1 Assignments: May 26th, 2026

We first recall a few definitions and the regularity lemma used in the exercise.

Definition 1.1 (matching)**.** *A matching of a graph G is a set of pairwise vertex-disjoint edges in $E(G)$. A matching M is called induced if for any two distinct edges $e_1 = v_1v_2 \in M$ and $e_2 = u_1u_2 \in M$, none of the four possible cross edges v_iu_j is an edge of G.*

Definition 1.2 (regular pair)**.** *For two disjoint vertex sets X, Y, write*

$$d(X,Y) = \frac{e(X,Y)}{|X||Y|}.$$

The pair (X, Y) is ϵ-regular if for all $X' \subseteq X$ and $Y' \subseteq Y$ with $|X'| \geq \epsilon|X|$ and $|Y'| \geq \epsilon|Y|$, we have

$$|d(X', Y') - d(X, Y)| < \epsilon.$$

Given an ϵ-regular partition $\mathcal{P} = \{V_0, V_1, \dots, V_k\}$ and a density threshold α, the reduced graph $R(\mathcal{P}, \alpha)$ has vertex set $[k]$, and ij is an edge if (V_i, V_j) is ϵ-regular and $d(V_i, V_j) > \alpha$. Here an ϵ-regular partition means that all but at most ϵk^2 pairs among (V_i, V_j), $1 \leq i < j \leq k$, are ϵ-regular.

Definition 1.3 (error edges)**.** *With respect to a reduced graph $R(\mathcal{P}, \alpha)$, the error edges are the edges of G which are not represented by edges of R. Concretely, these include edges incident with V_0, edges inside one cluster V_i, edges between irregular pairs, and edges between regular pairs whose density is at most α. After deleting all these edges, we denote the remaining graph by G_R.*

Theorem 1.1 (Szemerédi Regularity Lemma)**.** *For every $\epsilon > 0$ and every integer m, there exists an integer $M = M(\epsilon, m)$ such that every graph G on n vertices admits an ϵ-regular partition*

$$\mathcal{P} = \{V_0, V_1, \dots, V_k\}$$

with

$$m \leq k \leq M, \qquad |V_0| \leq \epsilon n, \qquad |V_1| = \cdots = |V_k|.$$

Moreover, all but at most ϵk^2 pairs (V_i, V_j) are ϵ-regular.

Exercise 1.1.
A matching of a graph G is a set of pairwise vertex-disjoint edges in $E(G)$. A matching M is said to be induced if for any two edges $e_1 = v_1v_2 \in M$, $e_2 = u_1u_2 \in M$, v_iu_j is not an edge of G for any distinct $i, j \in [2]$.
For a given constant c, let G be an n-vertex graph with $e(G) > cn^2$. Assume that G is a union of n induced matchings.
Apply Regularity Lemma on G with $\epsilon = \frac{c}{10}$ and $m = \epsilon^{-1}$, and obtain an ϵ-regular partition

$\mathcal{P} = \{V_0, V_1, \ldots, V_k\}$. Let $R = R(\mathcal{P}, 2\epsilon)$ be the reduced graph of G. Prove the following claims.

(a) Let G_R denote the graph after cleaning all "error" edges with respect to R. Compute $e(G_R)$.

(b) Prove that there is an induced matching M in G_R with more than $\frac{cn}{2}$ edges.

(c) Let $U_i = V_i \cap V(M)$ for each $i \in [k]$. Show that

$$\left| \bigcup_{\{i:\ |U_i| \geq \epsilon |V_i|\}} U_i \right| > \frac{9}{10} cn.$$

(d) Prove that there is a pair (U_i, U_j) from some ϵ-regular pair (V_i, V_j) such that $e(U_i, U_j) > |U_i|$.

(e) Prove that M is not an induced matching. That is to say, such G does not exist.

Proof. Let $s = |V_1| = \cdots = |V_k|$, and put $\epsilon = c/10$.

(a) Let E_{err} be the set of edges deleted in the cleaning step. Then

$$e(G_R) = e(G) - |E_{\mathrm{err}}|.$$

We now compute an upper bound for $|E_{\mathrm{err}}|$ from the four types of error edges.

The edges incident with V_0 are at most ϵn^2. The edges inside the clusters are at most

$$k\binom{s}{2} \leq \frac{ks^2}{2} \leq \frac{n^2}{2k} \leq \frac{\epsilon n^2}{2},$$

because $k \geq m = \epsilon^{-1}$. The irregular pairs contribute at most

$$\epsilon k^2 s^2 \leq \epsilon n^2$$

edges. Finally, the regular pairs of density at most 2ϵ contribute at most

$$2\epsilon \binom{k}{2} s^2 \leq \epsilon k^2 s^2 \leq \epsilon n^2$$

edges.

Thus

$$|E_{\mathrm{err}}| \leq \epsilon n^2 + \frac{\epsilon n^2}{2} + \epsilon n^2 + \epsilon n^2 = \frac{7}{2}\epsilon n^2.$$

Therefore

$$e(G_R) = e(G) - |E_{\mathrm{err}}| > cn^2 - \frac{7}{2}\epsilon n^2 = \frac{13}{20} cn^2 > \frac{c}{2} n^2.$$

(b) Write G as a union of n induced matchings $M_1, \ldots, M_n$. Deleting edges cannot destroy the property of being an induced matching, so each $M_t \cap E(G_R)$ is still an induced matching in G_R. These n matchings cover all edges of G_R, and therefore one of them, say M, has

$$|M| \geq \frac{e(G_R)}{n} > \frac{cn}{2}.$$

(c) Let

$$L = \{i \in [k] : |U_i| \geq \epsilon |V_i|\}.$$

Since $|M| > \frac{cn}{2}$, the matching M covers

$$|V(M)| = 2|M| > cn$$

vertices. On the other hand, the contribution from the clusters outside L is less than

$$\sum_{i \notin L} \epsilon |V_i| \leq \epsilon n = \frac{c}{10} n.$$

Thus most vertices covered by M must lie in the large clusters:

$$\left| \bigcup_{i \in L} U_i \right| > |V(M)| - \frac{c}{10} n > cn - \frac{c}{10} n = \frac{9}{10} cn.$$

(d) By the definition of L, the number of vertices of M lying in clusters outside L is less than

$$\sum_{i \notin L} \epsilon |V_i| \leq \epsilon n = \frac{c}{10} n.$$

Also, no edge of G_R is incident with V_0, so no endpoint of an edge of M lies in V_0. Hence fewer than $\frac{c}{10} n$ edges of M can meet a small cluster. Since $|M| > \frac{cn}{2}$, some edge $xy \in M$ has both endpoints in large clusters. Write $x \in U_i$ and $y \in U_j$, where $i, j \in L$.

Since xy is an edge of G_R, the pair (V_i, V_j) corresponds to an edge of the reduced graph R. Therefore (V_i, V_j) is ϵ-regular and

$$d(V_i, V_j) > 2\epsilon.$$

Also, by the choice of $i, j \in L$,

$$|U_i| \geq \epsilon |V_i|, \qquad |U_j| \geq \epsilon |V_j|.$$

By ϵ-regularity,

$$d(U_i, U_j) > d(V_i, V_j) - \epsilon > \epsilon.$$

Hence

$$e(U_i, U_j) > \epsilon |U_i||U_j|.$$

Here k is bounded only in terms of ϵ and m, so the cluster sizes tend to infinity as n grows. Since $j \in L$, we have $|U_j| \geq \epsilon |V_j|$; hence, for all sufficiently large n, $\epsilon |U_j| > 1$. Consequently,

$$e(U_i, U_j) > \epsilon |U_i||U_j| > |U_i|.$$

(e) The matching M contains at most $|U_i|$ edges between U_i and U_j, because every vertex of U_i is incident with at most one edge of M. However, by (d),

$$e(U_i, U_j) > |U_i|.$$

Thus there is an edge between U_i and U_j which is not one of the matching edges of M. Both of its endpoints are covered by M, and since the edge itself is not in M, these endpoints lie on two distinct matching edges. This is exactly a forbidden cross edge for an induced matching.

Therefore the assumed graph G cannot exist for sufficiently large n. $\square$