# Combinatorics of Sets

Po-Shen Loh

June 2011

## 1 Warm-Up

1. Let $\mathcal{F}$ be a collection of subsets $A_1, A_2, \ldots$ of $\{1, \ldots, n\}$, such that for each $i \neq j$, $A_i \cap A_j \neq \emptyset$. Prove that $\mathcal{F}$ has size at most $2^{n-1}$.

   **Solution:** For each set $S \in 2^{[n]}$, observe that at most one of $S$ and $\overline{S}$ is contained in $\mathcal{F}$.

2. Suppose that $\mathcal{F}$ above has size exactly $2^{n-1}$. Must there be a common element $x \in \{1, \ldots, n\}$ which is contained by every $A_i$?

   **Solution:** No. First, observe that in [3], one can create an intersecting family with the four sets $\{1, 2\}$, $\{2, 3\}$, $\{3, 1\}$, and $\{1, 2, 3\}$. Then for every $n \geq 3$, one can blow up this construction by taking all sets $S$ which are obtained by taking the union of one of these 4 sets together with an arbitrary subset of $\{4, \ldots, n\}$.

3. Let $\mathcal{F}$ be a family of sets, each of size exactly 3, such that:

   (a) Every pair of sets intersects in a single element.

   (b) Every pair of elements in the ground set $X = \bigcup_{L \in \mathcal{F}} S$ is contained in a unique set $L \in \mathcal{F}$.

   Suppose that $\mathcal{F}$ has more than one set. Prove that the ground set $X$ has exactly 7 elements, and show that such a family $\mathcal{F}$ exists.

   **Solution:** Let $L$ be the number of "lines," i.e., sets in $\mathcal{F}$, and let $n$ be the size of the ground set. Observations:

   - The degree of any element is exactly $\frac{n-1}{2}$, because of (b).
   - The number of lines is $L = \frac{1}{3}\binom{n}{2}$ because of (a).

   But by double-counting, we must have:

$$\sum_{v \in X} \binom{d_v}{2} = \binom{L}{2}.$$

   Substituting our expressions for $d_v$ and $L$, we conclude that

$$\begin{aligned} n \cdot \frac{1}{2}\left(\frac{n-1}{2} \cdot \frac{n-3}{2}\right) &= \frac{L(L-1)}{2} = \frac{n(n-1)}{6} \cdot \frac{(n-3)(n+2)}{6} \\ \frac{1}{4} &= \frac{n+2}{36} \\ 7 &= n\,. \end{aligned}$$

   The Fano plane provides the desired construction.

## 2 Designs

1. (TST 2005/1.) Let $n$ be an integer greater than 1. For a positive integer $m$, let $X_m = \{1, 2, \dots, mn\}$. Suppose that there exists a family $\mathcal{F}$ of $2n$ subsets of $X_m$ such that:

   (a) each member of $\mathcal{F}$ is an $m$-element subset of $X_m$;

   (b) each pair of members of $\mathcal{F}$ shares at most one common element;

   (c) each element of $X_m$ is contained in exactly 2 elements of $\mathcal{F}$.

   Determine the maximum possible value of $m$ in terms of $n$.

   **Solution:** Count in two ways:

$$\sum_v \binom{d_v}{2} = \#\text{ set pairs intersect} \le \binom{2n}{2}.$$

   But all $d_v = 2$, so we have $mn \le n(2n-1)$, i.e., $m \le 2n-1$. Equality is possible: take $2n$ lines in general position in $\mathbb{R}^{\nvDash}$, and let their $\binom{2n}{2} = mn$ intersection points be the points.

2. (USAMO 2011/6.) Let $X$ be a set with $|X| = 225$. Suppose further that there are eleven subsets $A_1, \dots, A_{11}$ of $X$ such that $|A_i| = 45$ for $1 \le i \le 11$ and $|A_i \cap A_j| = 9$ for $1 \le i < j \le 11$. Prove that $|A_1 \cup \cdots \cup A_{11}| \ge 165$, and give an example for which equality holds.

   **Solution:** The $|X| = 225$ condition is unnecessary. Count in two ways:

$$\sum_{v \in X} \binom{d_v}{2} = \binom{11}{2} \cdot 9 = 495.$$

   But also $\sum_{v \in X} d_v = 11 \cdot 45 = 495$. This suggests that equality occurs when all $d_v = \binom{d_v}{2}$, which is precisely at $d_v = 3$.

   Indeed, by Cauchy-Schwarz, if we let Let $n = |X|$, we find

$$\begin{gathered}\left(\sum_{v \in X} 1 d_v\right)^2 \le \left(\sum_{v \in X} 1\right)\left(\sum_{v \in X} d_v^2\right) \\ 495^2 \le n \cdot 3 \cdot 495,\end{gathered}$$

   where we deduced $\sum d_v^2 = 3 \cdot 495$ from $\sum \binom{d_v}{2}$ and $\sum d_v$. We conclude that $n \ge 165$.

   For the construction, take 11 planes in general position in $\mathbb{R}^3$, and let their $\binom{11}{3} = 165$ points of intersection be the points in the set.

3. A collection of subsets $L_1, \dots, L_m$ in the universe $\{1, \dots, n\}$ is called a *projective plane* if:

   (a) Every pair of sets (called "lines") intersects in a single element.

   (b) Every pair of elements in the ground set $X = \bigcup_{L \in \mathcal{F}} S$ is contained in a unique set $L \in \mathcal{F}$.

   Actually, there are two families of degenerate planes which satisfy the two conditions above, but are not considered to be projective planes. They are:

   (a) $L_1 = \{1, \dots, n\}$, $L_2 = \{1\}$, $L_3 = \{1\}$, $L_4 = \{1\}$, ...

   (b) $L_1 = \{2, 3, \dots, n\}$, $L_2 = \{1, 2\}$, $L_3 = \{1, 3\}$, $L_4 = \{1, 4\}$, ..., $L_n = \{1, n\}$.

   It is well-known that for every projective plane, there is an $N$ (called the "order" of the plane) such that:

(a) Every line contains exactly $N+1$ points, and every point is on exactly $N+1$ lines.

(b) The total number of points is exactly $N^2+N+1$, which is the same as the total number of lines.

4. For every prime power $p^n$, there exists a projective plane of that order.

   **Solution:** Take the finite field $\mathbb{F}_q$ of order $q=p^n$. Let the points be dimension-1 subspaces of $\mathbb{F}_q^3$, and let the lines be dimension-2 subspaces (i.e., the collection of dimension-1 subspaces that lie within a given dimension-2 subspace).

   Every pair of dimension-2 subspaces intersects in a subspace, which has integer dimension. That dimension is obviously strictly less than 2, but also cannot be 0 or else we would have had mutually independent bases of the pair of dimension-2 subspaces, already yielding dimension 4 for the parent space. Hence every pair of dimension-2 subspaces intersects in a dimension-1 subspace, i.e., containing a single point.

   Also, every pair of dimension-1 subspaces determines a pair of linearly independent basis vectors, whose span is already dimension-2. So For every pair of dimension-1 subspaces, there is a unique dimension-2 subspace that contains both of them.

   Observe that by calculation, the number of dimension-1 subspaces inside a fixed dimension-2 subspace is exactly $\frac{q^2-1}{q-1}=q+1$, because there are $q^2$ nonzero vectors in the dimension-2 subspace, and every pair of dimension-1 subspaces intersects only at the zero vector. Hence the order is $q=p^n$.

5. (Open.) What are the possible orders of projective planes? All known projective planes have prime power order, but it is unknown whether, for example, there is a projective plane of order 12.

# 3 Graphs and partitioning

1. Construct a bipartite graph in which all degrees are equal, and every pair of vertices on the same side has exactly 1 common neighbor. Show that this must achieve the maximum possible number of edges in any $C_4$-free bipartite graph with the same number of vertices.

   **Solution:** Take a projective plane, and create a graph by putting the points on the left, and the lines on the right, connecting a point to a line if they are incident. The codegree conditions are satisfied by the definition of a projective plane.

   Now suppose we have a $C_4$-free bipartite graph with $n$ vertices on each side. Count the number of $K_{1,2}$ with 2 points on the Right side. This is

$$\binom{n}{2} \geq \sum_v \binom{d_v}{2} \geq n\binom{\overline{d}}{2},$$

   with equality precisely when the degrees are all equal, and when every pair on the Right side has codegree exactly 1. Since we are optimizing the average degree, this implies the result.

2. Construct a non-bipartite graph in which all degrees are equal, and every pair of vertices has exactly 1 common neighbor. Show that this must achieve the maximum possible number of edges in any $C_4$-free graph with the same number of vertices.

   **Solution:** Consider $\mathbb{F}^3$, and associate every vertex with a 1-dimensional subspace. Join two vertices by an edge if their two subspaces are orthogonal. But given two distinct subspaces, there is a unique 1-dimensional subspace orthogonal to both of them.

   Another way to see this is to say that we join two vertices if one of the dimension-1 subspaces lies within the orthogonal complement (dimension-2 subspace) of the other vertex. Then if we are given two distinct dimension-1 subspaces, the number of common neighbors is the number of common dimension-2 subspaces that contain both, and this is the same calculation as above.

   For tightness, take the same sum of all $K_{1,2}$ throughout the graph.

3. Let $n$ be odd. Partition the edge set of $K_n$ into $n$ matchings with $\frac{n-1}{2}$ edges each.

   **Solution:** Spread the $n$ vertices around a circle. Take parallel classes.

4. Let $n$ be even. Partition the edge set of $K_n$ into $n-1$ matchings with $\frac{n}{2}$ edges each.

   **Solution:** Spread $n-1$ vertices around a circle, and let the final vertex be the origin. Take parallel classes, along with the orthogonal radius from the origin.

5. Find (nontrivial) infinite families of $t$ and $n$ for which it is possible to partition the edges of $K_n$ into disjoint copies of edges corresponding to $K_t$.

   **Solution:** Suppose $n = q^2 + q + 1$ for a prime power $q$. Take a projective plane, and identify the vertices of $K_n$ with the points. The cliques $K_t$ correspond to the lines.

   Alternatively, suppose $n = q^2$ for a prime power $q$. Take the affine plane, i.e., points corresponding to $\mathbb{F}_q^2$. Lines correspond to the points in affine lines, i.e., with $q$ points each. It is clear that every pair of points determines a unique line, hence every edge is in a unique $K_q$.

# 4 Extremal set theory

1. (Erdős-Ko-Rado.) Let $n \geq 2k$ be positive integers, and let $\mathcal{C}$ be a collection of pairwise-intersecting $k$-element subsets of $\{1, \ldots, n\}$, i.e., every $A, B \in \mathcal{C}$ has $A \cap B \neq \emptyset$. Prove that $|\mathcal{C}| \leq \binom{n-1}{k-1}$.

   **Remark.** This corresponds to the construction which takes all subsets that contain the element 1.

   **Solution:** Pick a random $k$-set $A$ from $2^{[n]}$ by first selecting a random permutation $\sigma \in S_n$, and then picking a random index $i \in [n]$. Then define $A = \{\sigma(i), \ldots, \sigma(i+k-1)\}$, with indices after $n$ wrapping around, of course. It suffices to show that $\mathbb{P}[A \in \mathcal{C}] \leq k/n$.

   Let us show that conditioned on any fixed $\sigma$, $\mathbb{P}[A \in \mathcal{C} | \sigma] \leq k/n$, which will finish our problem. But this is equivalent to the statement that $\mathcal{C}$ can only contain $\leq k$ intervals (wrapping after $n$) of the form $\{i, \ldots, i+k-1\}$, which is easy to show.

2. (Non-uniform Fisher's inequality.) Let $\mathcal{C} = \{A_1, \ldots, A_r\}$ be a collection of distinct subsets of $\{1, \ldots, n\}$ such that every pairwise intersection $A_i \cap A_j$ $(i \neq j)$ has size $t$, where $t$ is some fixed integer between 1 and $n$ inclusive. Prove that $|\mathcal{C}| \leq n$.

   **Solution:** Consider the $n \times r$ matrix $\mathbf{A}$, where the $i$-th column of $\mathbf{A}$ is the characteristic vector of $A_i$. Then, $\mathbf{A}^T\mathbf{A}$ is a $r \times r$ matrix, all of whose off-diagonal entries are $t$. We claim that the diagonal entries are all $> t$. Indeed, if there were some $|A_i|$ which were exactly $t$, then the structure of $\mathcal{C}$ must look like a "flower," with one set $A_j$ of size $t$, and all other sets fully containing $A_j$ and disjointly partitioning the elements of $[n] \setminus A_j$ among them. Any such construction has size at most $1 + (n-t) \leq n$, so we would already be done.

   Therefore, $\mathbf{A}^T\mathbf{A}$ is nonsingular by Lemma below, and the previous argument again gives $r \leq n$.

   **LEMMA:** Let $\mathbf{A}$ be a square matrix over $\mathbb{R}$, for which all non-diagonal entries are all equal to some $t \geq 0$, and all diagonal entries are strictly greater than $t$. Then $\mathbf{A}$ is nonsingular.

   **Proof.** If $t = 0$, this is trivial. Now suppose $t > 0$. Let $\mathbf{J}$ be the all-ones square matrix, and let $\mathbf{D} = \mathbf{A} - t\mathbf{J}$. Note that $\mathbf{D}$ is nonzero only on the diagonal, and in fact strictly positive there. We would like to solve $(t\mathbf{J} + \mathbf{D})\mathbf{x} = \mathbf{0}$, which is equivalent to $\mathbf{D}\mathbf{x} = -t\mathbf{J}\mathbf{x}$. Let $s$ be the sum of all elements in $\mathbf{x}$, and let the diagonal entries of $\mathbf{D}$ be $d_1, \ldots, d_n$, in order. Then, we have $d_i x_i = -ts \Rightarrow x_i = -(t/d_i)s$. But since $t$ and $d_i$ are both strictly postive, this forces every $x_i$ to have opposite sign from $s$, which is impossible unless all $x_i = 0$. Therefore, $\mathbf{A}$ is nonsingular.

# 5 Combinatorics and geometry

1. (Happy ending problem.) Given any 5 distinct points in the plane, no 3 collinear, show that some 4 are in *convex position*, i.e., forming the vertices of a convex quadrilateral.

   **Solution:** If the convex hull has size 4 or more, already done. Otherwise, convex hull is a triangle $ABC$ and there are 2 internal points. Consider the line $\ell$ determined by those internal points. Since no 3 collinear, it divides $\{A, B, C\}$ into two groups, one of which has size 2. Use those 2, plus the 2 internal points.

2. (Erdős-Szekeres.) For every integer $n$, there is some finite $N$ such that the following holds. Given any $N$ distinct points in the plane, no 3 collinear, some $n$ are in convex position.

   **Remark.** It is conjectured that $N = 1 + 2^{n-2}$ suffices for all $n \geq 3$, and known that $N \geq 1 + 2^{n-2}$ is required. The best known upper bound is of order $4^n/\sqrt{n}$.

   **Solution:** If a set $S$ of points has the property that every 4-subset is in convex position, then all of $S$ is in convex position. To see this, suppose there was some point $P$ strictly inside the convex hull of $S$. Triangulate the convex hull using diagonals, and $P$ will be strictly inside one of the triangles, say $ABC$. Then $PABC$ is concave, contradiction.

   Now suppose we have $R^{(4)}(5, n)$ many points. For every 4-set of points, color the corresponding 4-edge red if they are not in convex position, blue otherwise. Our hypergraph Ramsey bound implies that there must either be 5 vertices with all 4-edges red, or $n$ vertices with all 4-edges blue.

   But the 5 vertices with all 4-edges red contradicts the Happy Ending Problem, so we must have the latter. By our opening remark, in fact the $n$ points are all in convex position.

3. (Caratheodory.) A *convex combination* of points $x_i$ is defined as a linear combination of the form $\sum_i \alpha_i x_i$, where the $\alpha_i$ are non-negative coefficients which sum to 1.

   Let $X$ be a finite set of points in $\mathbb{R}^d$, and let $\text{cvx}(X)$ denote the set of points in the convex hull of $X$, i.e., all points expressible as convex combinations of the $x_i \in X$. Show that each point $x \in \text{cvx}(X)$ can in fact be expressed as a convex combination of only $d + 1$ points of $X$.

   **Solution:** Given a convex combination with $d + 2$ or more nonzero coefficients, find a new vector with which to perturb the nonzero coefficients. Specifically, seek $\sum_i \beta_i x_i = 0$ and $\sum_i \beta_i = 0$, which is $d + 1$ equations, but with $d + 2$ variables $\beta_i$. So there is a non-trivial solution, and we can use it to reduce another $\alpha_i$ coefficient to zero.

4. (Radon.) Let $A$ be a set of at least $d + 2$ points in $\mathbb{R}^d$. Show that $A$ can be split into two disjoint sets $A_1 \cup A_2$ such that $\text{cvx}(A_1)$ and $\text{cvx}(A_2)$ intersect.

   **Solution:** For each point, create an $\mathbb{R}^{d+1}$-vector $v_i$ by adding a "1" as the last coordinate. We have a non-trivial dependence because we have at least $d + 2$ vectors in $\mathbb{R}^{d+1}$, say $\sum_i \alpha_i v_i = 0$. Split $A = A_1 \cup A_2$ by taking $A_1$ to be the set of indices $i$ with $\alpha_i \geq 0$, and $A_2$ to be the rest.

   By the last coordinate, we have
   $$\sum_{i \in A_1} \alpha_i = \sum_{i \in A_2} (-\alpha_i) \,.$$

   Let $Z$ be that sum. Then if we use $\alpha_i/Z$ as the coefficients, we get a convex combination from $A_1$ via the first $d$ coordinates, which equals the convex combination from $A_2$ we get by using $(-\alpha_i)/Z$ as the coefficients.

5. (Helly.) Let $C_1, C_2, \ldots, C_n$ be sets of points in $\mathbb{R}^d$, with $n \geq d + 1$. Suppose that every $d + 1$ of the sets have a non-empty intersection. Show that all $n$ of the sets have a non-empty intersection.

   **Solution:** Induction on $n$. Clearly true for $n = d + 1$, so now consider $n \geq d + 2$, and assume true for $n - 1$. Then by induction, we can define points $a_i$ to be in the intersection of all $C_j$, $j \neq i$. Apply Radon's Lemma to these $a_i$, to get a split of indices $A \cup B$.

Crucially, note that for each $i \in A$ and $j \in B$, the point $a_i$ is in $C_j$. So, each $i \in A$ gives $a_i \in \bigcap_{j \in B} C_j$, and hence the convex hull of points in $A$ is entirely contained in all $C_j$, $j \in B$.

Similarly, the convex hull of points in $B$ is entirely contained in all $C_j$, $j \in A$. Yet Radon's Lemma gave intersecting convex hulls, so there is a point in both hulls, i.e., in all $C_j$, $j \in A \cup B = [n]$.

# 6 Bonus problems

1. (From Peter Winkler.) The 60 MOPpers were divided into 8 teams for Team Contest 1. They were then divided into 7 teams for Team Contest 2. Prove that there must be a MOPper for whom the size of her team in Contest 2 was strictly larger than the size of her team in Contest 1.

   **Solution:** In Contest 1, suppose the team breakdown was $s_1 + \cdots + s_8 = 60$. Then in the $i$-th team, with $s_i$ people, say that each person did $\frac{1}{s_i}$ of the work. Similarly, in Contest 2, account equally for the work within each team, giving scores of $\frac{1}{s_i'}$.

   However, the total amount of work done by all people in Contest 1 was then exactly 8, and the total amount of work done by all people in Contest 2 was exactly 7. So somebody must have done strictly less work in Contest 2. That person saw
   $$\frac{1}{s_i'} < \frac{1}{s_i},$$
   i.e., the size of that person's team on Contest 2 was strictly larger than her team size on Contest 1.

2. (MOP 2008.) Let $\mathcal{F}$ be a collection of $2^{n-1}$ subsets $A_1, A_2, \ldots$ of $\{1, \ldots, n\}$, such that for each $i \neq j \neq k$, $A_i \cap A_j \cap A_k \neq \emptyset$. Prove that there is a common element $x \in \{1, \ldots, n\}$ that is contained in every $A_i$.

3. (Sperner capacity of cyclic triangle, also Iran 2006.) Let $A$ be a collection of vectors of length $n$ from $\mathbb{Z}_3$ with the property that for any two distinct vectors $a, b \in A$ there is some coordinate $i$ such that $b_i = a_i + 1$, where addition is defined modulo 3. Prove that $|A| \leq 2^n$.

   **Solution:** For each $a \in A$, define the $\mathbb{Z}_3$-polynomial $f_a(\mathbf{x}) := \prod_{i=1}^n (x_i - a_i - 1)$. Observe that this is multilinear. Clearly, for all $a \neq b \in A$, $f_a(b) = 0$, and $f_a(a) \neq 0$; therefore, the $f_a$ are linearly independent, and bounded in cardinality by the dimension of the space of multilinear polynomials in $n$ variables, which is $2^n$.