A Permutation Regularity Lemma

Joshua N. Cooper*

Institute for Theoretical Computer Science
ETH-Zürich, Zürich, Switzerland
cooper@cims.nyu.edu

Submitted: May 17, 2004; Accepted: Mar 4, 2006; Published: Mar 14, 2006
Mathematics Subject Classification: 05D40

Abstract

We introduce a permutation analogue of the celebrated Szemerédi Regularity Lemma, and derive a number of consequences. This tool allows us to provide a structural description of permutations which avoid a specified pattern, a result that permutations which scatter small intervals contain all possible patterns of a given size, a proof that every permutation avoiding a specified pattern has a nearly monotone linear-sized subset, and a "thin deletion" result. We also show how one can count sub-patterns of a permutation with an integral, and relate our results to permutation quasirandomness in a manner analogous to the graph-theoretic setting.

1 Introduction

The Szemerédi Regularity Lemma, a tool developed in the early 1970's in service of the combinatorial milestone now known as the Szemerédi Theorem, has turned out to be one of the most useful tools in graph theory ever discovered. In essence, it says that any graph can be approximated by a small collection of random-like graphs. This powerful structural characterization allows one to answer questions about graphs by taking such a "Szemerédi partition" and then addressing the question by using known facts about random graphs. A number of variants of the Regularity Lemma (or Uniformity Lemma, as it is sometimes called) have emerged since the publication of the original. Versions of it giving structural decompositions of hypergraphs have been used in many contexts, and a few results have addressed the difficult case of sparse graph regularity. The reader is encouraged to read the excellent surveys of Komlós and Simonovits [9] and Kohayakawa and Rödl [8] to learn about how and where the Lemma is used, how it is proved, and what its limitations are.

*Supported by NSF grant DMS-0303272.

An idea intimately related to regularity – quasirandomness – was introduced by Chung, Graham, and Wilson in [2]. They show that a surprisingly large collection of random-like properties of graphs are in fact equivalent. Then, in a series of remarkable papers, Chung and Graham applied similar analyses to hypergraphs, subsets of $\mathbb{Z}_n$, tournaments, and other combinatorial objects. The following decade witnessed a flurry of generalizations and elaborations appearing in the literature, with much of the work exploring the connections between regularity and quasirandomness. In particular, Simonovits and Sós [11] showed how quasirandomness is equivalent to the property of having a Szemerédi partition into pieces whose regular pairs have density 1/2.

The author defined quasirandom permutations in [3] and proved that several classes of simple arithmetic functions almost always give rise to quasirandom permutations ([4]). The central paradigm is the same: a large collection of natural, random-like properties are mutually equivalent. However, the connections with regularity break down in this realm, as it has not been possible so far to bridge the worlds of graph quasirandomness and permutation quasirandomness.

In the present paper, we remedy this situation by proving a regularity lemma for permutations and analogizing the basic results used alongside the graph Regularity Lemma. The main result (Theorem 2) says that the ground set of any permutation may be decomposed into a small exceptional set and a bounded number of intervals in the remaining points so that the action of the permutation is randomlike on each such interval. Our hope is that this tool will help address the nascent realm of "extremal permutation" problems and lead to other work analogous to that of Extremal Graph Theory.

Examples of extremal permutation problems include:

1. For any permutation τ, give a structural description of the permutation that *avoids* τ, i.e., $\sigma|_I$ is not order-isomorphic to τ for any index set I. The problem of showing that the number of such σ is at most exponential in the number of symbols is commonly known as the "Stanley-Wilf Conjecture", and was recently solved by Marcus and Tardos [10].

2. For a given permutation τ, which permutation σ has the maximum number of "copies" of τ, in the above sense? We write the number of such copies as $\Lambda^\tau(\sigma)$. This question has seen a number of advances in the past ten years, following Herb Wilf's address at the 1992 SIAM Conference on Discrete Math. One particularly nice addition to the recent literature in this realm is [7].

3. Given permutations τ and τ', what is the expected value of $\Lambda^{\tau'}(\sigma)$ in the space of permutations σ chosen uniformly among those permutations on n symbols which avoid τ? What is the maximum value of $\Lambda^{\tau'}(\sigma)$ among all those permutations σ which avoid τ?

4. Call a sequence of permutations $\{\sigma_i\}_{i=1}^\infty$, σ_i a permutation of n_i symbols with $n_i \to \infty$, *asymptotically k-symmetric* if, for each τ, a permutation on k symbols, $\Lambda^\tau(\sigma_i) =$

$\binom{n_i}{k}(1+o(1))/k!$. Does there exist, for all k, a sequence which is asymptotically k-symmetric but not asymptotically $(k+1)$-symmetric? This question of R. L. Graham appears in [3] and is open except for $k = 1, 2, 3$.

The rest of the paper is as follows. In the next section, we define regularity and uniformity for permutations and prove the existence of a regular/uniform partition. Then, in Section 3, we address Problem 1 above with structural results about permutations which avoid a given pattern. These results are used in Section 4 to show that only a small number of pairs of points need be deleted to destroy all copies of a pattern in a permutation which has few of them to begin with. Section 5 provides a connection between permutation quasirandomness and regularity, and a proof of a new characterization of permutation quasirandomness. The following section contains a discussion of the (asymptotic) pattern counts one can compute given a regular partition of any permutation, and the final section contains a full proof of the permutation regularity lemma.

2 Regularity

We provide two versions of a permutation regularity result, the latter of which appears to be the more interesting and applicable, and we distinguish the two settings through the use of the terms "regular" and "uniform." The first result, concerning regularity, we state below but relegate the proof – which is quite standard – to Section 7.

We consider permutations to be elements of $\mathfrak{S}_n$, the set of bijective maps from $\mathbb{Z}_n$ to itself. For a permutation $\sigma \in \mathfrak{S}_n$ and subsets $S, T \subset \mathbb{Z}_n$, we write $p(S,T) = |\{(s,t) \in S \times T : \sigma(s) < t\}|$, and $d(S,T) = p(S,T)/|S||T|$. Throughout the rest of this paper, we consider only partitions in which each C_i, $i \geq 1$, is an interval. Though it is something of an abuse, we will often speak of a "partition of σ" instead of a partition of $\mathbb{Z}_n$. For integers s and t and an $\epsilon > 0$, say that the pair (C_s, C_t) is *ϵ-regular* if, for all intervals $I \subset C_s$ and $J \subset C_t$ with $|I| \geq \epsilon |C_s|$ and $|J| \geq \epsilon |C_t|$, we have

$$|d(I,J) - d(C_s, C_t)| \leq \epsilon.$$

Then we call $\mathcal{P}$ an *ϵ-regular partition* into k parts if $|C_s| = |C_t|$ for all $1 \leq s,t \leq k$, $|C_0| \leq \epsilon n$, and (C_s, C_t) is an ϵ-regular pair for all but ϵk^2 pairs (s,t). (If P has only this first property, it is called *equitable*.)

Our first theorem is the following.

Theorem 1 (Permutation Regularity). *Given $m \in \mathbb{N}^+$ and $\epsilon > 0$, there exist $M = M(\epsilon, m)$ and $N = N(\epsilon, m)$ so that any $\sigma : \mathbb{Z}_n \to \mathbb{Z}_n$ has an ϵ-regular partition into k (nonexceptional) intervals with $m \leq k \leq M$ if $n \geq N$.*

Note that this statement is very similar to the one gotten by taking applying the "standard" Regularity Lemma for graphs to the bipartite graph whose color classes are two copies of $\mathbb{Z}_n$, and so that there is an edge from s to t if $\sigma(s) < t$. The difference lies primarily in that the blocks of the partition must be intervals, and the two partitions of the color classes are actually the same.

We now prove a reformulation of this result which will be easier to use for some applications. Let $\mathcal{L}(S,\alpha)$, for a set $S \subset \mathbb{Z}_n$ and $\alpha \in [0,1]$, denote the fraction of elements of S whose image is less than αn, i.e., $|\sigma(S) \cap [0,\alpha n)|/|S|$. We say that two functions $f,g : [0,1] \to [0,1]$ are *ϵ-near* if, for each $\alpha \in [0,1]$, $g(\alpha-\epsilon)-\epsilon \leq f(\alpha) \leq g(\alpha+\epsilon)+\epsilon$. (We employ the convention that $g(\alpha) = g(0)$ for $\alpha < 0$ and $g(\alpha) = g(1)$ for $\alpha > 1$.) It is easy to see that this definition is symmetric in f and g.

Now, we say that a partition $\{C_j\}_{j=0}^k$ of σ is $(\epsilon,\mathcal{F})$-uniform, where $\mathcal{F} = \{f_s\}_{s=1}^k$, if it is equitable, $|C_0| \leq \epsilon n$, and, for each $s \in \{1,\ldots,k\}$ and every interval $I \subset C_s$ with $|I| \geq \epsilon |C_s|$, $\mathcal{L}(I,\cdot)$ is ϵ-near f_s.

The following theorem, which we consider to be the main one of this paper, says essentially that permutations are, up to small deviations, concatenations of "deterministic" maps (ones which send all points into just a few small intervals) and "random" maps (ones which resemble the original map on each subinterval). Note the absence of any "exceptionality" other than the exceptional set itself, in contrast to the Graph Regularity Lemma, where exceptional pairs are unavoidable.

Theorem 2 (Permutation Uniformity). *Given $m \in \mathbb{N}^+$ and $0 < \epsilon < 1$, there exists $M = M(\epsilon,m)$ and $N = N(\epsilon,m)$ so that, if $n \geq N$, $\sigma : \mathbb{Z}_n \to \mathbb{Z}_n$ has an $(\epsilon,\mathcal{F})$-uniform partition $\{C_j\}_{j=0}^k$, with $m \leq k \leq M$, where $\mathcal{F}$ is a collection of k nondecreasing C^∞ functions $f_j : [0,1] \to [0,1]$.*

Proof. Without loss of generality, we may assume that $\epsilon < 1/2$. Apply Theorem 1, and take an $\epsilon^2/4$-regular partition of σ so that each C_j, $j \geq 1$, has cardinality $\leq \epsilon n/4$. (We may always do so by choosing a partition with even higher regularity if necessary.) Note that there can be at most $\epsilon k/2$ indices $s \in [k]$ so that there are more than $\epsilon k/2$ indices t with (s,t) being ϵ-irregular. Call all other s "good", and add each "bad" C_j to C_0 to create a new partition of σ. Then the new exceptional set has size at most $\epsilon^2 n/4 + (\epsilon k/2)(n/k) \leq \epsilon n$.

Fix a good s, and let A be any subset of $\mathbb{Z}_n$. Now, suppose $x,y \in \mathbb{Z}_n$ have the property that there is some $C_t \subset [x,y)$. Then

$$|C_t|^{-1}|\{(s,t) \in A \times C_t : \sigma(s) < t\}| \geq |\sigma(A) \cap [0,x)|,$$

so $d(A,C_t) \geq \mathcal{L}(A,x/n)$. Similarly, $d(A,C_t) \leq \mathcal{L}(A,y/n)$. In order to guarantee that there is such a C_t and that (s,t) is regular, it suffices to ensure that the gap between x and y is at least

$$(\epsilon k/2+1)|C_1| + \epsilon^2 n/4 \leq \epsilon n/2 + \epsilon n/4 + \epsilon^2 n/4 < \epsilon n,$$

since it should be the length of $\epsilon k/2+1$ C_j's plus all the points of C_0. Therefore, if we set $y = (\alpha+\epsilon)n$, $x = \alpha n$, we have

$$\mathcal{L}(X,\alpha) \leq d(X,C_t) \leq \mathcal{L}(X,\alpha+\epsilon).$$

On the other hand, we may take $z = (\alpha-\epsilon)n$, and there will be a $C_{t'} \subset [z,x)$, so that

$$\mathcal{L}(X,\alpha-\epsilon) \leq d(X,C_{t'}) \leq \mathcal{L}(X,\alpha) \leq d(X,C_t) \leq \mathcal{L}(X,\alpha+\epsilon). \tag{1}$$

If we take $X = I$, an interval of C_s of length at least $\epsilon|C_s|$, then we may apply (1) to get

$$d(I, C_{t'}) \leq \mathcal{L}(I, \alpha) \leq d(I, C_t).$$

Then, using the regularity of the partition, we see that

$$d(C_s, C_{t'}) - \epsilon \leq \mathcal{L}(I, \alpha) \leq d(C_s, C_t) + \epsilon.$$

Applying (1) once more, this time with $X = C_s$,

$$\mathcal{L}(C_s, \alpha - \epsilon) - \epsilon \leq \mathcal{L}(I, \alpha) \leq \mathcal{L}(C_s, \alpha + \epsilon) + \epsilon.$$

Since this analysis works for any $\alpha \in [\epsilon, 1-\epsilon)$, and the conclusion holds trivially otherwise, we may take $f_s(\alpha) = \mathcal{L}(C_s, \alpha)$.

Note that $\mathcal{L}(C_s, \alpha + 1/n) - \mathcal{L}(C_s, \alpha) \leq |C_s|^{-1} < 2kn^{-1}$. It is easy to see, then, that by choosing n large enough we may assume that all of the f_s are C^∞ and monotone. $\square$

3 Pattern Avoidance

Define $\Lambda^\tau(\sigma)$ for $\tau \in \mathfrak{S}_m$ and $\sigma \in \mathfrak{S}_n$ to be the number of occurrences of the pattern τ in σ, i.e., the number of "index sets" $\{x_0 < \ldots < x_{m-1}\} \subset \mathbb{Z}_n$ such that $\sigma(x_i) < \sigma(x_j)$ iff $\tau(i) < \tau(j)$.

Suppose that $\sigma \in \mathfrak{S}_n$ has a uniform partition P, and $\tau \in \mathfrak{S}_m$. If it is known that $\Lambda^\tau(\sigma) = o(n^m)$, what can be said about the f_s? In fact, something quite strong: that it concentrates almost all the mass of $\sigma(C_s)$ in at most $m-1$ very small intervals.

Theorem 3. *Suppose $\sigma \in \mathfrak{S}_n$, $\tau \in \mathfrak{S}_m$, $0 < \epsilon \leq (2m)^{-1}$, and n is sufficiently large. Choose $\{C_j\}_{j=0}^k$, an $(\epsilon, \mathcal{F})$-uniform partition of σ. If $\Lambda^\tau(\sigma) < (\epsilon n/2km)^m$, then, for each $1 \leq s \leq k$, there is a collection $\mathcal{I}$ of at most $m-1$ disjoint intervals in $[0,1)$, each of length at most 6ϵ, so that $|\sigma(C_s) \cap (n \cdot \bigcup \mathcal{I})| \geq |C_s|(1 - 7m\epsilon)$.*

Proof. Write $\mathcal{F} = \{f_s\}_{s=1}^k$. First we prove a claim: if $J_0, \ldots, J_{m-1}$ are disjoint intervals of $[0,1)$ which are separated from each other by at least 4ϵ, then, for some t, we have

$$f_s(\sup J_t) - f_s(\inf J_t) \leq 5\epsilon.$$

To see this, suppose the contrary, i.e., that there are m such intervals for which $f_s(\sup J_t) - f_s(\inf J_t) \geq 5\epsilon$. Then split C_s into m intervals $C_s^0, \ldots, C_s^{m-1}$ whose sizes differ by at most 1, and denote their density functions by $f_s^q(\cdot) = \mathcal{L}(C_s^q, \cdot)$. Writing $x_t = \inf J_t$ and $y_t = \sup J_t$, we have

$$\begin{aligned} f_s^q(y_t + 2\epsilon) - f_s^q(x_t - 2\epsilon) &\geq (\mathcal{L}(C_s, y_t + \epsilon) - \epsilon) - (\mathcal{L}(C_s, x_t - \epsilon) + \epsilon) \\ &\geq (f_s(y_t) - \epsilon) - (f_s(x_t) + \epsilon) - 2\epsilon \\ &= f_s(y_t) - f_s(x_t) - 4\epsilon \geq \epsilon, \end{aligned}$$

since $|C_s^q|/|C_s| \geq \epsilon$. Define $x'_t = \max\{0, x_t - 2\epsilon\}$ and $y'_t = \min\{1, y_t + 2\epsilon\}$, and note that the intervals $\{J'_t = [x'_t, y'_t)\}$ are disjoint, by the separation property of the J_t. Then the fact that $f_s^q(y'_t) - f_s^q(x'_t) \geq \epsilon$ for each q and t implies that $|\sigma(C_s^q) \cap J'_{\tau(q)}| \geq \epsilon|C_s^q| \geq \epsilon n/2km$. If we take any $z_q \in C_s^q \cap \sigma^{-1}(J'_{\tau(q)})$, then $z_0, \ldots, z_{m-1}$ is a τ-pattern in σ, so we have at least $(\epsilon n/2km)^m$ such patterns, a contradiction.

Now, consider the following process: begin at 0, and find the first r so that $f_s(r) = 5\epsilon$ (or $r = 1$ if such a point does not exist). This is possible because f_s is monotone and continuous and $f_s(1) = 1$. Define $I_0 = [0, r)$. Then, let $I'_0 = [r, r + 4\epsilon)$. Now, begin at $r + 4\epsilon$, find the first r' so that $f_s(r') - f_s(r + 4\epsilon) = 5\epsilon$ (or $r' = 1$, again, if this is not possible), and define $I_1 = [r + 4\epsilon, r')$ and $I'_1 = [r', r' + 4\epsilon)$. Then define r'', I_2, and I'_2 similarly, and so on. This process must terminate in no more than $\lceil 1/(5\epsilon) \rceil$ steps, at which point the right-endpoint of the last interval defined is 1. In fact, it must terminate even sooner, by the claim above: if we have reached I'_m, then $I_0, \ldots, I_{m-1}$ provide a contradiction. Then the I'_l number at most $m - 1$ and each has length at most 4ϵ. Now, define I''_l to be the interval with left endpoint $x''_l = \min\{\inf I_l + \epsilon, 1\}$ and right endpoint $y''_l = \max\{\sup I_l - \epsilon, 0\}$. Then

$$\mathcal{L}(C_s, y''_l) - \mathcal{L}(C_s, x''_l) \leq (f_s(\sup I_l) + \epsilon) - (f_s(\inf I_l) - \epsilon) \leq 7\epsilon.$$

Therefore, the intervals which comprise the complement of $\bigcup_l I''_l$, each of which contains some I'_l, satisfy the conclusions of the theorem. $\square$

Define a permutation $\sigma \in \mathfrak{S}_n$ to be *m-universal* if $\Lambda^\tau(\sigma) > 0$ for each $\tau \in \mathfrak{S}_m$. Now, we say that a permutation σ has the $(\delta, \epsilon, \gamma)$-property if, for every interval I with $|I| \geq \delta n$ and every interval J with $|J| \leq \epsilon n$, we have $|\sigma(I) \cap J| \leq \gamma|I|$. That is, no sufficiently large interval is mapped too densely into any small interval. Our next result says that, for the appropriate parameters, this property implies universality. Note that, if we had instead stated that $|\sigma(I) \cap J| \geq \gamma|I|$ when $|J| \geq \epsilon n$, this would be immediate. With the reverse inequalities, however, it is far from obvious. On the other hand, if $\delta\gamma \geq \epsilon$, the statement would be vacuous. Therefore, in particular, it has content whenever $\delta < \epsilon$.

Proposition 4. *For each $m \geq 2$ and $\epsilon > 0$, there is a positive $\delta < \epsilon$ so that, for n sufficiently large, if $\sigma \in \mathfrak{S}_n$ has the $(\delta, \epsilon, m^{-1})$-property, then σ is m-universal.*

Proof. Suppose the contrary, so that there is some $\tau \in \mathfrak{S}_m$ with $\Lambda^\tau(\sigma) = 0$. Take $\epsilon' = \min\{\epsilon/6, m^{-2}/8\}$, and choose an $(\epsilon', \mathcal{F})$-uniform partition $\{C_j\}_{j=0}^k$. Let C_s be any block of the partition. Then, by Theorem 3, at least $|C_s|(1 - 7m\epsilon')/(m - 1)$ points of $I = C_s$ are mapped by σ into some interval J of length at most ϵn. However, if we take $\delta = 1/k$ and $\gamma = 1/(m - 1)$, then the fact that σ has the $(\delta, \epsilon, \gamma)$-property provides a contradiction, since $|I| \geq \delta n$, $J \leq \epsilon n$, and

$$\frac{|\sigma(I) \cap J|}{|I|} \geq \frac{1 - 7m\epsilon}{m - 1} > \frac{1}{m}.$$

$\square$

Now, we show that any permutation which avoids a given τ has a linear sized subpattern which is "nearly monotone". (Compare to the Erdős-Szekeres Theorem, which says that any permutation on n symbols has a $\sqrt{n}$-sized truly monotone subpattern.) Define a permutation $\rho \in \mathfrak{S}_r$ to be *δ-pseudomonotone* if either $\Lambda^{(01)}(\rho) \leq \delta\binom{r}{2}$ or $\Lambda^{(10)}(\rho) \leq \delta\binom{r}{2}$. Then we have the following.

Proposition 5. *For every $\delta > 0$ and $\tau \in \mathfrak{S}_m$, there is a $c > 0$ so that, for any permutation $\sigma \in \mathfrak{S}_n$ which avoids τ, with n sufficiently large, there is a set $I \subset \mathbb{Z}_n$ with $|I| \geq cn$ so that $\sigma|_I$ is δ-pseudomonotone.*

Proof. We may assume $m \geq 2$, and fix $\epsilon = \eta/14m$ with $\eta \leq 1$. By Theorem 3, σ has an $(\epsilon, \mathcal{F})$-uniform partition so that, for each $1 \leq s \leq k$, there is an interval I_s of length at most 6ϵ so that $|\sigma(C_s) \cap (n \cdot I_s)| \geq |C_s|(1 - 7m\epsilon)/(m-1)$. Order the C_s left-to-right. Suppose that T of the I_s intersect some fixed I_t. At least

$$T \cdot \frac{1 - 7m\epsilon}{m-1} \cdot |C_s| \geq \frac{T(1-7m\epsilon)(1-\epsilon)n}{k(m-1)}$$

points are mapped by σ into an interval of length at most $18\epsilon n$. Therefore,

$$T \leq \frac{18\epsilon k(m-1)}{(1-7m\epsilon)(1-\epsilon)} < 6\eta k.$$

Hence, we may iteratively pick $s_1, \ldots, s_{\lceil \eta^{-1}/6 \rceil}$ so that the I_{s_j} are mutually disjoint. By the Erdős-Szekeres Theorem, there is a subset $s'_1 < \cdots < s'_R$ of these s_j of size at least $\lceil \eta^{-1}/6 \rceil^{1/2}$ which is monotone with respect to the obvious ordering on the I_{s_j}. Let X be the union of the $\sigma^{-1}(I_{s'_j}) \cap C_{s'_j}$. The only pairs of elements of X which possibly display the opposite ordering to that of the intervals $I_{s'_j}$ are contained within a single set of the form $\sigma^{-1}(I_{s'_j}) \cap C_{s'_j}$. The fraction of pairs in X of this type is at most

$$\frac{R\binom{|C_s|}{2}}{\binom{R|C_s|/2(m-1)}{2}} \leq \frac{4(m-1)^2}{R} \leq 10\eta^{1/2}(m-1)^2.$$

If we let $\eta = \delta^2(m-1)^{-4}/100$, then the set X is δ-pseudomonotone and has cardinality at least $n \cdot (20(m-1)/\delta k)$. $\square$

4 Destroying Patterns

With the graph regularity lemma, one can prove that, if a graph G contains at most $o(n^m)$ copies of some m-vertex graph, then we may remove $o(n^2)$ edges to destroy *all* copies. Is there any hope of proving something analogous for permutations?

The first observation to make is that this is certainly not possible if one wishes to delete elements of the ground set. Consider the permutation

$$\sigma = (1, 0, 3, 2, 5, 4, \ldots, n-1, n-2)$$

for n even. It is clear that, even though this permutation has $\Lambda^{10}(\sigma) = O(n) = o(n^2)$, one must remove $\Omega(n)$ points to destroy all copies. Furthermore, the generalization of this construction to other patterns is a simple matter.

Clearly, deleting points of the ground set is not the proper analogue of removing edges. Let us instead attempt to "delete" pairs of points. We wish to choose a subset $\mathcal{S} \subset \binom{\mathbb{Z}_n}{2}$ so that *every copy of* τ *in* σ *contains (in its index set) both points of some element of* $\mathcal{S}$. To state it another way: if we do not count index sets in which pairs from $\mathcal{S}$ appear, there are no copies of the pattern τ. Any copy of τ containing such a pair we say is *destroyed* by the deletion of $\mathcal{S}$. The main result of this section says that, using $o(n^2)$ such deletions, we may destroy *all* copies of τ in a permutation σ which has $\Lambda^\tau(\sigma) = o(n^m)$.

Proposition 6. *Suppose that* $\sigma \in \mathfrak{S}_n$, $\tau \in \mathfrak{S}_m$, *and* $\Lambda^\tau(\sigma) = o(n^m)$. *Then we may delete at most* $o(n^2)$ *index pairs to destroy all copies of* τ.

Proof. Take an $(\epsilon, \mathcal{F})$-uniform partition $\{C_j\}_{j=0}^k$, $\epsilon < (2m)^{-1}$, and choose n large enough that $\Lambda^\tau(\sigma) < (\epsilon n/2km)^m$. By Theorem 3, for each $1 \leq s \leq k$, there is a collection $\mathcal{I}_s$ of at most $m-1$ disjoint intervals $\{I_s^j\}$ in $[0,1)$, each of length at most 6ϵ, so that $|\sigma(C_s) \cap (n \cdot \bigcup \mathcal{I}_s)| \geq |C_s|(1-7m\epsilon)$. We create a new collection of families $\mathcal{I}'_s$ of intervals as follows. Begin with the $\mathcal{I}_s$. If an interval $n \cdot \mathcal{I}_s$ receives fewer than $\epsilon|C_s|$ points of C_s under the action of σ, we remove it from the collection. Then each $\mathcal{I}'_s$ has at most $m-1$ elements and $|\sigma(C_s) \cap (n \cdot \bigcup \mathcal{I}'_s)| \geq |C_s|(1-8m\epsilon)$.

Now, delete all pairs which contain at least one point of C_0 or a point whose image does not fall into any of the $n \cdot \mathcal{I}'_s$. There are most $(8m+1)\epsilon n^2$ of these. Then, delete all pairs which contain two points from any one of the sets C_s. This uses at most $k\binom{|C_s|}{2} \leq n^2/2k$ pairs. Finally, delete all pairs whose elements are mapped to points within $12\epsilon n$ of each other by σ. There are at most $12\epsilon n^2$ of these. Hence, letting $\epsilon \to 0$ and $k \to \infty$, the result follows if we can show that the chosen deletion indeed destroys all copies of τ.

Suppose not. Then the index set on which τ appears, $i_0 < \ldots < i_{m-1}$ must have the following properties:

1. For each $r = 0, \ldots, m-1$, $\sigma(i_r) \in n \cdot I_s^j$ for some s and j.
2. For each $q = 0, \ldots, m-2$, $\sigma(i_{\tau^{-1}(q)}) < \sigma(i_{\tau^{-1}(q+1)}) - 12\epsilon n$.
3. If $\sigma(i_r) \in n \cdot I_s^j$ and $\sigma(i_{r'}) \in n \cdot I_{s'}^{j'}$, then $s \neq s'$.

Since each of the I_s^j have diameter at most 6ϵ, the first two properties imply that the $n \cdot I_s^j$ must be disjoint. Order these (dilated) intervals by increasing s, i.e., $s_0 < \cdots < s_{m-1}$, and call them $J_0, \ldots, J_{m-1}$. Because they are disjoint and $\sigma(i_r) \in J_r$ for each r, the intervals themselves are ordered like a copy of τ. Therefore, since the indices s are distinct, for *any* set of m indices drawn one from each of $\sigma^{-1}(J_0) \cap C_{s_0}, \cdots, \sigma^{-1}(J_{m-1}) \cap C_{s_{m-1}}$, σ restricted to this set is a copy of τ. This ensures that

$$\Lambda^\tau(\sigma) \geq (\epsilon|C_s|)^m \geq (\epsilon n/2k)^m,$$

a contradiction. ∎

5 Quasirandomness

In [3], the author proves that a number of random-like properties of permutations are equivalent to one another. In order to state the main result of that paper, a few definitions are necessary. Fix a permutation $\sigma \in \mathfrak{S}_n$. For any $S, T \subset \mathbb{Z}_n$ we define the *discrepancy* of S in T as

$$D_T(S) = \left| |S \cap T| - \frac{|S||T|}{n} \right|,$$

and we define the discrepancy of a permutation σ by

$$D(\sigma) = \max_{I,J} D_J(\sigma(I)),$$

where I and J vary over all intervals of $\mathbb{Z}_n$. Also, define

$$D^*(\sigma) = \max_{I,J} D_J(\sigma(I)),$$

where I and J vary only over "initial" intervals, i.e., intervals of the form $[0, x)$.

We say that a sequence $\{\sigma_i\}_{i=1}^{\infty}$ of permutations of $\mathbb{Z}_{n_1}$, $\mathbb{Z}_{n_2}$, ... is *quasirandom* if $D(\sigma_i) = o(n_i)$. Often the indices are suppressed, and we simply say that $D(\sigma) = o(n)$.

By $e(x)$, we mean $e^{2\pi i x}$. We also use the convention that the name of a set and its characteristic function are the same. The following is a portion of the main theorem in [3].

Theorem 7. *For any sequence of permutations $\sigma \in \mathfrak{S}_n$, integer $m \geq 2$, and real $\alpha > 0$, the following are equivalent:*

[UB] *(Uniform Balance) $D(\sigma) = o(n)$.*

[UB*] *(Uniform Star-Balance) $D^*(\sigma) = o(n)$.*

[SP] *(Separability) For any intervals $I, J, K, K' \subset \mathbb{Z}_n$,*

$$\left| \sum_{x \in K \cap \sigma^{-1}(K')} I(x) J(\sigma(x)) - \frac{1}{n} \sum_{x \in K, y \in K'} I(x) J(y) \right| = o(n).$$

[mS] *(m-Subsequences) For any permutation $\tau \in \mathfrak{S}_m$ and intervals $I, J \subset \mathbb{Z}_n$ with $|I| \geq n/2$ and $|J| \geq n/2$, we have $|I \cap \sigma^{-1}(J)| \geq n/4 + o(n)$ and*

$$\Lambda^{\tau}(\sigma|_{I \cap \sigma^{-1}(J)}) = \frac{1}{m!} \binom{|\sigma(I) \cap J|}{m} + o(n^m).$$

[2S] *(2-Subsequences) For any intervals $I, J \subset \mathbb{Z}_n$ with $|I| \geq n/2$ and $|J| \geq n/2$, we have $|I \cap \sigma^{-1}(J)| \geq n/4 + o(n)$ and*

$$\Lambda^{(01)}(\sigma|_{I \cap \sigma^{-1}(J)}) - \Lambda^{(10)}(\sigma|_{I \cap \sigma^{-1}(J)}) = o(n^2).$$

[E(α)] *(Eigenvalue Bound α) For all nonzero $k \in \mathbb{Z}_n$ and any interval I,*

$$\sum_{s \in \sigma(I)} e(-ks/n) = o(n|k|^{\alpha}).$$

[T] *(Translation) For any intervals I, J,*

$$\sum_{k \in \mathbb{Z}_n} \left(|\sigma(I) \cap (J+k)| - \frac{|I||J|}{n} \right)^2 = o(n^3).$$

Furthermore, for any implication between a pair of properties above, there exists a constant K so that the error term $\epsilon_2 n^k$ of the consequent is bounded by the error term $\epsilon_1 n^l$ of the antecedent in the sense that $\epsilon_2 = O(\epsilon_1^K)$.

In [11], the authors connect graph quasirandomness and regularity by showing that, essentially, a sequence of graphs is quasirandom if and only if they possess density 1/2 regular partitions with arbitrarily small ϵ. Here, we prove an analogous result for permutations. Let $O_1(x)$ denote some real number whose absolute value is at most x.

Proposition 8. *A sequence of permutations $\sigma_i \in \mathfrak{S}_{n_i}$, $i \geq 1$, $n_i \to \infty$, is quasirandom if and only if, for each $\epsilon > 0$, given any $(\epsilon, \{f_s\})$-uniform partition of σ_i with i sufficiently large, $f_s(x)$ is 2ϵ-near $id(x) = x$ for each s.*

Proof. Suppose that σ is quasirandom, and let P be an $(\epsilon, \mathcal{F})$-uniform partition. For any interval C_s, **[UB]** implies that

$$\mathcal{L}(C_s, \alpha) = \alpha + o(1).$$

Therefore, if we choose n_i large enough, we may ensure that $|\mathcal{L}(C_s, \alpha) - \alpha| \leq \epsilon$ for all $\alpha \in [0,1]$, which immediately implies that f_s is 2ϵ-near id for each s.

On the other hand, suppose σ_i has an $(\epsilon, \{f_s\})$-uniform partition $P = \{C_j\}_{j=0}^k$ for all sufficiently large i, where f_s is 2ϵ-near id for each s. It is easy to see that this implies (by sub-additivity) that P is $(3\epsilon, \{id\})$-uniform. We may assume that the C_j are ordered left-to-right. Choose $[0,x], [0,y) \subset \mathbb{Z}_{n_i}$. If, for some s, $C_j \subset [0,x)$ for all $1 \leq j \leq s$, choose the largest such s, and let $X = [0,x) \cap C_{s+1}$ (or $\emptyset$ if $s = k$). Otherwise, let $X = [0,x)$ and $s = 0$. Then

$$[0,x) = \left(\bigcup_{1 \leq j \leq s} C_j \right) \cup X \cup E$$

for some $E \subset C_0$. Therefore, $s|C_1| + |X| = x + O_1(\epsilon n_i)$. We may write

$$\left| \sigma\big([0,x)\big) \cap [0,y) \right| = \left| \bigcup_{1 \leq j \leq s} \sigma(C_j) \cap [0,y) \right| + |\sigma(X) \cap [0,y)| + |\sigma(E) \cap [0,y)|.$$

Clearly, $\left|\sigma(E) \cap [0, y)\right| = O_1(\epsilon n_i)$. Define $\beta = y/n_i$. Applying uniformity, then,

$$\begin{aligned}\left|\sigma\big([0,x)\big) \cap [0,y)\right| &= \sum_{1 \le j \le s} |C_1| \mathcal{L}(C_j, \beta) + |X| \mathcal{L}(X, \beta) + O_1(\epsilon n_i) \\ &= \sum_{1 \le j \le s} |C_1| (\beta + O_1(3\epsilon)) + |X| (\beta + O_1(3\epsilon)) + O_1(\epsilon n_i) \\ &= (s|C_1| + |X|)\beta + O_1(4\epsilon n_i),\end{aligned}$$

so we may conclude that

$$\left| \left|\sigma\big([0,x)\big) \cap [0,y)\right| - \frac{xy}{n_i} \right| \le 5\epsilon n_i.$$

If we take $\epsilon \to 0$, then σ_i is quasirandom, by **[UB*]**.

□

6 Counting Subpatterns

We wish to count how many occurrences of the pattern $\tau \in \mathfrak{S}_m$ appear in a permutation σ of $\mathbb{Z}_n$ with a given $(\epsilon, \mathcal{F})$-uniform partition $\{C_j\}_{j=0}^k$. Unfortunately, we have no control over the structure of the exceptional set C_0, so it is not possible to get an "exact" count this way. Nonetheless, if we write $X = \bigcup_j C_j$, it is easy to see that, for $n \ge \epsilon^{-1} \ge 2$,

$$\left| \binom{n}{m} - \binom{|X|}{m} \right| \le 2\epsilon n^m/(m-1)! \tag{2}$$

so that a count of the τ-patterns on X is going to be close to the same count on all of $\mathbb{Z}_n$. Note that we may also ignore all but the set of occurrences of τ all of whose symbols occur in different C_j's, since the number of these is $\binom{k}{m}|C_s|^m$, which is off from $\binom{n}{m}$ by at most

$$\left| \binom{n}{m} - \binom{k}{m}|C_s|^m \right| \le 2\epsilon n^m/(m-1)! \tag{3}$$

for n sufficiently large.

There is an additional obstruction to counting patterns that is more subtle than these two issues. Suppose the mass of $\sigma(C_s)$, for some s, were very tightly concentrated in some interval. If we use f_s as an estimate of its density function, then, since the condition of ϵ-nearness can "dislocate" the entire mass of $\sigma(C_s)$ by up to ϵ, the counts could be off by a significant amount. On the other hand, only index sets whose images have two points close to one another can be affected in this way. Since there are few of these, with some work, we are able to ignore them in the total count.

One way ensure that the counts are accurate is simply to posit that f_s does not concentrate its mass too tightly. Therefore, define f_s to be (B, ϵ)-Lipschitz if, for each x, $|f_s(x+\epsilon) - f_s(x)| \le B\epsilon$. For example, if we have a (quasi-)random permutation, we may take $f_s(x) = x$ for each s, a function which is $(1, \epsilon)$-Lipschitz for each ϵ.

The following lemma makes this idea rigorous. Define

$$\mathfrak{D}^r(\beta) = \{\mathbf{x} = (x_1, \ldots, x_r) \in [0, \beta)^r : \forall i, 1 \le i < r, x_i < x_{i+1}\}.$$

Lemma 9. *Let f_j, g_j, $j = 1, \ldots, r$, be cumulative distribution functions on $[0, a]$, $a > 0$, and suppose that for each j, f_j is ϵ-near g_j. If g_j is (B, ϵ)-Lipschitz for each j, then*

$$\left| \int_{\mathfrak{D}^r(\beta)} \alpha(x_1)\, dF_r - \int_{\mathfrak{D}^r(\beta)} \alpha(x_1)\, dG_r \right| \le r(a+1)(B+1)\epsilon.$$

for any $0 \le \beta \le a$, where $dF_r = df_1(x_1) \cdots df_r(x_r)$, $dG_r = dg_1(x_1) \cdots dg_r(x_r)$, and $\alpha : [0, a] \to [0, 1]$ is any nondecreasing function.

Proof. We prove the result by induction, by repeated application of integration by parts. First, we check that it is true for $r = 1$. Note that, since f_j is ϵ-near g_j,

$$f_j(\beta) = g_1(\beta) + O_1((B+1)\epsilon)$$

by the Lipschitz property, and we have

$$\begin{aligned}
\int_{\mathfrak{D}^1(\beta)} \alpha(x_1) df_1(x_1) &= \int_0^\beta \alpha(x_1) df_1(x_1) \\
&= \alpha(\beta) f_1(\beta) - \int_0^\beta f_1(x_1) d\alpha(x_1) \\
&= \alpha(\beta) g_1(\beta) + O_1((B+1)\epsilon) \\
&\qquad - \int_0^\beta [g_1(x_1) + O_1((B+1)\epsilon)] d\alpha(x_1) \\
&= \int_0^\beta \alpha(x_1) dg_1(x_1) + O_1((a+1)(B+1)\epsilon).
\end{aligned}$$

Now, suppose the result holds for $r - 1$, with $r > 1$. Then

$$\begin{aligned}
\int_{\mathfrak{D}^r(\beta)} dF_r &= \int_0^\beta \int_{\mathfrak{D}^{r-1}(x_r)} \alpha(x_1)\, dF_{r-1}\, df_r(x_r) \\
&= \int_0^\beta \left(\int_{\mathfrak{D}^{r-1}(x_r)} \alpha(x_1) dG_{r-1} + O_1((r-1)(a+1)(B+1)\epsilon) \right) df_r(x_r) \\
&= \int_0^\beta \int_{\mathfrak{D}^{r-1}(x_r)} \alpha(x_1)\, dG_{r-1}\, df_r(x_r) + O_1((r-1)(a+1)(B+1)\epsilon).
\end{aligned}$$

The function $\alpha_1(x_r) = \int_{\mathfrak{D}^{r-1}(x_r)} \alpha(x_1)\, dG_{r-1}$ is nondecreasing, nonnegative, and bounded by 1, so we may apply the $r = 1$ case to get

$$\int_{\mathfrak{D}^r(\beta)} dF_r = \int_{\mathfrak{D}^r(\beta)} dG_r + O_1(r(a+1)(B+1)\epsilon).$$

□

We wish to be able to count subpatterns in permutations which do not necessarily have the Lipschitz property, however. In order to be able to use this result, we have the following Lemma which says that convolving the c.d.f. of a permutation with a uniform distribution on a short interval preserves nearness and gives us a Lipschitz property. Therefore, fix $\delta > 0$ and, given a c.d.f. f on $[0,1]$, define $\tilde{f}(t) = \delta^{-1}\int_{-\delta}^{0} f(t+s)ds$, a c.d.f. on $[0, 1+\delta]$.

Lemma 10. *If f and g, c.d.f.'s on $[0,1]$, are ϵ-near, then $\tilde{f}$ and $\tilde{g}$ are ϵ-near. Furthermore, $\tilde{f}$ is $(2\epsilon\delta^{-1}, \epsilon)$-Lipschitz.*

Proof. To see the first claim, we write

$$\tilde{g}(t+\epsilon) - \tilde{f}(t) = \delta^{-1}\int_{t-\delta}^{t} g(s+\epsilon) - f(s)\, ds \leq \epsilon.$$

For the second claim,

$$\begin{aligned}\tilde{f}(t+\epsilon) - \tilde{f}(t) &= \delta^{-1}\left(\int_{t+\epsilon-\delta}^{t+\epsilon} f(s)\, ds - \int_{t-\delta}^{t} f(s)\, ds\right)\\ &= \delta^{-1}\left(\int_{t}^{t+\epsilon} f(s)\, ds - \int_{t-\delta}^{t+\epsilon-\delta} f(s)\, ds\right)\\ &\leq 2\epsilon\delta^{-1}.\end{aligned}$$

□

Now, fix a permutation $\tau \in \mathfrak{S}_m$. Write $\mathfrak{D}$ for $\mathfrak{D}^m(1)$, $\mathfrak{D}'$ for $\mathfrak{D}^m(1+\delta)$, and define a differential form $d\omega_f$ on $[0,1)^m$ as follows:

$$d\omega_f = |C_1|^m \sum_{1\leq s_0<\ldots<s_{m-1}\leq k} \bigwedge_{j=0}^{m-1} df_{s_{\tau^{-1}(j)}}(x_j)$$

We define $d\omega_g$, etc., analogously. Then we have the following.

Theorem 11. *Suppose $d\omega_f$ is defined as above, $\epsilon \leq 1/2$, and n is sufficiently large. Then*

$$\left|\Lambda^\tau(\sigma) - \int_{\mathfrak{D}} d\omega_f\right| < (20\epsilon^{1/2}m^2 + 4/k)n^m/(m-1)!$$

Proof. Define $g_j(x) = \mathcal{L}(C_j, x)$. Suppose that m elements $x_i \in \mathbb{Z}_n$ are chosen uniformly at random, $x_i \in C_{j_i}$, for some sequence $j_0 < \ldots < j_{m-1}$. Writing $j'_r = j_{\tau^{-1}(r)}$, the probability that their images under σ form a τ is precisely

$$\int_{\mathfrak{D}} dg_{j'_0}(x_1)\cdots dg_{j'_{m-1}}(x_m)$$

because dg_j represents the distribution of the images $\sigma(x_i)$ for x_i chosen at random from C_j, and we wish to compute the probability that

$$\sigma(x_{j'_0}) < \cdots < \sigma(x_{j'_{m-1}}).$$

Multiplying by the number $|C_1|^m$ of m-tuples, adding over all subsets $\{j_i\}$, and accounting for (2) and (3) yields

$$|\Lambda^\tau(\sigma) - \int_{\mathfrak{D}} d\omega_g| \leq 4\epsilon n^m/(m-1)!\,. \tag{4}$$

By Lemma 9 and Lemma 10 with $\delta = \epsilon^{1/2}$, the quantity

$$\left|\int_{\mathfrak{D}'} d\tilde{g}_{j'_0}(x_1)\cdots d\tilde{g}_{j'_{m-1}}(x_m) - \int_{\mathfrak{D}'} d\tilde{f}_{j'_0}(x_1)\cdots d\tilde{f}_{j'_{m-1}}(x_m)\right|$$

is bounded by $m(2+\epsilon^{1/2})(2\epsilon^{1/2}+1)\epsilon \leq 9m\epsilon$. Summing up again, we find

$$\left|\int_{\mathfrak{D}'} d\omega_{\tilde{f}} - \int_{\mathfrak{D}'} d\omega_{\tilde{g}}\right| \leq 9\epsilon n^m/(m-1)!\,. \tag{5}$$

Now, let $dG = k^{-1}\sum_{j=1}^k dg_j$ and $d\tilde{G} = k^{-1}\sum_{j=1}^k d\tilde{g}_j$. If we choose an m-tuple of points from each of these distributions, the distributions of their orderings with respect to increasing j coincide so long as each point is at least δ away from all the others, since we may view $d\tilde{G}$ as a random draw from the distribution dG followed by a random "jump" forward uniformly distributed in $[0,\delta]$. The probability that such an m-tuple has two points at most δ apart is bounded by the probability that some pair of its points are that close, i.e.,

$$\begin{aligned}
&\leq \binom{m}{2} \max_x \int_{x-\delta}^{x+\delta} dG \\
&= \binom{m}{2} \max_x \frac{1}{k}\sum_{j=1}^k \int_{x-\delta}^{x+\delta} dg_j \\
&= \binom{m}{2} \max_x \frac{1}{k}\sum_{j=1}^k (\mathcal{L}(C_j, x+\delta) - \mathcal{L}(C_j, x-\delta)) \\
&= \binom{m}{2} \max_x \frac{|\bigcup_j \sigma(C_j) \cap [(x-\delta)n, (x+\delta)n)|}{k|C_1|} \\
&\leq \frac{m^2}{2}\cdot\frac{2\delta n}{k|C_1|} \leq m^2\delta(1-\epsilon)^{-1} \leq 2m^2\delta.
\end{aligned}$$

Since $d\omega_g$ (or $d\omega_{\tilde{g}}$) is the same as the distribution of an unordered m-tuple drawn from dG (resp., $d\tilde{G}$) *minus* the event that two points are drawn from the same C_j,

$$\left|\int_{\mathfrak{D}} d\omega_g - \int_{\mathfrak{D}'} d\omega_{\tilde{g}}\right| \leq (2\epsilon^{1/2}m^2 + 2/k)n^m/m!\,, \tag{6}$$

where the second summand follows from the fact that

$$1-\binom{k}{m}\cdot\frac{m!}{k^m}\leq 1-(1-\frac{m-1}{k})^m<1/k.$$

Similarly, if we define $dF=k^{-1}\sum_{j=1}^k df_j$, $d\tilde{F}=k^{-1}\sum_{j=1}^s d\tilde{f}_j$, $F(x)=\int_0^x dF$, and $G(x)=\int_0^x dG$, the difference of the two integrals in question is bounded by $\binom{m}{2}\binom{n}{m}$ times the probability that two points chosen from dF are within δ, or

$$\begin{aligned}
&\leq\binom{m}{2}\max_x\int_{x-\delta}^{x+\delta}dF\\
&=\binom{m}{2}\max_x(F(x+\delta)-F(x-\delta))\\
&\leq\binom{m}{2}(\max_x(G(x+\delta+\epsilon)-G(x-\delta-\epsilon))+2\epsilon)\\
&\leq m^2(\delta+2\epsilon)(1-\epsilon)^{-1}\leq 6\epsilon^{1/2}m^2
\end{aligned}$$

and so,

$$\left|\int_{\mathfrak{D}}d\omega_f-\int_{\mathfrak{D}'}d\omega_{\tilde{f}}\right|\leq(6\epsilon^{1/2}m^2+2/k)n^m/m! \tag{7}$$

Putting together (4), (5), (6), and (7), we have

$$\left|\Lambda^\tau(\sigma)-\int_{\mathfrak{D}}d\omega_f\right|<(20\epsilon^{1/2}m^2+4/k)n^m/m!\,.$$

□

7 The Proof of Theorem 1

For the proof of Theorem 1 below, we are heavily indebted to [5], which we find to have the most comprehensible – if not the shortest – proof of the Regularity Lemma in the literature.

For disjoint sets $X,Y\subset\mathbb{Z}_n$, define the "index"

$$q(X,Y)=|X||Y|d^2(X,Y)/n^2.$$

Then extend this definition to a pair of partitions $\mathcal{X}$ of X and $\mathcal{Y}$ of Y by $q(\mathcal{X},\mathcal{Y})=\sum_{X'\in\mathcal{X},Y'\in\mathcal{Y}}q(X',Y')$. For a partition $\mathcal{P}=\{C_j\}_{j=1}^k$ of $\mathbb{Z}_n$, we write $q(\mathcal{P})=\sum_{i,j}q(C_i,C_j)$. If one set in the partition, C_0, has been designated as an *exceptional set*, then we treat C_0 as a collection of singletons in this sum. That is, we write $\tilde{\mathcal{P}}$ for the partition which refines $\mathcal{P}$ by splitting C_0 into singletons, then $q(\mathcal{P})=q(\tilde{\mathcal{P}},\tilde{\mathcal{P}})$. First of all, note that

$$q(\mathcal{P})=\sum_{i,j}\frac{|C_i||C_j|}{n^2}d^2(C_i,C_j)\leq n^{-2}\sum_{i,j}|C_i||C_j|=1.$$

Now, we have the following simple lemma, which says that refinement can only increase the index of a partition.

Lemma 12.

1. *Let $C, D \subset \mathbb{Z}_n$ (not necessarily disjoint). If $\mathcal{C}$ is a partition of C and $\mathcal{D}$ is a partition of D, then $q(\mathcal{C}, \mathcal{D}) \geq q(C, D)$.*

2. *If $\mathcal{P}, \mathcal{P}'$ are partitions of $\mathbb{Z}_n$ and $\mathcal{P}'$ refines $\mathcal{P}$, then $q(\mathcal{P}') \geq q(\mathcal{P})$.*

Proof.

1. Let $\mathcal{C} = \{C_1, \dots, C_k\}$ and $\mathcal{D} = \{D_1, \dots, D_l\}$. Then

$$\begin{aligned} q(\mathcal{C}, \mathcal{D}) &= \frac{1}{n^2} \sum_{i,j} \frac{p(C_i, D_j)^2}{|C_i||D_j|} \\ &\geq \frac{1}{n^2} \frac{(\sum_{i,j} p(C_i, D_j))^2}{\sum_{i,j} |C_i||D_j|} \\ &= \frac{1}{n^2} \frac{p(C, D)^2}{(\sum_i |C_i|)(\sum_j |D_j|)} \\ &= q(C, D), \end{aligned}$$

   where the inequality follows from Cauchy-Schwarz.

2. Let $\mathcal{P} = \{C_1, \dots, C_k\}$, and for $1 \leq i \leq k$, let $\mathcal{C}_i$ be the partition of C_i induced by $\mathcal{P}'$. Then

$$q(\mathcal{P}) = \sum_{i,j} q(C_i, C_j) \leq \sum_{i,j} q(\mathcal{C}_i, \mathcal{C}_j) = q(\mathcal{P}'),$$

   where the inequality follows from part (1).

□

The next lemma says that we may exploit irregular pairs to increase the index somewhat.

Lemma 13. *Let $\epsilon > 0$, and let $C, D \subset \mathbb{Z}_n$ be intervals. If (C, D) is not ϵ-regular with respect to σ, then there are partitions $\mathcal{C}$ and $\mathcal{D}$ of C and D, respectively, so that*

$$q(\mathcal{C}, \mathcal{D}) \geq q(C, D) + \epsilon^4 \frac{|C||D|}{n^2}.$$

Proof. Suppose (C, D) is not ϵ-regular, and choose intervals $C_1 \subset C$ and $D_1 \subset D$ with $|C_1| > \epsilon|C|$ and $|D_1| > \epsilon|D|$ so that $|\eta| > \epsilon$, where $\eta = d(C_1, D_1) - d(C, D)$. Let $\mathcal{C} = \{C_1, C_2, C_3\}$ and $\mathcal{D} = \{D_1, D_2, D_3\}$, where C_2 is the "left half" of $C \setminus C_1$; C_3 is the "right half"; and D_2 and D_3 are defined similarly. (That is, C_1 splits the interval C into three pieces: C_1 itself, one interval of points less than those of C_1 and one interval of points greater than those of C_1. Either, but not both, of these may be empty.)

For ease of notation, write $c_i = |C_i|$, $d_i = |D_i|$, $e_{ij} = p(C_i, D_j)$, $c = |C|$, $d = |D|$, and $e = p(C, D)$. Then, applying Cauchy-Schwarz again, we see

$$\begin{aligned} q(\mathcal{C}, \mathcal{D}) &= \frac{1}{n^2} \sum_{i,j} \frac{e_{ij}^2}{c_i d_j} = \frac{1}{n^2} \left(\frac{e_{11}^2}{c_1 d_1} + \sum_{(i,j) \neq (1,1)} \frac{e_{ij}^2}{c_i d_j} \right) \\ &\geq \frac{1}{n^2} \left(\frac{e_{11}^2}{c_1 d_1} + \frac{(e - e_{11})^2}{cd - c_1 d_1} \right). \end{aligned}$$

Since $e_{11} = c_1 d_1 e/cd + \eta c_1 d_1$, we have

$$\begin{aligned} n^2 q(\mathcal{C}, \mathcal{D}) &\geq \frac{1}{c_1 d_1} \left(\frac{c_1 d_1 e}{cd} + \eta c_1 d_1 \right)^2 + \frac{1}{cd - c_1 d_1} \left(\frac{cd - c_1 d_1}{cd} e - \eta c_1 d_1 \right)^2 \\ &= \frac{e^2}{cd} + \frac{\eta^2 c_1 d_1 cd}{cd - c_1 d_1} \\ &\geq \frac{e^2}{cd} + \epsilon^4 cd \end{aligned}$$

since $c_1 \geq \epsilon c$, $d_1 \geq \epsilon d$, and $\eta^2 > \epsilon^2$. □

The following lemma is the crux of the proof of Theorem 1.

Lemma 14. *Let $0 < \epsilon \leq 1/4$ and $k \in \mathbb{N}$, let σ be a permutation of $\mathbb{Z}_n$, and let $\mathcal{P}$ be an equitable partition of $\mathbb{Z}_n$ into $\{C_j\}_{j=0}^k$ with $|C_0| \leq \epsilon n$ and $|C_j| \geq 81^k$ for $j > 0$. If $\mathcal{P}$ is not ϵ-regular, then there is an equitable partition $\mathcal{P}' = \{C'_j\}_{j=0}^l$ of $\mathbb{Z}_n$ with exceptional set C'_0, where $k \leq l \leq k81^k$, such that $|C'_0| \leq |C_0| + n/9^k$ and*

$$q(\mathcal{P}') \geq q(\mathcal{P}) + \epsilon^5/2.$$

Proof. Let $c = |C_1|$. For all $1 \leq i, j \leq k$, define a partition $\mathcal{C}_{ij}^1$ of C_i and a partition $\mathcal{C}_{ji}^2$ of C_j as follows. If the pair (C_i, C_j) is ϵ-regular, then let $\mathcal{C}_{ij}^1 = \{C_i\}$ and $\mathcal{C}_{ji}^2 = \{C_j\}$. If not, then by the previous lemma, there are tripartitions $\mathcal{C}_{ij}^1$ and $\mathcal{C}_{ji}^2$ of C_i and C_j, respectively, so that

$$q(\mathcal{C}_{ij}^1, \mathcal{C}_{ji}^2) \geq q(C_i, C_j) + \frac{\epsilon^4 c^2}{n^2}.$$

For each $i = 1, \ldots, k$, let $\mathcal{C}_i$ be the partition of C_i that is the common refinement of every partition $\mathcal{C}_{ij}^r$. Note that $|\mathcal{C}_i| \leq 9^k$. Now, consider the partition

$$\mathcal{C} = \{C_0\} \cup \bigcup_{i=1}^k \mathcal{C}_i,$$

with C_0 as exceptional set. Then $\mathcal{C}$ refines $\mathcal{P}$ and $k \leq |\mathcal{C}| \leq k9^k$.

Let $\mathcal{C}_0 = \{\{x\} : x \in C_0\}$. If $\mathcal{P}$ is not ϵ-regular, then for more than ϵk^2 of the pairs (C_i, C_j), the partitions $\mathcal{C}_{ij}^1$ and $\mathcal{C}_{ji}^2$ are nontrivial. Hence, by Lemma 12,

$$\begin{aligned}
q(\mathcal{C}) &= \sum_{i,j\geq 1} q(\mathcal{C}_i, \mathcal{C}_j) + \sum_{i\geq 0} q(\mathcal{C}_0, \mathcal{C}_i) \\
&\geq \sum_{i,j\geq 1} q(\mathcal{C}_{ij}^1, \mathcal{C}_{ji}^2) + \sum_{i\geq 1} q(\mathcal{C}_0, \{C_i\}) + q(\mathcal{C}_0) \\
&\geq \sum_{i,j\geq 1} q(C_i, C_j) + \epsilon k^2 \frac{\epsilon^4 c^2}{n^2} + \sum_{i\geq 1} q(\mathcal{C}_0, \{C_i\}) + q(\mathcal{C}_0) \\
&= q(\mathcal{P}) + \epsilon^5 \left(\frac{kc}{n}\right)^2 \\
&\geq q(\mathcal{P}) + \epsilon^5/2,
\end{aligned}$$

since $kc \geq 3n/4$.

Now, $\mathcal{C}$ satisfies the conclusions of the theorem, except that it may not be equitable. To fix the situation, cut each non-exceptional block of $\mathcal{C}$ into a maximal collection of disjoint intervals of size $d = \lfloor c/81^k \rfloor$. Call the resulting set of intervals $\{C'_j\}_{j=1}^l$, and let $C'_0 = \mathbb{Z}_n \setminus \bigcup C'_j$. This new partition $\mathcal{P}'$ refines $\mathcal{C}$, so

$$q(\mathcal{P}') \geq q(\mathcal{C}) \geq q(\mathcal{P}) + \epsilon^5/2.$$

Since each set C'_j, $j > 0$, is contained in one of the sets C_i, but not more than 81^k sets can lie inside the same C_j, we also have $k \leq l \leq k81^k$. On the other hand, the sets $C'_1, \ldots, C'_l$ use all but at most d points from each nonexceptional block of $\mathcal{C}$. Therefore,

$$\begin{aligned}
|C'_0| &\leq |C_0| + d|\mathcal{C}| \\
&\leq |C_0| + \frac{c}{81^k} k 9^k \\
&= |C_0| + ck9^{-k} \\
&\leq |C_0| + n9^{-k}.
\end{aligned}$$

□

Now, since $q(\mathcal{P}) \leq 1$, this lemma cannot be applied ad infinitum. Indeed, we may now complete the proof of Theorem 1.

Proof of Theorem 1. Let $\epsilon > 0$ and $m \geq 1$. Without loss of generality, $\epsilon \leq 1/4$. Let $s = \lceil 2/\epsilon^5 \rceil$. If we apply Lemma 14 s times, we end up with a partition of σ which is ϵ-regular.

However, we need to choose the "starting" partition so that the exceptional set ends up with cardinality $\leq \epsilon n$ and the C_j are sufficiently large at each stage. With each iteration of the lemma, the size of the exceptional set can grow by at most $n/9^k$. Therefore, we wish to choose k large enough so that s increments of $n/9^k$ add up to at most $\epsilon n/2$, and

n large enough so that $|C_0| < k$ implies $|C_0| \leq \epsilon n/2$. (We can guarantee $|C_0| < k$ if we begin with an equitable k-partition.) So let $k \geq m$ be large enough so that $9^k \geq 2s/\epsilon$. Then $s/9^k \leq \epsilon/2$, and hence

$$k + \frac{s}{9^k}n \leq \epsilon n$$

whenever $n \geq 2k/\epsilon$.

Now, define $f(x) = x81^x$. We may take $M = \max\{f^s(k), 2k/\epsilon\}$. To deal with the second condition – that the blocks be sufficiently large at each stage – note that, after s steps, the nonexceptional blocks sizes are at least $n/(2M)^s$. Therefore, choosing $N = \max\{2M/\epsilon, 81^M(2M)^s\}$ suffices, and the proof is complete. $\square$

8 Concluding Remarks

The discussion of Section 3 is largely "local", i.e., the analysis is concerned with the internal structure of individual blocks of the uniform partition. Section 6 consists of a "global" analysis – it does not take into account the internal structure of the blocks, only their relationships with one another. On the other hand, the proofs of the main results of Sections 4 and 5, as well as that of Proposition 5, are both. It is here, in the interplay between local and global, that we believe the most interesting behavior resides. We suspect that such dual analysis may lead to a better understanding of extremal permutations in the senses of Problem 2 and 3 of the Introduction, perhaps using the results of Section 6. Theorem 11, in theory, gives a translation of these problems from combinatorial to analytic. We are hopeful that Theorem 11 can find application in algorithmic settings, e.g., in the vein of [6], or in other contexts where a "counting lemma" has been useful, such as the hypergraph-theoretic proof of the Szemerédi Theorem.

9 Acknowledgements

Thank you to Jim Propp, Vera Sós, and Joel Spencer for stimulating questions and invaluable discussions. Thanks also to the referee for helpful comments and suggestions.

References

[1] N. Alon and E. Friedgut, On the number of permutations avoiding a given pattern, *J. Comb. Theory Ser. A* **89** (2000), 133–140.

[2] F. R. K. Chung, R. L. Graham, and R. M. Wilson, Quasi-random graphs, *Combinatorica* **9** (1989), 345–362.

[3] J. N. Cooper, Quasirandom permutations, *J. Comb. Theory Ser. A* **106** 1 (2004), 123–143.

[4] J. N. Cooper, Quasirandom arithmetic permutations, J. Number Theory **114** (2005), no. 1, 153–169.

[5] R. Diestel, Graph Theory. Second Edition. *Graduate Texts in Mathematics* **173**, Springer-Verlag, New York, 2000.

[6] R. A. Duke, H. Lefmann, V. Rödl, A fast approximation algorithm for computing the frequencies of subgraphs in a given graph, SIAM J. Comput. 24 (1995), no. 3, 598–620.

[7] M. Hildebrand, B. E. Sagan, and V. R. Vatter, Bounding quantities related to the packing density of $1(l+1)l\cdots 2$, Adv. in Appl. Math. **33** (2004), no. 3, 633–653.

[8] Y. Kohayakawa, V. Rödl, Szemerédi's regularity lemma and quasi-randomness, *Recent advances in algorithms and combinatorics*, 289-351, CMS Books Math./Ouvrages Math. SMC, **11**, Springer, New York, 2003.

[9] J. Komlós and M. Simonovits, Szemerédi's regularity lemma and its applications in graph theory, *Combinatorics – Paul Erdős is Eighty*, vol. 2, D. Miklós, V. T. Sós, and T. Szőnyi, eds., Bolyai Mathematical Studies, pages 295–352. János Bolyai Mathematical Society, Budapest, Budapest, 1996.

[10] A. Marcus and G. Tardos, Excluded permutation matrices and the Stanley-Wilf conjecture, J. Combin. Theory Ser. A **107** (2004), no. 1, 153–160.

[11] M. Simonovits and V. T. Sós, Szemerédi's partition and quasirandomness, *Random Structures Algorithms* **2** (1991), no. 1, 1–10.

[12] E. Szemerédi, Regular partitions of graphs, *Problèmes Combinatoires et Théorie des Graphes*, Colloques Internationaux CNRS n. 260, Orsay, 1976.

[13] H. S. Wilf, The patterns of permutations, *Discrete Math.* **257** (2002), no. 2-3, 575-583.