star| < \Big(\varepsilon + \frac{1}{2t} + \varepsilon + \frac{\varrho}{2}\Big)n^2 \overset{(3.4)}{\leqslant} \Big(\varepsilon + \frac{\varrho}{4} + \varepsilon + \frac{\varrho}{2}\Big)n^2 \overset{(3.1)}{\leqslant} \varrho n^2 \,. \tag{3.5}$$

In order to verify that $G'$ is $F$-free, we suppose for a contradiction that $G'$ contains some copy of $F$. It follows from the definition of $E_\star$ that every edge of this copy lies in some $\varepsilon$-regular pair of density at least $\varrho$. This gives rise to some homomorphism $\varphi\colon F \longrightarrow R$ for the *reduced graph* $R$ defined by

$$V(R) = [t] \qquad \text{and} \qquad ij \in E(R) \iff (V_i, V_j) \text{ is } (\varepsilon, d_{ij})\text{-regular for some } d_{ij} \geqslant \varrho\,.$$

In particuar, $\varphi$, $F$, $R$, and $G'$ satisfy the assumptions of the counting lemma (Proposition 2.2), which yields

$$\begin{aligned}\hom(F,G) \geqslant \hom_\varphi(F,G') &\geqslant \Big(\prod_{uw\in E_F} d^\varphi_{uw} - \varepsilon\,|E_F|\Big)\prod_{u\in U}|V_{\varphi(u)}| \\ &\geqslant \big(\varrho^{|E_F|} - \varepsilon\,|E_F|\big)\Big(\frac{(1-\varepsilon)n}{t}\Big)^{|U|} \overset{(3.1)}{\geqslant} \frac{7}{8}\varrho^{|E_F|}\Big(\frac{n}{2T_0}\Big)^{|U|} \overset{(3.2)}{>} 2\eta n^{|U|}\,,\end{aligned}$$

which contradicts (3.3) and concludes the proof of the removal lemma. □

The removal lemma was generalised in several ways. Alon, Fischer, Krievelevich, and Szegedy [2] obtained a version for induced subgraphs. More precisely, this result asserts that any large graph $G = (V, E)$ containing at most $o(|V|^{|V_F|})$ induced copies of $F$ can be changed in $o(|V|^2)$ places by removing and adding edges such that the resulting graph $G'$ contains no induced copy of $F$ at all. Further extensions of Alon and Shapira [3, 4] allow to forbid not only a single graph $F$, but a possibly infinite family of graphs $\mathcal{F}$. These results rely on an iterated version of the regularity lemma and had some applications in the area of *property testing* in theoretical computer science.

Another line of research concerns quantitative aspects of the removal lemma. Owing to the use of the regularity lemma in the proof of the removal lemma, the constant $\eta = \eta(F, \varrho)$ is the reciprocal of some tower-type function of height polynomial in $1/\varrho$ and $|V(F)|$ and obtaining a better dependency is of great interest in extremal graph theory. Currently, the best dependency is due to Fox [16], who improved the height of the tower from a polynomial to a logarithmic dependency.

## §4. Triangle removal lemma and Roth's theorem

Ruzsa and Szemerédi [37] established a connection between the triangle removal lemma (Theorem 3.1 for $F = K_3$) and Roth's theorem [35,36] on arithmetic progressions of length three.

In 1936 Erdős and Turán considered the function

$$r_k(n) = \max\{|A|\colon A \subseteq [n] \text{ and } A \text{ contains no arithmetic progression of length } k\}$$

and conjectured $r_3(n) = o(n)$, i.e.,

$$\lim_{n\longrightarrow\infty} \frac{r_3(n)}{n} = 0\,.$$

This conjecture turned out to be difficult, which was indicated by lower bound constructions of Salem and Spencer [40] giving $r_3(n) \geqslant n^{1-o(1)}$ and Behrend [5], who showed

$$r_3(n) \geqslant \frac{n}{\exp(c\sqrt{\log n})}$$

for some constant $c > 0$. This lower bound is up to the constant $c$ the best known lower bound for $r_3(n)$ and we refer to [11,26,32] for more details and recent results in that direction. Roth verified the conjecture of Erdős and Turán and proved in [36] the following upper bound.

**Theorem 4.1** (Roth's theorem)**.** *There is some $c > 0$ such that $r_3(n) \leqslant c\frac{n}{\log\log(n)}$.* □

There is a great interest to further close the gap between the lower and the upper bound on $r_3(n)$ and several improvements on the upper bound were obtained by Heath-Brown [27], Szemerédi [46], Bourgain [9,10], Sanders [38,39], and Bloom [6]. In a recent breakthrough Bloom and Sisask [7] obtained the currently best upper bound

$$r_3(n) \leqslant \frac{n}{\big(\log(n)\big)^{1+\varepsilon}}$$

for some sufficiently small $\varepsilon > 0$. For longer arithmetic progressions ($k > 3$) the conjecture $r_k(n) = o(n)$ is also attributed to Erdős and Turán. In that direction Szemerédi [42] first addressed the case $k = 4$ before resolving it for every $k$ in [44].

**Theorem 4.2** (Szemerédi's theorem)**.** *For every $k \geqslant 3$ we have $r_k(n) = o(n)$.* □

In his proof Szemerédi introduced an early version of the regularity lemma, which was also used in the original approach to the triangle removal lemma of Ruzsa and Szemerédi [37]. Since then several different proofs of Szemerédi's theorem were found and inspired further research in different branches of mathematics like *ergodic theory* (pioneered by Furstenberg [21]), *harmonic analysis* (due to Gowers [23, 24]), and *extremal hypergraph theory* (developed by Rödl and his collaborators [19, 31, 34] and Gowers [25]) and we refer to Tao [47] for a more detailed discussion of these developments.

Below we derive a qualitative version of Roth's theorem as a corollary of the triangle removal lemma. This reduction is due to Ruzsa and Szemerédi [37]. First we deduce the following simple consequence of the removal lemma.

**Corollary 4.3.** *For every $\delta > 0$ there exists an $n_0$ such that the following holds. If a graph $G = (V, E)$ with $|V| = n \geqslant n_0$ has the property that every edge belongs to exactly one triangle, then $|E| \leqslant \delta n^2$.*

*Proof.* Let $\delta > 0$ be given. For the definition of $n_0$ we apply the triangle removal lemma (Theorem 3.1 for $F = K_3$) with $\varrho = \delta/3$, which yields some $\eta > 0$ and some integer $n_1$. We then set

$$n_0 = \max\left\{n_1, \frac{1}{\eta}\right\}.$$

Let $G = (V, E)$ with $|V| = n \geqslant n_0$ be given and suppose every edge belongs to precisely one triangle. In particular, the number of labeled triangles in $G$ satisfies

$$\hom(K_3, G) = 6 \cdot \frac{|E|}{3} = 2 \cdot |E| \tag{4.1}$$

and, hence, $G$ contains at most $n^2 \leqslant \eta n^3$ labeled triangles. The triangle removal lemma asserts that there is a set $E_\star \subseteq E$ of size at most $\varrho n^2$ such that $G' = (V, E \smallsetminus E_\star)$ is triangle-free. Since every edge from $E_\star$ can destroy at most 6 labeled triangles in $G$, we also have

$$\hom(K_3, G) \leqslant 6 \cdot |E_\star|$$

and, in view of (4.1) this yields the desired estimate

$$|E| = \frac{1}{2} \cdot |\hom(K_3, G)| \leqslant 3 \cdot |E_\star| \leqslant 3\varrho n^2 = \delta n^2 .$$

□

**Corollary 4.4** (Qualitative version of Roth's theorem)**.** *We have $r_3(n) = o(n)$.*

*Proof.* Let $\varepsilon > 0$ be arbitrary and for $n_0$ given by Corollary 4.3 applied with $\delta = \varepsilon/12$ we shall show that $r_3(n) \leqslant \varepsilon n$ for every $n \geqslant n_0$. For that we consider an arbitrary set $A \subseteq [n]$ without arithmetic progression of length three. In order to apply Corollary 4.3 we define an auxiliary graph $G_A = (X \cup Y \cup Z, E)$. The graph $G_A$ is tripartite with vertex classes $X = [n]$, $Y = [2n]$, and $Z = [3n]$ (considered as disjoint sets). The edge set of $G_A$ is the union of *defining* triangles, where for every $x \in X$ and $a \in A$ we include the defining triangle $K(x,a)$ with vertices

$$x \in X\,, \quad x+a \in Y\,, \quad \text{and} \quad 2x+a \in Z\,.$$

Clearly, $|V(G_A)| = 6n$ and every edge of $G_A$ is in at least one triangle. Moreover, since any two vertices of a defining triangle uniquely determine the third vertex, the defining triangles are mutually edge disjoint and we have

$$\big|E(G_A)\big| = 3n|A|\,. \tag{4.2}$$

Next we show that every edge of $G_A$ belongs to at most one triangle and we suppose for a contradiction that some edge belongs to two triangles. Owing to the disjointness of the defining triangles this means that second triangle is created by three defining triangles $K(x,a)$, $K(x,b)$, and $K(x',c)$, i.e., the vertices

$$x \in X\,, \quad x+b = x'+c \in Y\,, \quad \text{and} \quad 2x+a = 2x'+c \in Z$$

span a triangle. Consequently,

$$b = c + (x'-x) \qquad \text{and} \qquad a = c + 2(x'-x)\,,$$

which would mean that $c$, $b$, and $a$ form an arithmetic progression of length three with difference $|x'-x|$. Since $a$, $b$, $c \in A$, this contradicts the assumption that $A$ does not contain any three term progression. Therefore, every edge of $G_A$ belongs to precisely one triangle and Corollary 4.3 yields

$$|E(G_A)| \leqslant \delta \cdot |V(G_A)|^2 = \delta \cdot (6n)^2$$

and with (4.2) we arrive at

$$|A| = \frac{|E(G_A)|}{3n} \leqslant \delta \cdot 12n = \varepsilon n\,,$$

which concludes the proof of Corollary 4.4. □

## §5. Ramsey-Turán type problems

In this section we discuss another application of the regularity method. Erdős and Sós [14] started the investigation of the following function, which can be viewed as a common generalisation of the problems addressed by Ramsey's theorem [33] and by Turán's theorem [48]

$$\mathrm{RT}(n;k,\ell) = \max\{e(G)\colon |V(G)| = n\,,\ \omega(G) < k\,,\ \text{and}\ \alpha(G) < \ell\}\,,$$

with the convention that $\mathrm{RT}(n;k,\ell) = 0$, if no graph $G$ with $\omega(G) < k$ and $\alpha(G) < \ell$ on $n$ vertices exists. For example, if $n$ is at least the Ramsey number $r(k,\ell)$, then no such graphs exists, while $\mathrm{RT}(n;k,\ell)$ equals the Turán number $\mathrm{ex}(n,K_k)$ for $\ell > n$. The connection to Ramsey's theorem indicates that determining $\mathrm{RT}(n;k,\ell)$ for all values is at least as hard as determining all Ramsey numbers $r(k,\ell)$, which appears to be hopeless. Erdős and Sós set out to investigate the asymptotic behaviour of $\mathrm{RT}(n;k,o(n))$.

Since neighbourhoods of triangle-free graphs induce independent sets, it easy to see that $\mathrm{RT}(n;3,o(n)) = o(n^2)$. In [14] Erdős and Sós proved for every $k \geqslant 2$

$$\mathrm{RT}(n;2k-1,o(n)) = \Big(\frac{k-2}{k-1}+o(1)\Big)\binom{n}{2}.$$

This resolves the problem for odd cliques $K_{2k-1}$ and only the problem for even cliques remained open. The first open case was addressed by Szemerédi [43], who proved the following upper bound.

**Theorem 5.1.** *For every $\eta > 0$ there exist $\alpha > 0$ and $n_0$ such that*

$$\mathrm{RT}(n;4,\alpha n) \leqslant \Big(\frac{1}{8}+\eta\Big)n^2$$

*for every $n \geqslant n_0$.*

At the time it was not clear whether the obtained upper bound is sharp. This changed when Bollobás and Erdős [8] came up with a beautiful construction of $n$-vertex, $K_4$-free graphs with independence number $o(n)$ and $(1/8-o(1))n^2$ edges, which provides a matching lower bound for Theorem 5.1. The general case for even cliques was subsequently addressed by Erdős, Hajnal, Sós, and Szemerédi [13] (see the survey [41] for a more detailed discussion on Ramsey-Turán type problems).

Below we use Theorem 1.1 to derive Theorem 5.1, which again gives tower-type dependency between $\eta$ and $\alpha$. The original proof of Szemerédi is based on a simple lemma, which might be viewed as a very early version of a regularity for graphs (see Proposition A.2), which predates the lemma appearing in [44]. We include Szemerédi's original argument in Appendix A. This proof gives a double-exponential dependency between $\eta$ and $\alpha$. This was further improved by Fox, Loh, and Zhao [17] to a polynomial dependency and the optimal relation was recently obtained by Lüders and Reiher [28].

*Proof of Theorem 5.1.* Let $\eta > 0$ be given. We fix an auxiliary constant

$$\varrho = \frac{\eta}{4}$$

and for the intended application of the regularity lemma we fix

$$\varepsilon = \frac{\varrho^3}{4} < \frac{\eta}{8} \quad \text{and} \quad t_0 = \Big\lceil \frac{4}{\eta} \Big\rceil \tag{5.1}$$

and Theorem 1.1 yields a constant $T_0$. Finally, we set

$$\alpha = \frac{\varepsilon}{3T_0} < \frac{\eta}{2T_0} \quad \text{and} \quad n_0 = T_0 \tag{5.2}$$

and let $n \geqslant n_0$.

We consider a $K_4$-free graph $G = (V, E)$ with $|V| = n$ vertices and $\alpha(G) \leqslant \alpha n$. Let $V_0 \cup V_1 \cup \ldots \cup V_t = V$ with $t_0 \leqslant t \leqslant T_0$ be the vertex partition provided by the regularity lemma and consider the reduced graph $R = ([t], E_R)$ defined by

$$ij \in E_R \quad \Longleftrightarrow \quad (V_i, V_j) \text{ is } (\varepsilon, d_{ij})\text{-regular for some } d_{ij} \geqslant \varrho.$$

Below we verify the following two claims.

**Claim 5.2.** *The graph $R$ is $K_3$-free.*

**Claim 5.3.** *For every $1 \leqslant i < j \leqslant t$ we have $d(V_i, V_j) \leqslant \frac{1}{2} + 2\eta$.*

Before we verify Claims 5.2 and 5.3 we conclude the proof of Theorem 5.1 based on these claims. In order to establish an appropriate upper bound on $|E|$ we note that every edge $xy$ of $G$ satisfies at least one of the following statements

(*a*) $x$ or $y$ is in $V_0$, or
(*b*) $xy \in E_G(V_i)$ for some $i \in [t]$, or
(*c*) $xy \in E_G(V_i, V_j)$ for $1 \leqslant i < j \leqslant t$ such that $(V_i, V_j)$ is not $\varepsilon$-regular, or
(*d*) $xy \in E_G(V_i, V_j)$ for $1 \leqslant i < j \leqslant t$ such that $d(V_i, V_j) < \varrho$, or
(*e*) $xy \in E_G(V_i, V_j)$ and $ij \in E_R$.

Similarly as in the proof of Theorem 3.1 (see derivation of (3.5)) we observe that our choice of $\varrho$, $\varepsilon$, $t_0$ implies that there are at most

$$\varepsilon n^2 + \frac{n^2}{2t_0} + \varepsilon n^2 + \frac{\varrho n^2}{2} \leqslant \frac{\eta}{2} n^2 \,.$$

edges satisfying statements (*a*)–(*d*). Let $E'$ be the set of those edges, i.e., every edge in $E \smallsetminus E'$ satisfies statement (*e*). Claim 5.3 tells us

$$|E \smallsetminus E'| \leqslant \frac{1 + 4\eta}{2} \cdot \left(\frac{n}{t}\right)^2 \cdot |E_R| \,.$$

Mantel's theorem [29] (Turán's theorem for $K_3$) combined with Claim 5.2 implies

$$|E_R| \leqslant \frac{t^2}{4}$$

and, therefore, we arrive at

$$|E| = |E \smallsetminus E'| + |E'| \leqslant \frac{1 + 4\eta}{2} \cdot \left(\frac{n}{t}\right)^2 \cdot \frac{t^2}{4} + \frac{\eta}{2} n^2 = \left(\frac{1}{8} + \eta\right) n^2$$

as desired and it is left to verify both claims. □

*Proof of Claim 5.2.* Suppose for a contradiction that $i$, $j$, $k \in [t]$ span a triangle in $R$. In particular, $(V_i, V_j)$, $(V_i, V_k)$, and $(V_j, V_k)$ are $\varepsilon$-regular with density at least $\varrho$ and Proposition 2.2 for $F = K_3$ yields at least

$$(\varrho^3 - 3\varepsilon)|V_i||V_j||V_k| \overset{(5.1)}{=} \varepsilon|V_i||V_j||V_k|$$

triangles in $G[V_i \cup V_j \cup V_k]$. This means that there is some edge $xy \in E(V_i, V_j)$ that belongs to $\varepsilon|V_k|$ triangles, i.e.,

$$|N(x) \cap N(y) \cap V_k| \geqslant \varepsilon\,|V_k| \geqslant \varepsilon \cdot (1-\varepsilon)\frac{n}{t} \geqslant \frac{\varepsilon}{2T_0}n \overset{(5.2)}{>} \alpha n\,.$$

Since $\alpha(G) \leqslant \alpha n$, there exists some edge $zz'$ contained in $N(x) \cap N(y) \cap V_k$ and the vertices $x$, $y$, $z$, and $z'$ span a $K_4$ in $G$, which is a contradiction. □

*Proof of Claim 5.3.* Suppose for a contradiction that $d(V_i, V_j) > 1/2 + 2\eta$ for some distinct indices $i$, $j \in [t]$. Let $U_i$ be the vertices in $V_i$ with at least $(1/2 + \eta)|V_j|$ neighbours in $V_j$. Since

$$\Big(\frac{1}{2} + 2\eta\Big)|V_i||V_j| < e(V_i, V_j) < |V_i \smallsetminus U_i| \cdot \Big(\frac{1}{2} + \eta\Big)|V_j| + |U_i||V_j| \leqslant \Big(\frac{1}{2} + \eta\Big)|V_i||V_j| + |U_i||V_j|$$

we have

$$|U_i| \geqslant \eta|V_i| > \frac{\eta}{2T_0}n \overset{(5.2)}{>} \alpha n\,.$$

It follows from $\alpha(G) \leqslant \alpha n$ that there is some edge $uu' \in E(U_i)$ and the definition of $U_i$ implies

$$|N(u) \cap N(u') \cap V_j| \geqslant 2\eta|V_j| > \alpha n\,.$$

This yields an edge in $|N(u) \cap N(u') \cap V_j|$, which together with $u$ and $u'$ spans a $K_4$ in $G$, contradicting that $G$ is $K_4$-free. □

## Appendix A. Szemerédi's proof of Theorem 5.1

In this appendix we reproduce Szemerédi's proof of the following quantitative form of Theorem 5.1 from [43].

**Theorem A.1.** *For sufficiently small* $\eta > 0$ *Theorem 5.1 holds for* $\alpha = 2^{-2^{C\log(1/\eta)/\eta^2}}$ *for some* $C > 8$ *and sufficiently large* $n_0$.

A.1. **A weak predecessor of the regularity lemma.** The main tool in the proof of Theorem A.1 is the following lemma, Proposition A.2, which might be viewed as a weak predecessor regularity lemma for graphs. For a graph $G = (V, E)$, a vertex $v \in V$ and a set $U \subseteq V$, we denote by $N(v, U)$ and $\deg(v, U)$ the neighbourhood and the degree of $v$ within the set $U$, i.e.,

$$N_G(v, U) = \{u \in U \colon uv \in E\} \quad \text{and} \quad \deg_G(v, U) = |N_G(v, U)|$$

and when the graph $G$ is clear from the context, then we simply write $N(v)$ and $\deg(v)$ and drop the subscript.

**Proposition A.2.** *For all positive reals $\gamma$, $\varepsilon$, and $\delta$ and every $n$-vertex graph $G = (V, E)$ the following holds. For every set $U \subseteq V$ there exist subsets $U^\star \subseteq U$ and $V^\star \subseteq V$ such that*

*(i)* $|U^\star| \geqslant \gamma^{1/\varepsilon}|U|$,
*(ii)* $\deg(u, V^\star) \geqslant \deg(u) - \delta n$ *for every* $u \in U^\star$, *and*
*(iii)* *for every subset $W \subseteq V^\star$ with $|W| \geqslant \varepsilon n$ we have*

$$\big|\{u \in U^\star \colon \deg(u, W) < \delta|W|\}\big| < \gamma U^\star .$$

*Proof.* Given $\gamma$, $\varepsilon$, $\delta$, $G$, and $U$. We define iteratively the following finite sequences of sets $(U_i)$, $(V_i)$, and $(W_i)$, where $U_0 = U$, $V_0 = V$, and $W_0 = \varnothing$. If $U^\star = U_i$ and $V^\star = V_i$ satisfy property *(iii)* of Proposition A.2, then we stop and it will become clear from the proof, that in this case $U^\star$ and $V^\star$ also satisfy properties *(i)* and *(ii)*.

On the other hand, if $U_i$ and $V_i$ do not satisfy property *(iii)*, then there is a subset $W_{i+1} \subseteq V_i$ with

$$|W_{i+1}| \geqslant \varepsilon n \tag{A.1}$$

such that at least $\gamma|U_i|$ vertices from $U_i$ have degree at most $\delta|W_{i+1}|$ into $W_{i+1}$. Let $U_{i+1}$ be the set of those vertices, i.e.,

$$U_{i+1} = \{u \in U_i \colon \deg(u, W_{i+1}) < \delta|W_{i+1}|\} .$$

Consequently,

$$|U_{i+1}| \geqslant \gamma|U_i| \geqslant \gamma^{i+1}|U|$$

and if we show that this procedure stops after at most $\lfloor 1/\varepsilon \rfloor$ steps, then *(i)* follows. Moreover, we set

$$V_{i+1} = V_i \smallsetminus W_{i+1} .$$

Since $U_{i+1} \subseteq U_i \subseteq \cdots \subseteq U_0 = U$ and since the sets $W_j$ are mutually disjoint, it follows from the definition of $U_j$ that for every $u \in U_{i+1}$

$$\begin{aligned}\deg(u, V_{i+1}) &= \deg(u, V_i) - \deg(u, W_{i+1}) \\ &> \deg(u, V_i) - \delta|W_{i+1}| \\ &\geqslant \deg(u) - \delta|W_1 \cup W_2 \cup \ldots \cup W_{i+1}|\,.\end{aligned}$$

This way we also ensure property *(ii)* after the final iteration.

Finally, since property *(iii)* of the lemma holds trivially as soon as $V_j = \varnothing$, it follows from (A.1) and

$$|V_{i+1}| = |V_i| - |W_{i+1}| \leqslant |V_i| - \varepsilon n \leqslant |V_0| - (i+1)\varepsilon n = n - (i+1)\varepsilon n$$

that the procedure stops after at most $\lfloor 1/\varepsilon \rfloor$ iterations. □

A.2. **Proof of Theorem A.1.** Szemerédi's proof of Theorem A.1 relies on two applications of Proposition A.2. First we apply it to the given graph $G = (V, E)$ with $U = V$ and obtain subsets $U_1^\star$ and $V_1^\star$. In the second application we set $U = V_1^\star$ and obtain sets $U_2^\star$ and $V_2^\star$. Choosing $\delta$ for both applications to be sufficiently smaller than $\eta$, property *(ii)* of Proposition A.2 combined this with $K_4$-freeness yields that $V_1^\star$ and $V_2^\star$ contain more than half of all vertices. In fact, the intersection of both sets contains linearly many vertices. Moreover, since $U_2^\star$ and $V_1^\star \cap V_2^\star$ are subsets of $V_1^\star$ a careful choice of $\varepsilon$ for the first application of Proposition A.2 allows us to apply property *(iii)* with $W$ being these two sets. As a result we obtain a vertex $u_1 \in U_1^\star$ which has a 'large' neighbourhood in $U_2^\star$ and in $V_1^\star \cap V_2^\star$. However, owing to the second application of Proposition A.2, where we appeal to property *(iii)* with $W = N(u_1^\star, V_1^\star \cap V_2^\star)$, we then find a vertex $u_2 \in N(u_1, U_2^\star)$ with a 'large' neighbourhood in $N(u_1, V_1^\star \cap V_2^\star)$. Therefore, the $K_4$-freeness of $G$ implies that this large neighbourhood in $N(u_1, V_1^\star \cap V_2^\star)$ is an independent set, which contradicts the assumption on the independence number of $G$. Below we give the details of this outline.

*Proof of Theorem A.1.* Given $\eta > 0$ we set

$$\alpha = \sqrt{\eta}\left(\frac{1}{2}\right)^{(4/\eta)^{1+8/\eta^2}} \qquad \text{and} \qquad n_0 = \lceil 4/\eta \rceil\,.$$

Without loss of generality we may assume that $\eta$ is sufficiently small such that

$$\alpha \leqslant \min\left\{\frac{\eta^{3.5}}{32}, \sqrt{\eta}\left(\frac{\eta}{4}\right)^{1+8/\eta^2}\right\}\,. \tag{A.2}$$

Suppose $G = (V, E)$ is a graph with $|V| = n \geqslant n_0$, independence number $\alpha(G) < \alpha n$ and

$$|E| > (1/8 + \eta)n^2\,. \tag{A.3}$$

We will show that $G$ must contain a clique on four vertices.

First we move away from the average degree condition given in (A.3) to a minimum degree condition[1]. This idea is often used in extremal graph theory and it is easy to check that from (A.3) and from the assumption on $n_0$ it follows that there exists an induced subgraph $G' = (V', E')$ in $G$ with

$$|V'| = m \geqslant \sqrt{\eta} n \quad \text{and} \quad \delta(G') \geqslant (1/4 + \eta/2)m\,,$$

where we denote by $\delta(G')$ the minimum degree in $G'$. For the rest of the proof we only focus on $G'$.

As discussed in the outline we apply Proposition A.2 twice. However, in the first application we have to 'foresee' the second application, which we do by the following careful choice of $\varepsilon$. We set

$$\gamma_1 = \frac{1}{2}\,, \quad \varepsilon_1 = \left(\frac{\eta}{4}\right)^{1+8/\eta^2} \quad \text{and} \quad \delta_1 = \frac{\eta}{4}$$

and apply Proposition A.2 to $G'$ with $U = V'$. Proposition A.2 yields subsets $U_1^\star$ and $V_1^\star \subseteq V'$ satisfying properties *(i)*-*(iii)*. Before we move to the second application of Proposition A.2 we note that we can assume that

$$|V_1^\star| \geqslant \left(\frac{1}{2} + \frac{\eta}{4}\right) m\,. \tag{A.4}$$

In fact, owing to

$$|U_1^\star| \geqslant \gamma_1^{1/\varepsilon_1} m \geqslant (1/2)^{(4/\eta)^{1+8/\eta^2}} \sqrt{\eta} n = \alpha n$$

it follows from $\alpha(G) < \alpha n$, that there exists an edge $uv$ contained in $U_1^\star$. Moreover, due to property *(ii)* of Proposition A.2 the choice of $\delta_1$ guarantees that

$$d(u, V_1^\star) \geqslant \delta(G') - \delta_1 m \geqslant (1/4 + \eta/4)m\,,$$

and the same lower bound holds for the vertex $v$. Hence, if (A.4) would fail, then

$$|N(u, V_1^\star) \cap N(v, V_1^\star)| \geqslant \frac{\eta}{4} m \geqslant \frac{\eta^{3/2}}{4} n \overset{(A.2)}{\geqslant} \alpha n$$

yields an edge in the joint neighbourhood of $u$ and $v$, which results in a copy of $K_4$ in $G$.

For the second application of Proposition A.2 we set

$$\gamma_2 = \frac{\eta}{4}\,, \quad \varepsilon_2 = \frac{\eta^2}{8} \quad \text{and} \quad \delta_2 = \frac{\eta}{4}\,.$$

We apply the lemma to $G'$ with $U = V_1^\star$ and obtain sets $U_2^\star \subseteq V_1^\star$ and $V_2^\star \subseteq V'$. It is easy to check that

$$|U_2^\star| \geqslant \gamma_2^{1/\varepsilon_2} |V_1^\star| \geqslant \varepsilon_1 m \geqslant \varepsilon_1 \sqrt{\eta} n \overset{(A.2)}{\geqslant} \alpha n$$

[1] Actually here we slightly deviate from the original argument in [43].

and the same argument as for establishing (A.4) yields $|V_2^\star| \geqslant (1/2 + \eta/4)m$. Consequently,

$$|V_1^\star \cap V_2^\star| \geqslant \frac{\eta}{2}m\,. \tag{A.5}$$

In particular, $U_2^\star$ and $V_1^\star \cap V_2^\star$ are both subsets of $V_1^\star$ satisfying

$$|U_2^\star| \geqslant \varepsilon_1 m \quad \text{and} \quad |V_1^\star \cap V_2^\star| \geqslant \varepsilon_1 m\,.$$

Therefore, we can appeal to property *(iii)* from the first application of Proposition A.2 for $W = U_2^\star$ and $W = V_1^\star \cap V_2^\star$. Since $\gamma_1 = 1/2$ this gives rise to a vertex $u_1 \in U_1^\star$ such that

$$d(u_1, U_2^\star) \geqslant \delta_1 |U_2^\star| = \frac{\eta}{4}|U_2^\star| = \gamma_2 |U_2^\star|$$

and

$$d(u_1, V_1^\star \cap V_2^\star) \geqslant \delta_1 |V_1^\star \cap V_2^\star| \overset{(A.5)}{\geqslant} \delta_1 \frac{\eta}{2}m = \frac{\eta^2}{8}m = \varepsilon_2 m\,.$$

Consequently, we can appeal to property *(iii)* from the second application of Proposition A.2 for $W = N(u_1, V_1^\star \cap V_2^\star)$, which then yields a vertex $u_2 \in N(u_1, U_2^\star)$ with

$$d(u_2, N(u_1, V_1^\star \cap V_2^\star)) \geqslant \delta_2 |N(u_1, V_1^\star \cap V_2^\star)| \geqslant \delta_2 \varepsilon_2 m \geqslant \delta_2 \varepsilon_2 \sqrt{\eta} n \overset{(A.2)}{\geqslant} \alpha n\,.$$

Hence, the assumption on $\alpha(G)$ yields an edge in the joint neighbourhood of $u_1$ and $u_2$, which gives rise to a copy of $K_4$. □

**Acknowledgement.** I thank Tibor Szabó for providing a copy of [43] and helping with the translation.

## References

[1] N. Alon, R. A. Duke, H. Lefmann, V. Rödl, and R. Yuster, *The algorithmic aspects of the regularity lemma*, J. Algorithms **16** (1994), no. 1, 80–109. MR1251840 ↑3

[2] N. Alon, E. Fischer, M. Krivelevich, and M. Szegedy, *Efficient testing of large graphs*, Combinatorica **20** (2000), no. 4, 451–476. MR1804820 ↑3

[3] N. Alon and A. Shapira, *Every monotone graph property is testable*, SIAM J. Comput. **38** (2008), no. 2, 505–522. MR2411033 ↑3

[4] ———, *A characterization of the (natural) graph properties testable with one-sided error*, SIAM J. Comput. **37** (2008), no. 6, 1703–1727. MR2386211 ↑3

[5] F. A. Behrend, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U. S. A. **32** (1946), 331–332. MR0018694 ↑4

[6] T. F. Bloom, *A quantitative improvement for Roth's theorem on arithmetic progressions*, J. Lond. Math. Soc. (2) **93** (2016), no. 3, 643–663. MR3509957 ↑4

[7] T. F. Bloom and O. Sisask, *Breaking the logarithmic barrier in Roth's theorem on arithmetic progressions*, available at arXiv:2007.03528. Submitted. ↑4

[8] B. Bollobás and P. Erdős, *On a Ramsey-Turán type problem*, J. Combinatorial Theory Ser. B **21** (1976), no. 2, 166–168. MR0424613 ↑5

[9] J. Bourgain, *On triples in arithmetic progression*, Geom. Funct. Anal. **9** (1999), no. 5, 968–984. MR1726234 ↑4

[10] ———, *Roth's theorem on progressions revisited*, J. Anal. Math. **104** (2008), 155–192. MR2403433 ↑4

[11] M. Elkin, *An improved construction of progression-free sets*, Israel J. Math. **184** (2011), 93–128. MR2823971 ↑4

[12] P. Erdős, P. Frankl, and V. Rödl, *The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent*, Graphs Combin. **2** (1986), no. 2, 113–121. MR932119 ↑3

[13] P. Erdős, A. Hajnal, V. T. Sós, and E. Szemerédi, *More results on Ramsey-Turán type problems*, Combinatorica **3** (1983), no. 1, 69–81. MR716422 ↑5

[14] P. Erdős and V. T. Sós, *Some remarks on Ramsey's and Turán's theorem*, Combinatorial theory and its applications, II (Proc. Colloq., Balatonfüred, 1969), North-Holland, Amsterdam, 1970, pp. 395–404. MR0299512 ↑5

[15] P. Erdős and P. Turán, *On some sequences of integers*, J. London Math. Soc. **11** (1936), no. 4, 261–264. MR1574918 ↑

[16] J. Fox, *A new proof of the graph removal lemma*, Ann. of Math. (2) **174** (2011), no. 1, 561–579. MR2811609 ↑3

[17] J. Fox, P.-S. Loh, and Y. Zhao, *The critical window for the classical Ramsey-Turán problem*, Combinatorica **35** (2015), no. 4, 435–476. MR3386053 ↑5

[18] J. Fox and L. M. Lovász, *A tight lower bound for Szemerédi's regularity lemma*, Combinatorica **37** (2017), no. 5, 911–951. MR3737374 ↑1

[19] P. Frankl and V. Rödl, *Extremal problems on set systems*, Random Structures Algorithms **20** (2002), no. 2, 131–164. MR1884430 ↑4

[20] Z. Füredi, *Extremal hypergraphs and combinatorial geometry*, Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994), Birkhäuser, Basel, 1995, pp. 1343–1352. MR1404036 ↑3

[21] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256. MR0498471 ↑4

[22] W. T. Gowers, *Lower bounds of tower type for Szemerédi's uniformity lemma*, Geom. Funct. Anal. **7** (1997), no. 2, 322–337. MR1445389 ↑1

[23] ———, *A new proof of Szemerédi's theorem for arithmetic progressions of length four*, Geom. Funct. Anal. **8** (1998), no. 3, 529–551. MR1631259 ↑4

[24] ———, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), no. 3, 465–588. MR1844079 ↑4

[25] ———, *Hypergraph regularity and the multidimensional Szemerédi theorem*, Ann. of Math. (2) **166** (2007), no. 3, 897–946. MR2373376 ↑4

[26] B. Green and J. Wolf, *A note on Elkin's improvement of Behrend's construction*, Additive number theory, Springer, New York, 2010, pp. 141–144. MR2744752 ↑4

[27] D. R. Heath-Brown, *Integer sets containing no arithmetic progressions*, J. London Math. Soc. (2) **35** (1987), no. 3, 385–394. MR889362 ↑4

[28] C. M. Lüders and Chr. Reiher, *The Ramsey-Turán problem for cliques*, Israel J. Math. **230** (2019), no. 2, 613–652, DOI 10.1007/s11856-019-1831-4. MR3940430 ↑5

[29] W. Mantel, *Vraagstuk* XXVIII, Wiskundige Opgaven **10** (1907), 60–61. ↑5

[30] G. Moshkovitz and A. Shapira, *A short proof of Gowers' lower bound for the regularity lemma*, Combinatorica **36** (2016), no. 2, 187–194. MR3516883 ↑1

[31] B. Nagle, V. Rödl, and M. Schacht, *The counting lemma for regular $k$-uniform hypergraphs*, Random Structures Algorithms **28** (2006), no. 2, 113–179. MR2198495 ↑4

[32] K. O'Bryant, *Sets of integers that do not contain long arithmetic progressions*, Electron. J. Combin. **18** (2011), no. 1, Paper 59, 15. MR2788676 ↑4

[33] F. P. Ramsey, *On a problem of formal logic*, Proc. London Math. Soc. (2) **30** (1930), no. 1, 264–286, DOI 10.1112/plms/s2-30.1.264. MR1576401 ↑5

[34] V. Rödl and J. Skokan, *Regularity lemma for $k$-uniform hypergraphs*, Random Structures Algorithms **25** (2004), no. 1, 1–42. MR2069663 ↑4

[35] K. F. Roth, *Sur quelques ensembles d'entiers*, C. R. Acad. Sci. Paris **234** (1952), 388–390 (French). MR0046374 ↑4

[36] ———, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 104–109. MR0051853 ↑4

[37] I. Z. Ruzsa and E. Szemerédi, *Triple systems with no six points carrying three triangles*, Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Colloq. Math. Soc. János Bolyai, vol. 18, North-Holland, Amsterdam-New York, 1978, pp. 939–945. MR519318 ↑3, 4, 4

[38] T. Sanders, *On Roth's theorem on progressions*, Ann. of Math. (2) **174** (2011), no. 1, 619–636. MR2811612 ↑4

[39] ———, *On certain other sets of integers*, J. Anal. Math. **116** (2012), 53–82. MR2892617 ↑4

[40] R. Salem and D. C. Spencer, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U. S. A. **28** (1942), 561–563. MR0007405 ↑4

[41] M. Simonovits and V. T. Sós, *Ramsey-Turán theory*, Discrete Math. **229** (2001), no. 1-3, 293–340. Combinatorics, graph theory, algorithms and applications. MR1815611 ↑5

[42] E. Szemerédi, *On sets of integers containing no four elements in arithmetic progression*, Acta Math. Acad. Sci. Hungar. **20** (1969), 89–104. MR0245555 ↑4

[43] ———, *On graphs containing no complete subgraph with* 4 *vertices*, Mat. Lapok **23** (1972), 113–116 (1973) (Hungarian). MR0351897 ↑5, A, 1, A.2

[44] ———, *On sets of integers containing no $k$ elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245. Collection of articles in memory of Juriĭ Vladimirovič Linnik. MR0369312 ↑3, 4, 5

[45] ———, *Regular partitions of graphs*, Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), Colloq. Internat. CNRS, vol. 260, CNRS, Paris, 1978, pp. 399–401 (English, with French summary). MR540024 ↑1

[46] ———, *Integer sets containing no arithmetic progressions*, Acta Math. Hungar. **56** (1990), no. 1-2, 155–158. MR1100788 ↑4

[47] T. Tao, *What is good mathematics?*, Bull. Amer. Math. Soc. (N.S.) **44** (2007), no. 4, 623–634. MR2338369 ↑4

[48] P. Turán, *Eine Extremalaufgabe aus der Graphentheorie*, Mat. Fiz. Lapok **48** (1941), 436–452 (Hungarian, with German summary). MR0018405 ↑5



Fachbereich Mathematik, Universität Hamburg, Hamburg, Germany

*Email address*: `schacht@math.uni-hamburg.de`