Discrete Mathematics
REU 2004. Info:

Instructor: László Babai

8/13/04
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Points in the plane</td>
<td>7</td>
</tr>
<tr>
<td>1.1</td>
<td>Games</td>
<td>7</td>
</tr>
<tr>
<td>1.2</td>
<td>Ramsey Numbers</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>Games continued</td>
<td>11</td>
</tr>
<tr>
<td>2.1</td>
<td>Hamilton cycles</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>Graph theory</td>
<td>15</td>
</tr>
<tr>
<td>3.1</td>
<td>Inequalities</td>
<td>16</td>
</tr>
<tr>
<td>3.2</td>
<td>Graphs without short cycles</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>Graph theory continued</td>
<td>19</td>
</tr>
<tr>
<td>4.1</td>
<td>Puzzles</td>
<td>19</td>
</tr>
<tr>
<td>4.2</td>
<td>Set game and Szemerédi’s theorem</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>Density versions of coloring theorems</td>
<td>23</td>
</tr>
<tr>
<td>5.1</td>
<td>Hypergraphs</td>
<td>24</td>
</tr>
<tr>
<td>5.2</td>
<td>Steiner Triple Systems</td>
<td>25</td>
</tr>
<tr>
<td>5.3</td>
<td>Projective planes</td>
<td>27</td>
</tr>
<tr>
<td>5.4</td>
<td>Galois Planes</td>
<td>28</td>
</tr>
<tr>
<td>5.5</td>
<td>Internet and library resources</td>
<td>30</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>6</td>
<td>6.1 North-South Game</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>6.2 Number of unit distances in the plane $&lt; n^{3/2}$</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>6.3 Max distance occurs $\leq n$ times</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>6.4 Fitting $10^{20} + 1$ squares on a $(10^{10} + 0.1) \times (10^{10} + 0.1)$ square</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>6.5 Tiling $a \times b$ rectangle by rectangles with at least one integer dimension</td>
<td>32</td>
</tr>
<tr>
<td>7</td>
<td>7.1 Fishermen's clubs</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>7.2 Ramsey Theory</td>
<td>36</td>
</tr>
<tr>
<td>8</td>
<td>8.1 Steiner Triple Systems</td>
<td>39</td>
</tr>
<tr>
<td>9</td>
<td>9.1 Extremal Set Theory</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>9.1.1 Even town</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>9.1.2 Oddtown</td>
<td>44</td>
</tr>
<tr>
<td>10</td>
<td>10.1 Cauchy-Hilbert Matrix</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>10.2 Cauchy's Functional Equation</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>10.3 Number Theory</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>10.4 Ramsey Theory</td>
<td>49</td>
</tr>
<tr>
<td>11</td>
<td>11.1 Totally unimodular matrices</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>11.2 Latin squares</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>11.3 Counting perfect matchings in a bipartite graph; the permanent</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>11.4 Doubly Stochastic Matrices</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>11.5 Back to Latin squares</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>11.6 Orthogonal Latin Squares</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>54</td>
</tr>
</tbody>
</table>
## CONTENTS

12

12.1 Constructive Proofs of Negative Results in Ramsey Theory ........................................ 57
12.2 Bipartite Ramsey Theory ......................................................................................... 58
12.3 Hadamard Matrices .............................................................................................. 58

13

13.1 The Gale-Berlekamp Switching game ...................................................................... 61
13.2 The RAY-CHAUDHURI - WILSON theorem ............................................................ 63

14

14.1 Points in general position ...................................................................................... 65
14.2 Pattern in proofs of inequalities using the Linear Algebra Method ......................... 65
14.3 Additional Exercises in Probability Theory ............................................................ 67

15

15.1 Bases for Vector Spaces of Polynomials ................................................................ 69
15.2 Projective Representation of a Graph ..................................................................... 70
15.3 The Number of Zero-Patterns of a Sequence of Polynomials ................................. 71

16

16.1 Random Variables ................................................................................................ 73
16.2 Independence ......................................................................................................... 73
16.3 Conditional Probability .......................................................................................... 74

17

17.1 Puzzles .................................................................................................................. 77
17.2 Statistical Independence vs Linear Independence .................................................. 77
17.3 Algorithmic Application ....................................................................................... 78
17.4 Algebraic Coding Theory ..................................................................................... 79

18

18.1 Projective dimension of a graph, continued ............................................................. 81
18.2 Theorem of Milnor-Thom ..................................................................................... 83
19 Matrix rigidity (Valiant 1978) ........................................................................... 85

20 Matrix Rigidity continued ....................................................................................... 89

21 Linear Programming .............................................................................................. 91
   21.2 Shannon Capacity of a graph ............................................................................. 92
      21.2.1 Lovász’s theta function (Lovász’s capacity) ........................................... 94

22 0, 1-measures .......................................................................................................... 95
   22.2 Sign-Rigidity (Paturi-Simon) .......................................................................... 95
Chapter 1

1st day, Monday 6/21/04 (Scribe: Charilaos Skiadas, Eric Patterson, and Travis Schedler)

1.1 Points in the plane

Exercise 1.1 (Sylvester). Given \( n \) points in the plane, not all on the same line, show that there is a line in the plane that passes through exactly two of them. What if we switch the role of points and lines, i.e. given \( n \) lines in the plane such that not all of them pass through the same point, is it true that there is a point that belongs to exactly two of them? What if in addition, no two of the lines are parallel?

Exercise 1.2 (Erdős-E. Klein). Show that there is a constant \( c \), independent of \( n \), such that given \( n \) points in the plane, the number of pairs of points with distance 1 is less than \( cn^{3/2} \).

Note: It is known that this number is in fact less that \( cn^{5/4} \) (for possibly a different constant \( c \)).

OPEN QUESTION: Is it less than \( cn^{1.1} \)?

1.2 Games

Exercise 1.3. Let \( n \) be even, and consider an \( n \times n \) chessboard, with two opposite corners removed. Show that it cannot be covered with dominoes. (A domino covers two neighboring squares, and dominoes don’t overlap.) (An “Aha” problem.)

Exercise 1.4 (John Conway). Consider an infinite chessboard: that means, a chessboard with infinitely many squares in all directions. We will draw one line which cuts the chessboard into infinite halves, the North and the South half. This could be viewed as the coordinate plane where the integer lattice (i.e. points \((a, b)\) for \(a\) and \(b\) integers) are the corners of squares, and the \(x\)-axis is the equator separating the board into the North and South parts. The North side has some army of pieces, all identical, with at most one on each square, and the South
side is empty. North’s goal is to take its pieces and invade as deep as possible into South’s territory. To do this, the only move allowed is to take two of North’s pieces which are adjacent (possibly diagonally adjacent) and have one jump (or step) over the other piece, thereby killing the second piece, ending up across on the other side. In other words, two adjacent pieces can be replaced with a third piece which makes a line of three adjacent pieces with the first two. Prove that North cannot invade too far into South’s territory: say, prove that North cannot advance beyond 100 steps south of the equator. (Note that North may have begun with an infinite number of pieces.) (Note: 8 is the maximum depth.)

**Exercise 1.5 (Paul Erdős).** a) Take \( n + 1 \) numbers between 1 and \( 2n \), inclusive. (assigned to Lajos Pósa at Pósa’s age of 12). Prove that two of them are relatively prime. (an “aha” problem)

b) Again take \( n + 1 \) numbers between 1 and \( 2n \), inclusive. Prove that one is a factor of another. (an “ahaaa” problem)

**Exercise 1.6.** Take a rectangular pool table, and a large collection of pennies, all round of the same radius (assume that the pool table fits at least one penny). Suppose two people take turns, each turn consisting of placing another penny on the pool table so that it does not overlap with any other penny and does not dangle off the side. The first player who cannot fit a penny on the board loses. Prove that the first player has a winning strategy.

**Divisor Game.** Fix an integer \( n > 1 \). Player 1 starts the game by picking a divisor of \( n \). Player 2 picks another divisor of \( n \) that is not a divisor of the divisor picked by Player 1. Play alternates between the two players with each player choosing a divisor of \( n \) that is not a divisor of any divisor already picked. Eventually, one player is forced to choose \( n \), and that player loses. A winning strategy for Player 1 was found in class for \( n = 30, 12 \). For primes \( p, q \), and \( r \), we generalized winning strategies for \( n = pq, p^k, p^kq \).

**Exercise 1.7.** Find a winning strategy for Player 1 when \( n = p^2q^2 \).

**Exercise 1.8.** Prove that Player 1 has a winning strategy for all \( n \geq 2 \). (*Hint:* Prove the existence of a winning strategy. No explicit strategy is known that works for all \( n \).)

On an \( n \times n \) chessboard, we can “infect” any initial configuration of cells. Then the infection spreads: a cell becomes infected if at least two adjacent cells are infected (diagonal neighbors do not count). How few cells can we initially infect so that the whole board becomes infected? Note that if we infect the \( n \) cells on a diagonal, the whole board becomes infected.

**Exercise 1.9.** Show that in order to get the entire board infected, we need at least \( n \) initially infected cells.

### 1.3 Ramsey Numbers

**Ramsey Game.** Player 1 has a blue pencil, and Player 2 has a red pencil. Given \( n \) points in the plane, no three on a line, Player 1 starts by connecting a pair of points with a blue
1.3. RAMSEY NUMBERS

line. Player 2 connects a different pair with a red line. The players continue taking turns by assigning their color to an uncolored pair of points. Player 1 loses if a blue triangle is formed, that is three of the \( n \) points are connected to each other with blue lines. Similarly, Player 2 loses if a red triangle is formed.

For the game \( n = 5 \), it is possible that no player loses: color the five sides of a regular pentagon blue, and color the five diagonals red. In this way, no triangles are made. We will see that this is not possible in the \( n = 6 \) game, but first we introduce some notation.

**Definition 1.10 (Erdős-Rado Arrow Notation).** The notation \( n \rightarrow (k, \ell) \) means that for any red and blue coloring of the \( \binom{n}{2} \) pairs of \( n \) points, there is either a subset of \( k \) points all of whose pairs are colored red or a subset of \( \ell \) points all of whose pairs are colored blue.

We saw above that \( 5 \not\rightarrow (3, 3) \) and we prove

**Theorem 1.11.**

\[ 6 \rightarrow (3, 3) \]

**Proof.** Select one point, call it 1. Since 1 is paired with five other points, at least three of these pairings are the same color, say blue. Denote these three other numbers by 2, 3, and 4. If a pair among \( \{2, 3, 4\} \) is also blue, say \( \{2, 4\} \), then \( \{1, 2, 4\} \) is a blue triangle. If no pair among \( \{2, 3, 4\} \) is blue, then all three pairs are red, so \( \{2, 3, 4\} \) is a red triangle. \[ \Box \]

Notice that the arrow notation is more general than the Ramsey game because the arrow notation does not put any restrictions on the proportion of blue or red pairs. In particular, we found a coloring in the case \( n = 6 \) so that there are 9 red pairs but no red triangles (the three houses, three utilities example).

**Exercise 1.12.** Show \( 10 \rightarrow (3, 4) \).
Chapter 2

2nd day, Tuesday 6/22/04 (Scribe: Charilaos Skiadas, Eric Patterson, and Travis Schedler)

2.1 Games continued

Exercise 2.1 (Dominoes). Prove: if we remove two opposite corners from the chessboard, the board cannot be covered by dominoes. (Each domino covers two neighboring cells of the chessboard.) Look for an “Ah-ha” proof: clear, convincing, no cases to distinguish.

Exercise 2.2 (Triominoes). Remove a corner from a 101 × 101 chessboard. Prove that the rest cannot be covered by triominoes. A triomino is like a domino except it consists of three squares in a row; each cell can cover one cell on a chessboard. Each triomino can either “stand” or “lie.” Look for an “Ah-ha” proof.

Exercise 2.3 (Knight’s trail). Consider a knight moving around on a 4 × 4 chessboard. We let the knight start at any cell of our choosing, and we wish to guide it through 15 moves so it never steps on a previously visited cell. So, after the 15 moves, the knight will have visited each cell. Prove that this is impossible. Find an “Ah-ha” proof.

Graph of knight moves on a 4 × 4 chessboard.

Hint: Show that from any trail of the knight’s moves, if you delete 4 cells, the trail splits into no more than 5 connected parts. (A set $S$ of cells is connected from the knight’s point of view if the knight can move from any cell of $S$ to any other cell of $S$ without leaving $S$.)
Exercise 2.4. A mouse finds a $3 \times 3 \times 3$ chunk of cheese, cut into 27 blocks (cubes), and wishes to eat one block per day, always moving from a block to an adjacent block (a block that touches the previous block along a face). Moreover, the mouse wants to leave the center cube last. Prove that this is impossible. Find two “Ah-ha!” proofs; one along the lines of the solution of knight’s trail problem, the other inspired by the solution of the dominoes problem.

Exercise 2.5. 96 kids wait for us to split an $8 \times 12$ chocolate bar along the grooves into 96 small rectangles. It is up to us in what order we do the splitting; we can start, for instance, by breaking the 7 long grooves and then split each of the 8 long $(1 \times 12)$ pieces; or we can start with the short grooves, or halve the bar each time, or any other way. The one thing we are not permitted to do is stack the pieces. At any one time, we have to pick up one piece and break it into two.

Each break takes us 1 second. Find the fastest method. (This is another “Ah-ha” problem.)

Exercise 2.6. $n$ teams play a straight-elimination tournament; there are no ties. How many matches do they need to play before the winner is declared? (An “Ah-ha” problem again.)

* * *

Study the handout for the basic definitions involving graphs, including complete graphs, (complete) bi(multi-)partite graphs, (induced) subgraphs, the degree of a vertex, complement of a graph and Hamilton cycles. In particular, recall that an isomorphism from the graph $G = (V; E)$ to the graph $H = (W; F)$ is a bijection $f : V \to W$ between the vertices that preserves the adjacency relation, i.e. for any two vertices $x, y \in V$, $x, y$ are adjacent in $G$ iff (if and only if) $f(x), f(y)$ are adjacent in $H$. $G$ is said to be isomorphic to $H$ if there exists an isomorphism from $G$ to $H$.

Exercise 2.7. Show that isomorphism of graphs is a transitive relation.

Exercise 2.8. Show that the number of functions $f : A \to B$ is $|B|^{|A|}$.

* * *

We defined decision tree and the related concepts of root, node, leaves, parent, and child. For example, flipping $n$ coins can be made into a decision tree with $2^n$ leaves. Similarly, we can make a decision tree from the moves in a chess game, but the tree is more complicated. For example, the leaves will not all be at the same level since different games end after different numbers of moves. Furthermore, the possible moves at each node are dependent on the preceding moves or the history. Although the decision tree for a chess game is far too large to store in a computer (it has more configurations than the number of atoms in the earth), such trees can on principle be analyzed in a finite time for a winning strategy.

First, assign a value of B, W, or D to the leaves of the tree if the outcome is a black win, a white win, or a draw, respectively. The nodes above the leaves can be assigned a value of B,
2.2. HAMILTON CYCLES

W, or D if there is an optimal strategy for black to win, white to win, or a draw. Suppose the node $x$ is a white move. Then the value of $x$ is W if $\exists$ W child of $x$, B if all children are B, and D if there is no W child and $\exists$ D child.

**Definition 2.9.** A **strategy function** for white is a function from the set of histories preceding a white move to the set of possible next moves. A **winning** strategy function for white makes moves to nodes with value W.

If a winning strategy exists for white, then the root must have value W, and there is a W child all of whose children are W. Each of these children must have a W child all of whose children are W, etc.

**Theorem 2.10.** If a finite game only has win or lose outcomes, then one of the players has a winning strategy.

This theorem can be used to solve the Divisor Game problem. (Hint: Proof by contradiction.)

**Exercise 2.11.** Exercise 6.1.5 p. 44 of the notes.

2.2 Hamilton cycles

**Definition 2.12.** A **Hamilton cycle** is a subgraph that is an $n$-cycle in a graph on $n$ vertices. A graph is **Hamiltonian** if it has a Hamilton cycle.

**Exercise 2.13.** The $k \times \ell$ grid is Hamiltonian if and only if $k\ell$ is even and $k, \ell \geq 2$.

**Method 1.** Make the vertices of the grid into the cells of a chess board. *Hint:* How many steps do you move to get back to your starting color? How many steps do you move to get all the cells?

**Method 2.** *Sam’s hint:* Use the lemma:

**Lemma 2.14.** If $G$ is Hamiltonian then removing $k$ vertices splits $G$ into at most $k$ connected components.

**Definition 2.15.** A graph $G$ is bipartite if its vertices can be colored red and blue such that adjacent vertices have different colors.

Observe: if $G$ is bipartite and Hamiltonian, then the two parts are equal. Exercise 2.13 is a consequence of this.

**Method 3.** *Alex’s hint:* What goes up, must come down. What moves left, must return to the right.
Exercise 2.16 (Dirac). Prove: if every vertex in a graph $G$ of $n$ vertices has degree $\geq n/2$, then $G$ is Hamiltonian (i.e. it contains a Hamiltonian cycle).

Exercise 2.17. Let $G$ be a graph with $n$ vertices and $m$ edges. If $m \geq \binom{n-1}{2} + 2$ then $G$ is Hamiltonian.

Definition 2.18. A graph $G$ is bipartite if the set of vertices $V$ partitions into two disjoint subsets, $V_1$ and $V_2$ (i.e. so that $V_1 \cup V_2 = V$ and $V_1 \cap V_2 = \emptyset$) such that no vertex of $V_1$ is adjacent to any other vertex of $V_1$, and no vertex of $V_2$ is adjacent to any other vertex of $V_2$. In other words, all edges of $G$ connect one vertex from $V_1$ and one vertex from $V_2$.

In other words, a graph is bipartite iff it is a subgraph of a complete bipartite graph (cf. p. 44 of the notes).

Exercise 2.19. Prove that a graph $G$ is bipartite iff (if and only if) $G$ has no odd cycle.


Chapter 3

3rd day, Wednesday 6/23/04 (Scribe: Justin Sinz and Travis Schedler)

3.1 Graph theory

Exercise 3.1. Let $R$ be an equivalence relation on a set $\Omega$. Prove that we can obtain from $R$ a (unique) partition of $\Omega$ (such that the partition induces the original equivalence relation in the natural way).

Exercise 3.2. Recall: Every connected graph contains a spanning tree. Prove this by induction on the number of edges (i.e., complete the proof given in class). Recall that if we remove an edge from a cycle in a connected graph, the graph remains connected, and that this decreases the (finite) number of cycles (cycles have no repeated vertices).

Exercise 3.3. (A)

(a) In every tree, there exists a vertex that is common to all the longest paths (i.e., paths of greatest length).

(b) In every tree, if the length of a longest path is odd then there is an edge common to all the longest paths.

(c)* Disprove (a) for connected graphs.

Exercise 3.4. (B)
(Helly-type Theorem) Let $R_1, \ldots, R_k$ be subtrees of a tree, and suppose that for all $i, j$ we have $V(R_i) \cap V(R_j) \neq \emptyset$. Then $V(R_1) \cap \ldots \cap V(R_k) \neq \emptyset$.

Exercise 3.5. (C)
Let $G$ be a connected graph. Let $P_1, P_2$ be longest paths in $G$. Prove that $V(P_1) \cap V(P_2) \neq \emptyset$.

Note: (C) and (B) imply (A).
Definition 3.6. A legal coloring of a graph $G$ is an assignment of a “color” to each vertex such that adjacent vertices have different “colors”.

The chromatic number $\chi(G)$ of $G$ is defined to be the minimum number of colors necessary to color $G$ legally.

Exercise 3.7. The graph $G$ is bipartite if and only if $\chi(G) \leq 2$.

Exercise 3.8. Construct a graph $G$ such that $K_3 \not\subseteq G$ and $\chi(G) = 4$.

Hint: 11 vertices, 5-fold symmetry.

3.2 Inequalities

For a set of $n$ positive real numbers, $x_1, \ldots, x_n$ we can define several means:

1. the arithmetic mean $A(x_1, \ldots, x_n) = \frac{x_1 + x_2 + \cdots + x_n}{n}$;
2. the quadratic mean $Q(x_1, \ldots, x_n) = \sqrt{\frac{x_1^2 + \cdots + x_n^2}{n}} = \sqrt{A(x_1^2, \ldots, x_n^2)}$;
3. the geometric mean $G(x_1, \ldots, x_n) = \sqrt[n]{x_1 \cdots x_n}$;
4. the harmonic mean $H(x_1, \ldots, x_n) = \frac{n}{\frac{1}{x_1} + \cdots + \frac{1}{x_n}} = \frac{1}{A(\frac{1}{x_1}, \ldots, \frac{1}{x_n})}$.

Exercise 3.9. Prove, for any $x_1, \ldots, x_n$, the inequalities

\[ Q(x_1, \ldots, x_n) \geq A(x_1, \ldots, x_n) \geq G(x_1, \ldots, x_n) \geq H(x_1, \ldots, x_n). \]

Hint: $Q \geq A$ is easy; $A \geq G$ is difficult; $G \geq H$ follows immediately from $A \geq G$.

Exercise 3.10 (JENSEN’S INEQUALITY). There is a generalization of all of the above inequalities: the discrete version of Jensen’s inequality. Let $I \subseteq \mathbb{R}$ be a finite or infinite interval and let $f : I \to \mathbb{R}$ be a function. If $f$ satisfies the inequality

\[ (\forall x, y \in I) f \left( \frac{x + y}{2} \right) \geq \frac{f(x) + f(y)}{2}, \]

then

\[ f \left( \frac{x_1 + \cdots + x_n}{n} \right) \geq \frac{f(x_1) + \cdots + f(x_n)}{n}, \]

i.e.

\[ f(A(x_1, \ldots, x_n)) \geq A(f(x_1), \ldots, f(x_n)). \]

Similarly, if:

\[ (\forall x, y \in I) f \left( \frac{x + y}{2} \right) \leq \frac{f(x) + f(y)}{2}, \]

then

\[ f(A(x_1, \ldots, x_n)) \leq A(f(x_1), \ldots, f(x_n)). \]
Note: A continuous function satisfying inequality (3.2) is **concave** and a continuous function satisfying inequality (5.5) is **convex**. However, not every function satisfying either (or even both) inequalities is continuous. (Prove!)

**Exercise 3.11.** Prove Jensen’s inequality. Hint: First show inductively that the inequality is true for \( n = 2^k \) (that is, by induction on \( k \)). Then, show that the inequality for \( 2^k \) implies the inequality for any \( n < 2^k \) as well, thus proving the inequality for all \( n \).

**Exercise 3.12.** Deduce Exercise 3.9 from Exercise 3.11.

### 3.3 Graphs without short cycles

We proved in class the

**Theorem 3.13 (Kővári-Sós-Turán).** (cf. Exercise 6.1.22 in the text) There exists \( c > 0 \) such that, if \( \mathcal{G} \) is a graph with \( m \) edges and \( n \) vertices, and \( \mathcal{G} \) has no 4-cycles, i.e. \( \mathcal{G} \not\subseteq C_4 \), then \( m \leq cn^{3/2} \). In other words, \( m = O(n^{3/2}) \).

In fact we showed that, as \( m \) tends to infinity, \( m \leq \frac{1}{2}n^{3/2} \).

The following exercise strengthens this result:

**Exercise 3.14.** Show that, in the above situation of \( \mathcal{G} \) with \( m \) edges, \( n \) vertices, s.t. \( \mathcal{G} \not\subseteq C_4 \), that

\[
m \leq \frac{n^{3/2} + n}{2} \quad (3.7)
\]

Hint: use the following inequality from the proof of the theorem:

\[
\frac{n-1}{2} \geq \left( \frac{2m}{n} \right) \quad (3.8)
\]

The following exercise shows the tightness of the bound:

**Exercise 3.15.** Show that there exists \( c' > 0 \) such that for infinitely many choices of \( n \), there exists a graph \( \mathcal{G} \not\subseteq C_4 \) with \( m \) edges and \( n \) vertices, such that \( m > c'n^{3/2} \).

**Exercise 3.16.** Show that \((\forall k)(\exists \mathcal{G} \not\subseteq K_3), (\chi(\mathcal{G}) = k)\). In words, there exist graphs which do not contain triangles which have arbitrarily high chromatic number.

**Exercise 3.17.** Show that \((\forall k, \ell)(\exists \mathcal{G} \text{ such that } \mathcal{G} \text{ has no odd cycles of length } \ell, \text{ and } \chi(\mathcal{G}) = k)\). In words, show that for any odd \( \ell \geq 1 \), there exist graphs with arbitrarily high chromatic number which do not contain any odd cycles of length \( 3, 5, \ldots, \ell \).

**Remark 3.18.** Exercise 3.17 remains valid if we drop the word ”odd”: 
Theorem 3.19 (Erdős 1960). \((\forall k, \ell)(\exists G\text{ such that } G\text{ has no cycles of length } \leq \ell, \text{ and } \chi(G) = k)\). In words, for any \(\ell \geq 1\), there exist graphs with arbitrarily high chromatic number which do not contain any cycles of length \(3, 4, 5, \ldots, \ell\).

Erdős proved this in 1960 using the probabilistic method but did not exhibit any specific examples of such graphs. No explicit construction was known of such graphs until 30 years later, around 1990.
4.1 Graph theory continued

**Exercise 4.1.** If $G$ is connected with $n \geq 2$ vertices, then there exists a vertex $v$ such that $G \setminus v$ is connected (the latter means the subgraph we get by throwing out the vertex $v$ and all edges incident with $v$). Note that we cannot pick just any vertex, as evidenced by the case of throwing out the center of a star (which leaves a highly disconnected graph).

**Definition 4.2.** A graph $G$ is $k$-edge-connected between vertices $x$ and $y$ if there are $k$ edge-disjoint paths between $x$ and $y$.

**Definition 4.3.** A graph $G$ is $k$-vertex-connected between vertices $x$ and $y$ if there are $k$ vertex-disjoint paths between $x$ and $y$.

Suppose we found 57-edge-disjoint paths between $x$ and $y$, and we think 57 is optimal. How can we demonstrate that 58 such paths cannot be found?

**Ben's Conjecture:** If there are no more than 57 edge-disjoint paths from $x$ to $y$, then there is a partition $V = A \cup B$ such that $x \in A$, $y \in B$, and the maximum number of edges between $A$ and $B$ is less than or equal to 57.

**Definition 4.4.** An $(x, y)$-cut of a graph $G = (V, E)$ is a partition $V = A \cup B$ such that $x \in A$ and $y \in B$.

**Theorem 4.5 (Menger's Theorem).** The maximum number of edge-disjoint paths between $x$ and $y$ equals the minimum number of edges between the two parts of an $(x, y)$-cut.

4.2 Puzzles

**Exercise 4.6.** If we have a $10^{10} \times 10^{10}$-size square that we want to cover with $1 \times 1$-squares, evidently we can fit $10^{20}$ squares. Show that if we have a $(10^{10} + 0.1) \times (10^{10} + 0.1)$-square then we can fit $10^{20} + 1$ squares.
Exercise 4.7. Suppose we have a big rectangle which is $a \times b$ in size, and we tile it with
various size rectangles $a_n \times b_n$ (situated parallel to the sides of the original rectangle), where
for each $n$, at least one of $a_n$ or $b_n$ is an integer. Show that either $a$ and $b$ is an integer.

4.3 Set game and Szemerédi’s theorem

Set Game. The Set game is played with a deck of cards. Each card has four properties with
three possible values: a shape (oval,diamond, squiggle), a color (red, green, purple), a shading
(blank, solid, shaded), and a number (1,2,3). Each card is distinct, so there are $3^4 = 81$ cards.
Twelve cards are placed face up on a table. The object is to find a “set” among the twelve
cards. A “set” is formed by three cards so that each property is either the same on all cards
or different on all cards. The first player to find a “set” gets that “set,” after which three new
cards from the deck are placed face up to replace them. The player who finds the most “sets”
wins.

We can generalize the Set game so that there are $n$ properties on each card, and each
property has three possible values. In this $n$-dimensional Set game we have $3^n$ cards. We are
interested in how many cards we can draw from the deck without finding a “set.” Let $\alpha(n)$
denote the maximum number.

Let $\mathbb{Z}_3$ be the set $\{0, 1, 2\}$ with addition mod 3, so $2 + 2 = 1$ for example. Let $\mathbb{Z}_3^n$ be the set
$\{(x_1, \ldots, x_n) : x_i \in \mathbb{Z}_3\}$. For $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $\mathbb{Z}_3^n$, we define addition by
$x + y = (x_1 + y_1, \ldots, x_n + y_n)$. In the Set game ($n = 4$), the cards correspond to the quadruples
in $\mathbb{Z}_3^4$; for instance [oval, green, blank, 3] ↔ [1, 1, 0, 2].

Observation 4.8. Three distinct elements $x, y, z \in \mathbb{Z}_3^4$ form a set if and only if $x + y + z = 0$.

Similarly, we can work in $\mathbb{Z}_3^n$ for the $n$-dimensional Set game. So an alternative definition
for $\alpha(n)$ is the maximum number of elements of $\mathbb{Z}_3^n$ without a solution to $x + y + z = 0$ where
$x, y, z$ are distinct.

Definition 4.9. A function $f : \mathbb{N} \to \mathbb{R}$ is supermultiplicative if $f(n + m) \geq f(n) \cdot f(m)$.

Exercise 4.10. Prove: if $f$ is supermultiplicative, then $L = \lim_{n \to \infty} \sqrt[n]{f(n)}$ exists, and $L \geq \sqrt{f(n)}$ for all $n$.

Exercise 4.11. Prove $\alpha(n)$ is supermultiplicative.

Observe that $2^n \leq \alpha(n) \leq 3^n$, so $2 \leq L \leq 3$ by the previous two exercises, where $L = \lim_{n \to \infty} \sqrt[n]{\alpha(n)}$.

Exercise 4.12. Prove: $\alpha(4) \geq 20$. Infer: $L \geq \sqrt[20]{20} \approx 2.115$.

Exercise 4.13. Prove: $\alpha(2) = 4$. Infer $\alpha(4) \leq 36$. 

Exercise 4.14. Prove: $\alpha(3) \geq 10$. Infer: $L \geq \sqrt[10]{10} \approx 2.154$.
4.3. SET GAME AND SZEMERÉDI’S THEOREM

Exercise 4.14. Prove: $\alpha(3) = 9$. Infer $\alpha(4) \leq 27$.

Exercise 4.15. Prove: $\alpha(4) \leq 24$.


Conjecture 4.17 (Open Problem). $\lim_{n \to \infty} \sqrt[3]{\alpha(n)} = 3$.

Exercise 4.18. Color the integers red and blue. Is there a set of 3 red or 3 blue integers that form an arithmetic progression? Find a coloring of the positive integers such that there is no arithmetic sequence of 3 red or 3 blue integers in the first $n$ numbers. What is the largest $n$ for which you can find an example?

Theorem 4.19 (Van der Waerden, 1928). If we color the positive integers with $k$ colors, then for all $\ell$ there is an $\ell$-term arithmetic progression in one color.

Theorem 4.20 (Szemerédi, 1975). For all $\epsilon > 0$ and $\ell$, there is an $N$ such that if $A \subseteq \{1, \ldots, N\}$ and $|A| \geq \epsilon N$, then $A$ contains an $\ell$-term arithmetic progression.

Exercise 4.21. Prove that Szemerédi’s Theorem implies van der Waerden’s Theorem.

History:

- Erdős and Turán conjectured this result in 1936.
- In 1952, K. F. Roth found an analytic proof of the result for $\ell = 3$. This result, in addition to another major result, were part of his Fields Medal citation in 1958.
- In the late 1960s, E. Szemerédi found a combinatorial proof for $\ell = 3$.
- Later, E. Szemerédi found a combinatorial proof for $\ell = 4$.
- In 1975, E. Szemerédi found a combinatorial proof of the whole conjecture.
- Subsequently, Furstenberg found an analytic proof of Szemerédi’s Theorem. His methods led to further generalizations. For example:

Theorem 4.22. $\alpha(n) = o(3^n)$.

On April 8, 2004, a major breakthrough was announced in a very old problem that has fascinated mathematicians for ages:

Theorem 4.23 (Green-Tao, 2004). For all $\ell$, there is an $\ell$-term arithmetic progression among prime numbers.

Ben Green and Terence Tao announced this result in their paper “The Primes Contain Arbitrarily Long Arithmetic Progressions.” The proof uses Szemerédi’s Theorem and the ideas of Furstenberg’s proof. It is posted on arXiv.
Chapter 5
5th day, Friday 6/25/04 (Scribe: Justin Sinz and Charilaos Skiadas)

5.1 Density versions of coloring theorems

**Notation:** 

\[ [k] = \{1, 2, \ldots, k\}; \quad [k]^n = \{(x_1, \ldots, x_n) : x_i \in [k]\}. \]

**Definition 5.1.** A subset \( L \subseteq [k]^n \) with \( k \) elements is called a **combinatorial line** if we can arrange the elements in an order so that in every coordinate either all the entries are the same or they form the sequence \( 1, 2, \ldots, k \), in that order.

**Note:** A combinatorial line is a “SET” in the game “SET.” The converse is not true. (why?)

**Theorem 5.2 (Hales-Jewett, 1963).** For all \( k, \ell \) there exists an \( n_0 \) such that for all \( n \geq n_0 \) and any \( k \)-coloring of \([\ell]^n\), there exists a combinatorial line in one color.

**Note:** This theorem implies Van der Waerden’s theorem:

**Theorem 5.3 (Van der Waerden, 1927).** For any \( k, \ell \) there exists \( N_0 \) such that for all \( N \geq N_0 \), if we color \([N]\) with \( k \) colors, then there is an \( \ell \)-term arithmetic progression in one color.

The idea is to let \( N_0 = \ell^{n_0} \), and to associate with every number in \([N]\) the digits in its expansion in the base-\( \ell \) number system. Then any combinatorial line would correspond to an arithmetic progression. (But not conversely. Why?)

Recall Szemerédi’s Theorem, which is the “density version” of van der Waerden’s Theorem:

**Theorem 5.4 (Szemerédi, 1975).** For all \( \varepsilon > 0 \) and \( \ell \) there exists \( N_0 \) such that for all \( N \geq N_0 \), if \( A \subseteq \{1, \ldots, N\} \) and \( |A| \geq \varepsilon N \) then \( A \) contains an \( \ell \)-term arithmetic progression.

The density version of the Hales-Jewett Theorem is:
Theorem 5.5 (Furstenberg-Katznelson). For all $\epsilon > 0$ and all $\ell$ there exists $n_0$ such that for all $n \geq n_0$, if $A \subseteq [\ell]^n$ and $|A| \geq \epsilon \ell^n$, then $A$ contains a combinatorial line.

Exercise 5.6. Use the Furstenberg-Katznelson Theorem to prove that $\alpha(n) = o(3^n)$ in the notation of the fourth set of notes to this course. (Recall that we used $\alpha(n)$ to denote the maximum number of cards in the $n$-dimensional “SET” game without a “SET.”)

5.2 Hypergraphs

Definition 5.7. (a) A hypergraph is a pair $\mathcal{H} = (V, E)$, where $V$ is a set of vertices and $E$ is a set of edges, an edge simply being a subset of $V$.

(b) A hypergraph is called $t$-uniform, if every edge has $t$ vertices.

Note:

(a) A graph is a 2-uniform hypergraph.

(b) If a hypergraph on $n$ vertices is $t$-uniform, then $|E| \leq \binom{n}{t}$.

(c) Given a set $V$ of vertices, $|V| = n$, the number of $t$-uniform hypergraphs on $V$ is $\binom{n}{t}$.

   The number of $t$-uniform hypergraphs with $m$-edges is $\binom{n}{m}$.

(d) The “SET” game is based on a 3-uniform hypergraph with 81 vertices, where the edges are the “SETs.”

Exercise 5.8. Count the total number of “SETs” in the $n$-dimensional “SET” game. (Self-check: for $n = 1$, your formula should give 1.) What number do you get for $n = 4$ (the actual “SET” game)?

Extremal set theory is concerned with problems of the following type: determine (or estimate) the maximum number of edges of a hypergraph satisfying certain conditions. We have seen some examples of this already in graph theory, for instance the Mantel-Turán Theorem and the Kővári–Sós–Turán Theorem. There are many more that deal with hypergraphs. Ruzsa and Szemerédi, in examining the combinatorial roots of Roth’s Theorem on 3-term arithmetic progressions, were lead to the following natural problem in extremal set theory:

Question: What is the maximum number of edges of a 3-uniform hypergraph satisfying the following two conditions:

(C1) No two edges share more than one point.
5.3. **STEINER TRIPLE SYSTEMS**

(C2) There are no triangles, i.e., there is no set of three edges that intersect pairwise but have an empty intersection.

Note that the maximum possible number of edges in a 3-uniform hypergraph is \( \binom{n}{3} \sim \frac{n^3}{6} \), the total number of triples of vertices. Condition (C1) already forces a lower (quadratic) rate of growth:

**Exercise 5.9.** Prove: if a 3-uniform hypergraph on \( n \) vertices satisfies condition (C1) then it has at most \( n(n-1)/6 \) edges.

*Hint.* For every pair of vertices there is at most one edge containing them; and every edge contains 3 pairs of vertices.

When we add condition C2, the rate of growth of the number of edges drops even further, below quadratic:

**Lemma 5.10 (Ruzsa-Szemerédi).** The maximum number \( m(n) \) of edges in a 3-uniform hypergraph satisfying conditions (C1) and (C2) is \( m(n) = o(n^2) \), i.e., \( \lim_{n \to \infty} m(n)/n^2 = 0 \).

We will see next time that this implies Roth’s theorem (i.e., Szemerédi’s theorem for \( \ell = 3 \)):

**Roth’s Theorem.** Given \( \epsilon > 0 \) there exists \( N_0 \) such that for all \( N \geq N_0 \), if \( |A| \subseteq [N] \) and \( |A| \geq \epsilon N \), then \( A \) contains a 3-term arithmetic progression.

**Exercise 5.11.** Assuming that we have proved Roth’s theorem for sets \( A \) consisting only of odd numbers, and prove Roth’s theorem for all sets.

5.3 **Steiner Triple Systems**

**Definition 5.12.** A Steiner Triple System (STS) is a hypergraph \( \mathcal{H} \) (in which the edges are called lines) such that

- every line consists of exactly 3 points (=vertices) (i.e., \( H \) is 3-uniform);
- every pair of (distinct) points is contained in a unique line.

In other words, a STS satisfies condition (C1) above and is maximal with that property: The number of edges is exactly equal to the upper bound obtained above: \( |E| = n(n-1)/6 \), where \( n \) is the number of vertices. In particular we must have that \( n(n-1) \) is divisible by 6.

**Exercise 5.13.** With \( n \) as above, prove that \( n \) is odd.

**Exercise 5.14.** Deduce from the preceding two exercises that the number of points of a STS is congruent to 1 or 3 mod 6.
This necessary condition on \( n \) also turns out to be sufficient.

We construct a STS \( \mathcal{F} \) on 7 vertices (with 7 edges) as follows: Take the vertices of a(n equilateral) triangle along with the bisectors of the edges and the (in)center of the triangle. Let the edges of the STS consist of the following:

(a) the 3 vertices on each edge of the triangle (for a total of 3 edges).

(b) the 3 vertices on each angle bisector.

(c) the 3 vertices on the (natural) inscribed circle.

This configuration is called the “Fano Plane” see Figure 5.1

\[ \begin{array}{cccccccc}
0 & 1 & 2 & 3 & 4 & 5 & 6 \\
a & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\
b & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\
c & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\
d & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\
e & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\
f & 0 & 1 & 0 & 0 & 0 & 1 & 1 \\
g & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\
\end{array} \]

Figure 5.1: The Fano Plane

The incidence matrix for the Fano plane is as follows (the rows correspond to lines, columns to points, “1” indicates incidence):
Exercise 5.15. Prove that the Fano plane has 168 automorphisms (line-preserving permutations).

The group of automorphisms of the Fano plane is a very important group: it is the second smallest finite simple group. “Simple groups” to “groups” are like “atoms” to “chemical compounds,” so according to this analogy, the automorphism group of the Fano plane is the “Helium of group theory.”

Exercise 5.16. Construct STSs for \( n = 13 \) and for \( n = 3^k \) and \( n = 2^k - 1 \) for all \( k \geq 1 \).

Exercise 5.17. Given STSs \( \mathcal{H}_1 \) and \( \mathcal{H}_2 \) with \( n_1 \) and \( n_2 \) points, respectively, construct a STS with \( n = n_1 n_2 \) points.

Exercise 5.18. Given a STS with \( k \) points, construct (explicitly) STSs on \( 3k, 3k - 2, \) and \( 3k - 6 \) points.

Exercise 5.19. From Exercises 5.16 and 5.18 above, deduce that if \( n \) is congruent to 1 or 3 mod 6 then there exists a STS on \( n \) points.

5.4 Projective planes

Definition 5.20. An incidence geometry is a set \( P \) of “points,” a set \( L \) of “lines,” and an incidence relation \( I \subseteq P \times L \).

Notation 5.21. If \( (p, \ell) \in I \) then we say that \( p \) is incident with \( \ell \) and we write \( p \leftrightarrow \ell \). If \( (p, \ell) \notin I \), we write \( p \not\leftrightarrow \ell \).

Definition 5.22. The dual of the incidence geometry \( (P, L, I) \) is the incidence geometry \( (L, P, I^{-1}) \) (we switch the roles of points and lines; the same pairs remain incident).

Definition 5.23. A projective plane is an incidence geometry satisfying the following three axioms:

Axiom 1. (\( \forall p_1 \neq p_2 \in P \)(\( \exists! \ell \in L \))(\( p_1 \leftrightarrow \ell \) and \( p_2 \leftrightarrow \ell \)).

Axiom 2. (\( \forall \ell_1 \neq \ell_2 \in L \)(\( \exists! p \in P \))(\( p \leftrightarrow \ell_1 \) and \( p \leftrightarrow \ell_2 \)).

Axiom 3. (Non-degeneracy) \( \exists p_1, p_2, p_3, p_4 \in P \) such that no three are on the same line.

Exercise 5.24. Prove that the dual of a projective plane is a projective plane. (Note the dual of Axiom 1 is Axiom 2 and vice versa. State the dual of Axiom 3 and prove that it follows from Axioms 1–3.)
Unless expressly stated otherwise, all projective planes considered here will be finite. For \( p \in P \), let \( \text{deg}(p) \), the \textit{degree of} \( p \), be the number of lines incident with \( p \). For \( \ell \in L \), let \( \text{rk}(\ell) \), the \textit{rank of} \( \ell \), be the number of points incident with \( \ell \).

The following observation is an immediate consequence of Axioms 1 and 2 and does not require Axiom 3.

**Exercise 5.25.** If \( p \dashv \ell \), then \( \text{deg}(p) = \text{rk}(\ell) \).

**Exercise 5.26.** In a projective plane, \( \forall p_1, p_2 \in P, \exists \ell \in L, p_1 \dashv \ell \) and \( p_2 \dashv \ell \).

The proof of this requires some care.

The following is immediate from the preceding two exercises.

**Exercise 5.27.** In a projective plane, all points have the same degree.

Use the fact that the dual of a projective plane is a projective plane (Ex: 5.24) to infer:

**Exercise 5.28.** In a projective plane, all lines have the same rank.

**Exercise 5.29.** In a projective plane, the degree of every point and the rank of each line is the same.

In other words, projective planes are \textit{regular} and uniform, and their degree and rank are equal.

For reasons of tradition, we denote this common value by \( n + 1 \). Every point of the plane is thus incident with \( n + 1 \) lines and every line is incident with \( n + 1 \) points. The number \( n \) is called the \textit{order} of the projective plane.

**Proposition 5.30.** \( |P| = |L| = n^2 + n + 1 \).

The smallest projective plane is the \textit{Fano plane}, which has order \( n = 2 \) (see Figure 5.1).

### 5.5 Galois Planes

A class of projective planes called \textit{Galois planes} is constructed as follows. Let \( F \) be a finite field of order \( q \). Let \( F^3 \) be the 3-dimensional space over \( F \). We define the \textit{inner product} over \( F^3 \) in the usual way: for \( u = (\alpha_1, \alpha_2, \alpha_3) \) and \( v = (\beta_1, \beta_2, \beta_3) \) we set \( u \cdot v = \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3 \). We say that \( u \) and \( v \) are \textit{perpendicular} if \( u \cdot v = 0 \).

Let us say that two nonzero vectors \( u, v \in F^3 \) are equivalent if \( u = \lambda v \) for some \( \lambda \in F \).

Let \( S \) be the set of equivalence classes on \( F^3 - 0 \). Note that each equivalence class has \( q - 1 \) elements and therefore the number of equivalence classes is \((q^3 - 1)/(q - 1) = q^2 + q + 1\).
Set $P = L = S$ and let us say that $p \in P$ and $\ell \in L$ are incident if $u \cdot v = 0$ where $u \in p$ ($u$ is a vector in the equivalence class $p$) and $v \in \ell$. The coordinates of $u$ are called homogeneous coordinates of $p$ (they are not unique—every point has $q - 1$ triples of homogeneous coordinates); similarly, the coordinates of $v$ are called homogeneous coordinates of $\ell$.

**Exercise 5.31.** Prove that this definition gives a projective plane of order $q$. It is called a Galois plane after Évariste Galois (1811–1832), the discoverer of finite fields and of modern algebra.

**Exercise 5.32.** Prove that the Fano plane is a Galois plane (necessarily over the field of order 2) by assigning homogeneous coordinates to the points and lines of the Fano plane.

A set of points is **collinear** if there is a line with which all of them are incident.

We say that four points are in **general position** if no three of them are collinear.

A **collineation** is a transformation of the projective plane consisting of a permutation of the points and a permutation of the lines which preserves incidence.

**Theorem 5.33 (Fundamental Theorem of Projective Geometry.).** If $a_1, \ldots, a_4$ and $b_1, \ldots, b_4$ are two quadruples of points in general position in a Galois plane then there exists a collineation $\varphi$ such that $(\forall i)(\varphi(a_i) = b_i)$.

**Exercise 5.34.** (Prove!) Use Theorem 5.33 to show that the Fano plane has 168 collineations.

**Exercise 5.35.** Consider the projective plane $\Pi = PG(2, F)$ over the field $F$. for $i = 1, 2, 3$, let $p_i = (a_i, b_i, c_i)$ be three points in $\Pi$, given by homogeneous coordinates $(a_i, b_i, c_i \in F$, not all zero). Prove: the three points are collinear (lie on a line) if and only if the $3 \times 3$ determinant $|a_i b_i c_i|$ is zero.

**Exercise 5.36.** A projective plane $\Pi_1 = (P_1, L_1, I_1)$ is a **subplane** of the projective plane $\Pi_2 = (P_2, L_2, I_2)$ if $P_1 \subset P_2$, $L_1 \subset L_2$, and the incidence relation $I_1$ is the restriction of $I_2$ to $P_1 \times L_1$. Prove: if $\Pi_1$ is a proper subplane of $\Pi_2$ then $n_1 \leq \sqrt{n_2}$ (where $n_i$ is the order of $\Pi_i$).

**Exercise 5.37.** Let $P(n)$ be the number of projective planes of order $n$. Prove: $P(n) < (ne)^{(n+1)^3}$. Hint, first prove that

$$P(n) \leq \binom{n^2+n+1}{n^2+n+1}.$$  

**Exercise 5.38.** Let us consider the Galois plane $PG(2, 5)$ (over the field of 5 elements).

1. How many points does this plane have, and what is the number of points per line?

2. Points are given by “homogeneous coordinates.” Determine whether or not the points given by $a = [1, 4, 0], b = [3, 2, 2]$, and $c = [4, 1, 2]$ are collinear (belong to the same line). (Coordinates are mod 5.) Prove your answer.
* * *

Puzzle 5.39. Find two infinite sets of nonnegative integers, $A, B$, such that every non-negative integer can be written in a unique way as $a + b$, with $a \in A$ and $b \in B$.

**Note:**

(a) We necessarily have that $A \cap B = \{0\}$.

(b) If $A$ is allowed to be finite, then we could take, for our favorite number $d > 0$, $A = \{0, \ldots, d - 1\}$ and $B = d\mathbb{Z}$. The hard part is to make both $A$ and $B$ infinite.

5.6 Internet and library resources

Regarding arithmetic progressions of prime number, please check Ivars Peterson’s “Math Trek” article, “Progressive Primes,” at

http://www.maa.org/mathland/mathtrek_04_26_04.html

and the new article by Ben Green and Terence Tao: “The primes contain arbitrarily long arithmetic progressions.” Peterson’s article includes a link to the Green – Tao paper.

Regarding the card game “SET,” consult the article “The card game SET” by B. L. Davis and Diane Maclagan at

http://math.stanford.edu/~maclagan/papers/set.pdf and some of the references mentioned in that paper, especially

Chapter 6

6th day, Monday 6/28/04 (Scribe: Eric Patterson and Travis Schedler)

6.1 North-South Game

**Exercise 6.1.** Assign positive real numbers to the cells in the grid so that

(a) the sum of the North squares is finite,

(b) a move never increases the sum of occupied squares, and

(c) there is a cell (in the South) with a number greater than the sum of all the cells in the North.

(Conclusion: that square cannot be reached.)

6.2 Number of unit distances in the plane \(< n^{3/2}\)

**Exercise 6.2.** First prove geometrically that a unit distance graph does not contain a \(K_{2,3}\). Then prove the Kővári-Sós-Turán theorem for \(K_{2,3}\).

**Exercise 6.3.** In \(\mathbb{R}^4\), find \(n\) points with \([n^{3/4}]\) unit distances between them (for all \(n\)).

**Exercise 6.4.** Find a method of shifting a grid and doubling points to get \(n\) points with \(cn\log n\) unit distances.

**Exercise 6.5.** What numbers can be written as \(a^2 + b^2\) in many ways? In particular, what is the smallest number \(n\) that you can write as \(a^2 + b^2\) in at least 200 ways?

**Exercise 6.6.** Use the previous exercise to show that there is a \(d_n\) such that the distance \(d_n\) occurs \(n^{1+\frac{c}{\log n}}\) times in the \(\sqrt{n} \times \sqrt{n}\) grid. This is the best known result for this problem.

**Exercise 6.7.** Prove \((\log x)^{100} = o(x)\). Show \(n^{\frac{c}{\log n}} > \log n\). (Thus, the best known result is in fact better than \(cn\log n\).)
6.3 Max distance occurs \( \leq n \) times

**Exercise 6.8.** If each vertex in a graph has degree \( \leq 2 \), then the number of edges is \( \leq n \).

Let \( S \) be a set of points in the plane. Let \( G(S) = (S, E) \) be the “max distance graph” defined by joining two points if they are at maximum distance.

**Exercise 6.9.** Prove: if \( \text{deg}(x) \geq 3 \), then \( x \) has a neighbor of degree one (a “dangling neighbor”).

**Exercise 6.10.** Prove: if \( G \) has the property in the previous exercise, hereditarily to all induced subgraphs, then \( m \leq n \).

**Exercise 6.11.** Use the preceding exercises to conclude that the max distance occurs at most \( n \) times.

6.4 Fitting \( 10^{20} + 1 \) squares on a \((10^{10} + 0.1) \times (10^{10} + 0.1)\) square

**Exercise 6.12.** Let the upper \((10^{10} - 10^6) \times (10^{10})\) rectangle be region \( A \), and let the rest of the square be region \( B \). So \( B \) is a \((10^6 + 0.1) \times (10^{10} + 0.1)\) rectangle. Tile region \( A \) by aligned squares \((10^{20} - 10^{16})\) of them, no surprise here). \( B \) is a long stripe. Squeeze in the square here by tilting the columns. You may find a number better than \( 10^6 \) for the width of this stripe.

**Exercise 6.13.** In an \((n + .01) \times (n + .01)\) square, we can put in \( n^2 + cn^\alpha \) squares for some constants \( c \) and \( \alpha \). Compute as large an \( \alpha \) as you can get.

6.5 Tiling \( a \times b \) rectangle by rectangles with at least one integer dimension

Let \( R \) be an \( a \times b \) rectangle, where \( a \) and \( b \) are positive real numbers. Suppose that we can tile \( R \) by “aligned” rectangles (the sides of the rectangles are parallel to the sides of \( R \)) such that each realigned rectangle has at least one side with integer length. Show that at least one side of \( R \) has integer length. Assume one corner of \( R \) is at the origin, and the sides of \( R \) are aligned with the axes.

There is a measure theory solution and a graph theory solution. This hint applies to the measure theory solution. Assign weights (positive or negative) to each rectangle in \( \mathbb{R}^2 \) so that if a rectangle has an integer side, the weight is zero. The weights should be additive (if a rectangle is cut up into rectangles, the weights should add up).

This can be done continuously or discretely. For a rectangle \( S \), we can define its weight as

\[
\int_{S} \sin(2\pi x) \sin(2\pi y) dx \ dy.
\]
To define the weight discretely, tile $\mathbb{R}^2$ on the half integers with a chess board coloring. If a region is white, its weight is its area. If a region is black, its weight is the negative of its area.

**Exercise 6.14.** Prove that the weights have the following properties:

(a) if a side of an aligned rectangle is an integer, then the measure is zero, and

(b) if the measure of an aligned rectangle is zero and one corner is the origin, then a side is an integer.

**Exercise 6.15.** Use the previous exercise to prove that at least one side of $R$ has integer length.
Puzzle 7.1. (a) A 3-way lamp (red, yellow, green) is connected to $n$ 3-way switches. Each configuration of the switches corresponds to a state of the lamp. In other words we have a function $f : \{3\}^n \rightarrow \{3\}$. It is known that if we change all the switches, then the lamp will switch also. Show that all but one of the switches are dummies. In other words, the function $f$ depends on only one of the variables: one switch determines the state of the lamp.

(b) Show that this is not necessarily true if we have infinitely many switches. In fact, we can arrange it so that the lamp does not change whenever we change finitely many switches. (But it does change if we change all of them.)

Puzzle 7.2. A hundred prisoners are on their way to a prison island. They know a lot about the workings of the prison. First of all, the moment they arrive they will be taken to individual cells, never to see or talk to each other again. Also, there is a special room that they are taken to, where they are being interrogated by the warden. They are taken there one at a time, but in unknown order. They have no way of knowing who else has entered the room before them. They know that each one of them will be called to the room infinitely many times, but they don't know in what order. In the room there is a two-way switch (up, down), which they get to see and are allowed to change if they want to. Nobody other than the prisoners will ever touch the switch. There is also a red button, which has an amazing property: If one of the prisoners presses it and by that time all the prisoners have been in this room at least once, then the game is over and they all go free. If one of them hasn’t, then they all spend the rest of their lives in prison. (By the way, the prisoners are immortal, so this will be a really long time.) The prisoners have a lot of time before they get to the prison to make their plans. Find a way for them to get out.

There are two versions to this problem. In the easier version, the prisoners know the initial position of the switch. In the other they don’t. Figure out a strategy for each problem.
7.1 Fishermen’s clubs

In a small fishing town by a lake, the \( n \) fishermen of the town are in the habit of forming as many clubs as they can. To control the situation a bit the state has formed a rule that any two clubs must share exactly one member. The question is raised, how many clubs can there be. It is easy to come up with \( n \) clubs. For instance we could have one really passionate person, who has formed \( n - 1 \) clubs, one with each other fisherman. (So all these clubs so far have exactly 2 members each.) Then all the other fishermen, having gotten tired of him, decide to form a club, for a total of \( n \) clubs. Or maybe he was faster than them and formed a club by himself. Notice though that not both of those last clubs can be formed, because they don’t share a member. Another way to form \( n \) clubs is if \( n = k^2 + k + 1 \) and there is a projective plane of order \( k \); the points correspond to the fishermen and the lines to the clubs. The natural question then is whether we could have more than \( n \) clubs. The surprising answer is given by the following theorem:

**Theorem 7.3 (Erdős-deBruijn, 1949).** Given a set with \( n \) elements, there cannot be more than \( n \) subsets with the property that any two of them share exactly one member.

A reasonable next step would be to ask: What if we allow them to have exactly \( k \) members in common? The answer remains the same:

**Theorem 7.4 (Fisher-R.C.Bose-Majumdar).** Given \( k \geq 1 \) and given a set with \( n \) elements, there cannot be more than \( n \) subsets with the property that any two of them share exactly \( k \) members.

The proof is not hard but is quite ingenious: it invokes an unexpected tool, linear algebra. The method, introduced by R.C. Bose, has since produced volumes of startling results.

**Definition 7.5.** (a) The **characteristic vector** (incidence vector) of a subset \( A \subseteq [n] \) is the vector \((x_1, \ldots, x_n)\) in \( \mathbb{R}^n \), where \( x_i = 1 \) if \( i \in A \) and \( x_i = 0 \) if \( i \notin A \).

(b) If \( A_1, \ldots, A_k \) are subsets of \([n]\), the **incidence matrix** of this set-system (hypergraph) is the \( k \times n \) matrix whose \( i \)-th row is the characteristic vector of \( A_i \).

**Exercise 7.6 (R.C. Bose, 1949).** Let \( k \geq 1 \) and let \( A_1, \ldots, A_k \subseteq [n] \). Show that if \((\forall i \neq j)(|A_i \cap A_j| = k)\), the characteristic vectors of the \( A_i \) are linearly independent.

Since there cannot be more than \( n \) linearly independent vectors in \( \mathbb{R}^n \) ("Fundamental Fact of Linear Algebra"), this proves Theorem 7.4.

(Note: Bose proved this for uniform hypergraphs; Majumdar extended the result to the nonuniform case.)
7.2 Ramsey Theory

Notation: If \( A \) is a set then \( \binom{A}{k} \) denotes the set of all \( k \)-subsets (subsets of size \( k \)) of \( A \). So if \( |A| = n \) then \( \left| \binom{A}{k} \right| = \binom{n}{k} \). Recall that \( \binom{n}{1} = \ldots = \binom{n}{n-1} = 1 \).

**Definition 7.7 (Erdős-Rado arrow notation).** \( \binom{[n]}{r} \rightarrow (s_1, \ldots, s_k) \) if for any partition of \( \binom{[n]}{r} = A_1 \cup \ldots \cup A_k \) there exists \( i, 1 \leq i \leq k \) and \( H \subseteq [n] \) such that \( |H| \geq s_i \) and \( \binom{H}{r} \subseteq A_i \). Absence of a superscript on \( (n) \) indicates that \( r = 2 \). The subscript is redundant and often omitted.

**Theorem 7.8 (Ramsey).** For all \( r, k, s_1, \ldots, s_k \) there exists \( n \) such that \( \binom{[n]}{r} \rightarrow (s_1, \ldots, s_k) \).

**Notation:** \( R_r(s_1, \ldots, s_k) \) denotes the smallest such \( n \) and is called a **Ramsey number**. If \( r = 2 \) then \( r \) is omitted.

**Exercise 7.9.** Prove that \( 17 \rightarrow (3,3,3) \).

We have \( 6 \rightarrow (3,3) \) and \( 17 \rightarrow (3,3,3) \). In general, \( [n!e] \rightarrow (3, \ldots, 3) \) (\( n \) parts, i.e., \( \binom{[n!e]}{2} \rightarrow (3, \ldots, 3) \)).

Here are some known Ramsey numbers.

- \( 6 \rightarrow (3,3) \), \( 5 \rightarrow (3,3) \), so \( R(3,3) = 6 \).
- \( 10 \rightarrow (3,4) \), \( 9 \rightarrow (3,4) \), so \( R(3,4) = 10 \).
- \( 17 \rightarrow (3,3,3) \), \( 16 \rightarrow (3,3,3) \) (\textit{not} 17), so \( R(3,3,3) = 17 \).

The \( R(5,5) \) is known to be at least 43 and this is conjectured to be the correct value. However, in order to prove that it is 43 (that is, for every graph \( G \) on 43 vertices either \( G \) or its complement contains a \( K_5 \)), we would either have to be quite clever or perform a computer search. A (naive) computer search would involve examining each of the \( 2^{\binom{43}{2}} = 2^{902} \) 2-colorings of the complete graph on 43 vertices. \( 2^{902} \) is rather large (larger than the number of particles in the known universe), so a computer search is unfeasible.

**Exercise 7.10.** Prove that Ramsey’s theorem implies the weak form of Szekeres’ theorem (i.e., it gives a “Happy Ending”).

**Theorem 7.11 (Erdős-Szekeres, 1935).** \( R(k, \ell) \leq \binom{k+\ell-2}{k-1} \), i.e., \( \binom{k+\ell}{k} \rightarrow (k+1, \ell+1) \).

From the above statement, it follows that \( R(k,k) < 4^k \). Using the probabilistic method, Erdős proved that \( 2^{k/2} < R(k,k) \).

More information about Ramsey numbers can be found in [CG98], the dynamic survey of Stanisław Radziszowski at http://www.combinatorics.org/ and at web-pages of Stanislaw Radziszowski (http://www.cs.rit.edu/~spr/homepage.html) and Brendan McKay (http://cs.anu.edu.au/~bdm/).
Chapter 8

8th day, Friday 7/2/04 (Scribe: Eric Patterson and Travis Schelldor)

8.1 Steiner Triple Systems

Exercise 8.1 (Puzzle). Let $A_1 \ldots A_n$ be a regular $n$-gon inscribed in the unit circle. Prove that $A_1A_2A_3 \ldots A_1A_n = n$. Here $XY$ is the length of the segment from $X$ to $Y$. Hint: $C$.

Recall [Definition 5.11] that a Steiner Triple System (STS) is a 3-uniform hypergraph [Definition 5.7(a,b)] such that for any pair of points (vertices) there is exactly one line passing through them.

Recall Exercise 5.12: Prove that in a STS, the number of vertices $n$ is odd.

Recall Exercise 5.18: If $n \equiv 1, 3 \pmod{6}$, then there exists a STS on $n$ points. (note that Exercise 5.13 is the converse of this statement). To get started on this, make sure to prove and use the

Lemma 8.2. For any $k$, we get a STS on $3^k$ points from the $k$-dimensional SET game (i.e. with $k$ characteristics).

The above lemma is part of Exercise 5.15.

Definition 8.3. The finite field $\mathbb{F}_3 = \{0, 1, 2\}$ is defined by taking arithmetic modulo 3. For any $k$ we can consider the $k$-dimensional $\mathbb{F}_3$-vector space, $\mathbb{F}_3^k = \{(\alpha_1, \ldots, \alpha_k) : \alpha \in \mathbb{F}_3\}$.

Definition 8.4. More generally, for any $p$ we can define $\mathbb{F}_p = \{0, 1, 2, \ldots, p - 1\}$, the finite field of order $p$, by taking arithmetic modulo $p$. We can also define the $k$-dimensional $\mathbb{F}_p$-vector space, denoted by $\mathbb{F}_p^k$.

Definition 8.5. For any vector space $F^k$ (where $F$ is any field such as $\mathbb{F}_p, \mathbb{C}, \mathbb{Q}, \mathbb{Q}[\sqrt{2}]$, etc.), we define a linear subspace to be a nonempty subset which is closed under addition and multiplication by scalars (elements of $F$). We define an affine subspace to be any translate of
a linear subspace (i.e. a linear subspace with origin moved to any given vector). (In particular, any linear subspace is also affine.) For any $v_1, \ldots, v_j \in F^k$, we can consider the subspace

$$\text{Span}(v_1, \ldots, v_s) = \left\{ \sum_{i=1}^{s} \alpha_i v_i : \alpha_i \in F \right\}. \quad (8.1)$$

We can also consider the affine span,

$$\text{Aff}(v_1, \ldots, v_s) = \left\{ \sum_{i=1}^{s} \alpha_i v_i : \alpha_i \in F, \sum_{i=1}^{s} \alpha_i = 1 \right\}. \quad (8.2)$$

Note that the linear (resp. affine) span is the smallest linear (resp. affine) subspace containing the given vectors. Also note that the span of $v_1, \ldots, v_s$ is the same as the affine span of $v_1, \ldots, v_s, 0$. Intuitively, the affine span is the smallest hyperplane of any dimension (not necessarily passing through the origin) containing the given vectors, and the linear span is the smallest hyperplane of any dimension which passes through 0 and contains the given vectors.

**Observation 8.6.** Note that if $\mathcal{T}$ is an affine subspace, and $v \in \mathcal{T}$, then $\mathcal{T} = \mathcal{U} + v$ where $\mathcal{U}$ is a linear subspace. Also, $\mathcal{U}$ is independent of the choice of $v$.

**Definition 8.7.** If $\mathcal{T}$ is an affine subspace $\mathcal{T} = v + \mathcal{U}$ where $\mathcal{U}$ is a linear subspace, then we define $\dim(\mathcal{T}) = \dim(\mathcal{U}) = \text{the maximum number of linearly independent vectors in } \mathcal{U}$.

Now we will consider some enumerative geometry over finite fields $F_p$.

**Exercise 8.8.** First of all, show that $|\mathbb{F}_p^k| = p^k$.

**Exercise 8.9.** Prove that, for all $x \neq y \in \mathbb{F}_p^k$, there exists a unique line through $x, y$, namely $\text{Aff}(x, y)$. (A line is by definition a set $\{v + tw \in \mathbb{F}_p^k : t \in \mathbb{F}_p\}$, for a fixed choice of $v, w \in \mathbb{F}_p^k$.)

**Exercise 8.10.** Show that the total number of lines in $\mathbb{F}_p^k$ is

$$\frac{\binom{p^k}{2}}{\frac{p}{p(p-1)/2}} = \frac{p^k(p^k - 1)}{p(p-1)}, \quad (8.3)$$

**Exercise 8.11.** Show that the total number of planes (i.e. two-dimensional affine subspaces) in $\mathbb{F}_p^k$ is

$$\frac{p^k(p^k - 1)(p^k - p)}{p^2(p^2 - 1)(p^2 - p)}. \quad (8.4)$$

**Definition 8.12.** For any $k$ and $p$, we define $AG(k, p)$ ("affine geometry") to be the set $\mathbb{F}_p^k$ with its affine lines, planes, etc.

**Exercise 8.13.** Verify that the "set" cards form the space $AG(4, 3)$, where the SETs are affine lines. (cf. Lecture 5, 6/25/04)
Exercise 8.14. Verify that the SET game $AG(k, 3)$ gives an STS for $n = 3^k$.

For the relevant definition of projective space and the Fundamental Theorem of Projective Geometry, see the Linear Algebra Methods in Combinatorics handout, pages 49 and 50.

Theorem 8.15 (Galois). A finite field of order $n$ exists iff $n = p^k$.

Consequence 8.16. If $n = p^k$, then there is a projective plane of order $n$.

Puzzle 8.17 (Euler's 36 Officers Problem). Given $n^2$ officers with $n$ ranks and $n$ divisions such that no officers have the same rank and division, put the officers in a $n \times n$ grid so that neither rank nor division occurs twice in any row or column. It can be done for any prime power $p$ but not for 6.

Exercise 8.18. Prove that the impossibility of Euler's 36 officers problem implies the nonexistence of a projective plane of order 6.
Chapter 9

9th day, Wednesday 7/7/04 (Scribe: Ivona Bezáková)

9.1 Extremal Set Theory

Notation: \([n] = \{1, \ldots, n\}\). The incidence vector of a set \(A \subseteq [n]\) is defined as \(\nu_A \in \mathbb{F}_2^n\) where \((\nu_A)_i = 1\) if and only if \(i \in A\). Standard inner product of two vectors \(a, b \in \mathbb{F}^n\) is defined as \(a \cdot b = \sum_{i=1}^{n} a_i b_i\). Notice that \(\nu_A \cdot \nu_B = |A \cap B|\) and \(\nu_A \cdot \nu_A = |A|\).

Theorem 9.1 (Fisher’s Inequality). Let \(A_1, \ldots, A_m \subseteq [n]\), and \(a \geq 1\). If \(|A_i \cap A_j| = a\) for every \(i \neq j\), then \(m \leq n\).

In 1949 R.C. Bose proved Fisher’s Inequality by showing linear independence of a well-chosen set of vectors over a well-chosen field, establishing the “linear algebra method” in combinatorics. The proof is based on the following lemma:

Lemma 9.2. The incidence vectors of the \(A_i\) are linearly independent.

If the size of one of the sets is \(a\), e.g. \(|A_i| = a\), then \(A_j \supseteq A_i\) for every \(j\) and the system forms a **sunflower** (no two sets intersect outside the common intersection of all sets).

9.1.1 Eventown

Suppose there is a town of \(n\) citizens and there are \(m\) clubs \(A_1, \ldots, A_m \subseteq [n]\). The lawmakers tend to create new laws and curiously examine the highest possible number of clubs under the current set of rules.

Rules in the Eventown:

0) The clubs have to be distinct.

1) Each club has even number of members.
(2) \(|A_i \cap A_j|\) is even for every \(i, j\).

Under rule (0), there could be total \(2^n\) clubs. Under rules (0) and (1), the total number of clubs drops to \(2^{n-1}\). If all three rules need to be satisfied, the “married couples solution” provides \(2^{n/2}\) clubs. This solution is maximal, i.e., the solution cannot be extended, adding another club would violate one of the rules. Is this solution also maximum, i.e., there is no solution with higher number of clubs?

Exercise 9.3. In Eventown, \(m \leq 2^{[n/2]}\). (The “married couples solution” is maximum.)

Hint. Prove that the statement follows from Exercise 9.5.

Exercise 9.4. Find another maximum solution which contains three clubs \(A_1, A_2, A_3\) such that \(|A_1 \cap A_2 \cap A_3|\) is odd. (So not all maximum solutions are isomorphic.)

Exercise 9.5. In Eventown every maximal set of clubs is maximum.

9.1.2 Oddtown

Rules in Oddtown:

(1) \(|A_i|\) is odd for every \(i\)

(2) \(|A_i \cap A_j|\) is even for every \(i \neq j\)

Exercise 9.6. The number of ways to create a system of \(n\) clubs in Oddtown is \(\geq 2^{n^2/8}\).

Theorem 9.7 (Oddtown Theorem, Berlekamp). In Oddtown, \(m \leq n\).

Lemma 9.8. Under Oddtown rules the incidence vectors of the clubs are linearly independent.

Exercise 9.9. Prove the previous lemma over the following fields: (a) over \(\mathbb{Q}\), (b) over \(\mathbb{F}_2\), (c) over \(\mathbb{R}\).

Definition 9.10. For \(A \subseteq [n]\), the incidence vector (or characteristic vector) of \(A\) is the vector \(v_A = (\alpha_1, \ldots, \alpha_n)\) where

\[
\alpha_i = \begin{cases} 
1 & \text{if } i \in A \\
0 & \text{if } i \notin A 
\end{cases}
\]

Let \(B\) be a matrix with rows \(v_{A_i}\). Part (a) implies that \(B\) has a full rank over \(\mathbb{Q}\). Notice that rank is invariant under extension of the field (Gaussian elimination process keeps all coefficients in the original field), therefore part (c) is a consequence of part (a). Notice that \(\mathbb{F}_2\) is not a subfield of \(\mathbb{Q}\). However, part (a) follows from (b) and the following exercise.

Exercise 9.11. Let \(A\) be a \((0, 1)\)-matrix, i.e., its entries are from \(\{0, 1\}\). Let \(rk_p(A)\) denote the rank of \(A\) over \(\mathbb{F}_p\) and let \(rk_0(A)\) be its rank over \(\mathbb{Q}\). Prove: \(rk_p(A) \leq rk_0(A)\).
Definition 9.12. Vectors \( a, b \in \mathbb{F}^n \) are **perpendicular**, denoted \( a \perp b \), if \( a \cdot b = 0 \). For \( S \subseteq \mathbb{F}^n \), the set \( S^\perp = \{ x \mid (\forall y \in S)(x \perp y) \} \) is called \( S \)-**perp**. A vector \( v \in \mathbb{F}^n \) is called **isotropic** if \( v \perp v \). A set \( U \subseteq \mathbb{F}^n \) is **totally isotropic** if \( U \perp U \).

**Exercise 9.13.** Prove \( S^\perp = (\text{Span } S)^\perp \).

**Exercise 9.14.** Let \( U \subseteq \mathbb{F}^n \). Prove: \( \dim(U) + \dim(U^\perp) = n \).

Let \( U \subseteq \mathbb{F}_2^n \) be a maximal set of clubs in Eventown. Then \( U \subseteq U^\perp \) and therefore \( \text{Span}(U) \subseteq \text{Span}(U)^\perp \). Since \( U \) is maximal, we can conclude that \( U = \text{Span}(U) \). By previous exercise, \( \dim(U) \leq n/2 \). Therefore \( |U| \leq |\mathbb{F}_2^{n/2}| = 2^{n/2} \), i.e., the married couples solution is optimal.

**Exercise 9.15.** For what \( p \) does there exist a nonzero isotropic vector in \( \mathbb{F}_p^2 \)? (Answer is appealing.)

**Exercise 9.16.** Prove that there exists an \( n \)-dimensional totally isotropic subspace over \( \mathbb{C}^{2n} \).

**Question:** How many sets can have pairwise intersection of size 0 or 1? If we take all sets of size at most 2, we get a set system of \( \binom{n}{2} + n + 1 \) sets.

**Exercise^+ 9.17.** Let \( A_1, \ldots, A_m \subseteq [n] \) be such that \( |A_i \cap A_j| \leq 1 \) for \( i \neq j \). Prove: \( m \leq \binom{n}{2} + n + 1 \)

**Hint.** Use linear algebra method. The trick lies in finding a good set of vectors in dimension \( \binom{n}{2} + n + 1 \).
Chapter 10

10th day, Friday 7/9/04 (Scribe: Ivona Bezáková)

10.1 Cauchy-Hilbert Matrix

The Hilbert matrix is defined as:

\[
\begin{bmatrix}
\frac{1}{1} & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} \\
\frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots & \frac{1}{n+1} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{1}{n} & \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \frac{1}{2n-1}
\end{bmatrix}
\]

The Hilbert matrix is a special case of a Cauchy matrix:

\[
\begin{bmatrix}
\frac{1}{\alpha_1-\beta_1} & \frac{1}{\alpha_1-\beta_2} & \frac{1}{\alpha_1-\beta_3} & \cdots & \frac{1}{\alpha_1-\beta_n} \\
\frac{1}{\alpha_2-\beta_1} & \frac{1}{\alpha_2-\beta_2} & \frac{1}{\alpha_2-\beta_3} & \cdots & \frac{1}{\alpha_2-\beta_n} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{1}{\alpha_n-\beta_1} & \frac{1}{\alpha_n-\beta_2} & \frac{1}{\alpha_n-\beta_3} & \cdots & \frac{1}{\alpha_n-\beta_n}
\end{bmatrix}
\]

Exercise 10.1. Prove: If all \(\alpha_i, \beta_i\) are distinct then the Cauchy-Hilbert matrix has full rank.

10.2 Cauchy’s Functional Equation

Let \(f : \mathbb{R} \to \mathbb{R}\) be a real function.

Definition 10.2. Cauchy’s functional equation specifies:

\[(\forall x, y)(f(x+y) = f(x) + f(y)).\]
Which functions are solutions to Cauchy’s functional equation (CFE)? Clearly, \( f(x) = cx \) is a solution.

We will make several observations: if \( f \) satisfies CFE then

1. \( f(0) = 0 \). (Why?)
2. Let \( c := f(1) \). Then \( (\forall x \in \mathbb{Z})(f(x) = cx) \).
3. \( f \left( \frac{p}{q} \right) = c \cdot \frac{p}{q} \) for \( p, q \in \mathbb{Z} \). (Why?)

Therefore \( (\forall x \in \mathbb{Q})(f(x) = cx) \).

**Exercise 10.3.** If \( f \) is continuous then \( (\forall x)(f(x) = cx) \).

**Exercise 10.4.** If \( f \) is a solution of CFE and it is bounded in an interval then \( (\forall x)(f(x) = cx) \).

This is a known fact from linear algebra:

**Definition 10.5.** Let \( F \) be a field and \( U, W \) vector spaces over \( F \). A function \( \varphi : V \to W \) is called linear if \((\forall x, y \in V)(\varphi(x + y) = \varphi(x) + \varphi(y)) \) and \((\forall x \in V)(\forall \alpha \in F)(\varphi(\alpha x) = \alpha \varphi(x)) \).

**Lemma 10.6.** Let \( F \) be a field and \( V, W \) be two vector spaces over \( F \). Suppose \( B \subseteq V \) be a basis and \( \varphi : B \to W \) be an arbitrary function. Then \( \varphi \) can be uniquely extended to a linear function \( \varphi : V \to W \).

**Hamel basis** \( H \) is a basis of \( \mathbb{R} \) over \( \mathbb{Q} \) (its existence is implied by Zorn’s Lemma). Its cardinality is **continuum**, denoted as \( c \). CFE translates to finding all \( \mathbb{Q} \)-linear (linear over \( \mathbb{Q} \)) functions \( f : \mathbb{R} \to \mathbb{R} \). Using previous lemma we can define the function \( f \) by assigning a value \( f(h) \) to every vector \( h \) in the Hamel basis.

**Corollary 10.7.** There are solutions to CFE not of the form \( f(x) = cx \). In fact in \( \alpha_1, \alpha_2, \ldots \) are linearly independent over \( \mathbb{Q} \) then \( f(\alpha_1), f(\alpha_2), \ldots \) can be independently prescribed.

### 10.3 Number Theory

**Exercise 10.8.** Pick two positive integers \( x, y \) at random, what is the probability that \( x, y \) are relatively prime? (Hint. What does “integer chosen at random” mean? We want all integers to be chosen with the same probability. First make sense out of the problem statement. You will define the probability as a limit. Assume the limit exists. Then use \( \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \).)

A little diversion: \( \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \) was stated by Fermat and proved by Leonhard Euler. Other examples of Fermat’s theorems that Euler worked on include Fermat’s Last Theorem stating that \( x^n + y^n = z^n, n \geq 3 \) implies \( xyz = 0 \) (recently proved by Andrew Wiles), and the following remarkable result:
Exercise 10.9. Let $p > 2$ be a prime. Prove that $p = a^2 + b^2$ for $a, b \in \mathbb{Z}$ if and only if $p \equiv 1 \mod 4$.

Exercise 10.10. Prove that $\mathbb{F}_p[\sqrt{-1}] = \{a + bi \mid a, b \in \mathbb{F}_p\}$ (“complex numbers mod $p$”) is a field if and only if $p \equiv -1 \mod 4$.

Exercise 10.11. Let $\{a_i\}_{i=0}^\infty$ be defined as $a_0 = 1$, $a_{n+1} = \sqrt{2}a_n$. Prove that $(\forall n)(a_n < 2)$.

Exercise 10.12. Prove that $a_n \rightarrow 2$ for the $a_i$ from the previous exercise.

Exercise 10.13. Find the maximum $r$ s.t. $r^r$ is bounded.

Exercise 10.14. Find two infinite sets $A, B \subseteq \mathbb{N} = \{0, 1, 2, \ldots\}$ such that for every $n \in \mathbb{N}$ there exist unique $a \in A$ and $b \in B$ such that $a + b = n$. (HINT. Look at the digits of the integers. Recall Hales-Jewett’s implies the Van der Waerden’s Theorem.)

10.4 Ramsey Theory

By Erdős-Szekeres $N = \binom{k+\ell-2}{k-1} \rightarrow (k, \ell)$, i.e. if $G$ is a graph with $N$ vertices then $G$ contains a $K_k$ or $\overline{G}$ contains a $K_{\ell}$. Turán’s example, a graph of $n^2$ vertices consisting of $n$ independent copies of $K_n$, shows that $n^2 \not\rightarrow (n+1, n+1)$. The following theorem states a much stronger result.

Theorem 10.15 (Erdős). $2^{n/2} \not\rightarrow (n+1, n+1)$

In other words, $N \not\rightarrow (2 \log_2 N + 1, 2 \log_2 N + 1)$. The proof is based on the “probabilistic method.”

Corollary 10.16. If $2^{\binom{n}{k}} \leq 2^{\binom{1}{2}}$ then $n \not\rightarrow (k, k)$.

Exercise 10.17. Prove: $\binom{n}{k} \leq \frac{n^k}{k!}$.

We know that $4^n \rightarrow (n+1, n+1)$ (why?), i.e. $N \rightarrow (1/2 \log N, 1/2 \log N)$. By the above $N \not\rightarrow (2 \log N + 1, 2 \log N + 1)$. There is a multiplicative gap of 4 between the lower bound and the upper bound.

Open problems.

1. Shrink this gap: either prove $N \rightarrow ((1/2 + \varepsilon) N, (1/2 + \varepsilon) N)$ or prove $N \not\rightarrow ((2 - \varepsilon) N, (2 - \varepsilon) N)$.

2. Prove: $((\exists \varepsilon)(\forall \varepsilon > 0)(\exists n_0)(\forall n > n_0)(n \rightarrow ((c - \varepsilon) n, (c - \varepsilon) n) and n \not\rightarrow ((c + \varepsilon) n, (c + \varepsilon) n))$
CHAPTER 10.
11.1 Totally unimodular matrices

**Definition 11.1.** A matrix is **totally unimodular** if the determinant of every square submatrix is 0, 1, or −1.

**Definition 11.2.** Let $G$ be a digraph with $n$ vertices and $m$ edges. The **incidence matrix** $A$ of $G$ is a $n \times m$ matrix with rows corresponding to vertices and columns corresponding to edges s.t.

$$A[v, e] = \begin{cases} 
0 & \text{if } v \text{ is not incident to } e, \\
1 & \text{if } e \text{ starts at } v, \\
-1 & \text{if } e \text{ ends at } v.
\end{cases}$$

**Exercise 11.3.** Prove: The incidence matrix of a digraph is totally unimodular.

11.2 Latin squares

**Definition 11.4.** Let $S$ be an $n$-element set of symbols. A $n \times n$ matrix with entries from $S$ is called a **Latin square** if each symbol appears exactly once in each row and in each column. $L(n)$ denotes the number of $n \times n$ Latin squares.

**Theorem 11.5.** For every $\varepsilon > 0$ there exists $n_0$ such that for every $n > n_0$

$$n^{(1-\varepsilon)n^2} \leq L(n) \leq n^{n^2}.$$ 

Rephrasing the theorem we get $\log L(n) \sim n^2 \log n$.

We say that two Latin squares $A$ and $B$ are **equivalent**, denoted $A \sim B$, if $A$ is obtained from $B$ by a sequence of the following: (1) permuting rows, (2) permuting columns, (3) permuting symbols, (4) permuting the roles of rows, column indices, and symbols.

An example of type (4) equivalence would replace entry $a$ in row $b$, column $c$ by entry $c$ in row $a$, column $b$. 

51
Exercise 11.6. The number of Latin squares equivalent to a given Latin square is \( \leq 6(n!)^3 \).

Exercise 11.7. Let \( \tilde{L}(n) \) be the number of inequivalent Latin squares of order \( n \). Prove \( \log \tilde{L}(n) \sim \log L(n) \).

It is easy to prove the upper bound from Theorem 11.5 since \( L(n) < (n!)^n < n^{n^2} \) (why?). The lower bound is related to counting perfect matchings in a bipartite graph.

11.3 Counting perfect matchings in a bipartite graph; the permanent

A graph \( G \) is bipartite if its vertex set can be partitioned into \( V_1 \) and \( V_2 \) so that every edge in \( G \) has one endpoint in \( V_1 \) and the other in \( V_2 \). A matching \( M \subseteq E(G) \) is a set of edges such that no two edges in \( M \) share an endpoint. A matching is perfect if it contains \( |V(G)|/2 \) edges.

Definition 11.8. Let \( A = (a_{i,j})_{n \times n} \) be a matrix. The permanent of \( A \) is defined as

\[
\text{per } A = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i,\sigma(i)},
\]

where \( S_n \) is the set of all permutations of \( [n] = \{1, \ldots, n\} \).

Let \( G \) be a bipartite graph with partitions \( V_1 \) and \( V_2 \) where \( V_1 = [k_1] \times \{1\} \) and \( V_2 = [k_2] \times \{2\} \). The incidence matrix of \( G \) is a \( k_1 \times k_2 \) matrix \( M = (\alpha_{i,j}) \) defined by

\[
\alpha_{i,j} = \begin{cases} 
1 & \text{if } (i,1) \sim (j,2), \text{ i.e., } (i,1) \text{ is adjacent to } (j,2), \\
0 & \text{otherwise.}
\end{cases}
\]

Theorem 11.9. Let \( G \) be a bipartite graph with \( |V_1| = |V_2| \) and let \( M \) be its incidence matrix. Then the number of perfect matchings of \( G \) is \( \text{per}(M) \).

Theorem 11.10. A regular bipartite graph of degree \( r \) with \( n + n \) vertices has \( > (r/e)^n \) perfect matchings.

11.4 Doubly Stochastic Matrices

Definition 11.11. A \( n \times n \) matrix is called stochastic if all its entries are nonnegative and every row sums to 1. Matrix \( A \) is doubly stochastic if both \( A \) and \( A^T \) are stochastic, i.e., the entries are nonnegative and every row and every column sums to 1.
11.4. **DOUBLY STOCHASTIC MATRICES**

Let $I$ be the identity matrix and $J$ be the all-ones matrix.

\[
I = \begin{pmatrix}
1 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & 0 & \ldots & 1
\end{pmatrix} \quad J = \begin{pmatrix}
1 & 1 & \ldots & 1 \\
1 & 1 & \ldots & 1 \\
\vdots & \ddots & \ddots & \vdots \\
1 & 1 & \ldots & 1
\end{pmatrix}
\]

**Example 11.12.** $I$ and $\frac{1}{n} J$ are doubly stochastic.

**Exercise 11.13.** $n! > (n/e)^n$. (HINT: $e^x = \sum_{n=0}^{\infty} x^n/n!$)

Clearly, $\text{per}(I) = 1$ and $\text{per}(\frac{1}{n} J) = \frac{n!}{n^n} > \frac{1}{e^n}$.

**Comment on how not to prove this inequality:** Notice that the Stirling’s formula

\[
n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}
\]

implies the inequality $n! > (\frac{n}{e})^n$ only for sufficiently large $n$, while the above method proves it for every $n$. A more precise Stirling’s formula

\[
n! = \left(\frac{n}{e}\right)^n \sqrt{2\pi n} \left(1 + \frac{\theta_n}{12n}\right),
\]

where $|\theta_n| \leq 1$, would imply the inequality for every $n$.

**Exercise 11.14.** If $A$ is doubly stochastic then $\text{per}(A) \leq 1$. The equality holds exactly when $A$ is a permutation matrix (i.e., $A$ has exactly one 1 in each row and column, otherwise zeros.)

**Theorem 11.15 (The Permanent Inequality).** If $A$ is doubly stochastic then

\[
\text{per} A \geq \text{per} \left(\frac{1}{n} J\right) = \frac{n!}{n^n}.
\]

The theorem was conjectured by van der Waerden and was known as van der Waerden’s Permanent Conjecture. It was proved independently by Egorichev and Falikman in 1980.

**Exercise 11.16.** Let $M$ be the incidence matrix of a regular bipartite graph of degree $r$. Prove that $\frac{1}{r} M$ is doubly stochastic.

We show that the permanent inequality implies Theorem [11.10]. Notice that $\text{per}(\alpha A) = \alpha^n \text{per}(A)$. Let $M$ be the incidence matrix of the bipartite graph from Theorem [11.10]. By applying the permanent inequality and Exercise [11.13] we get

\[
\text{per}(M) = r^n \text{per} \left(\frac{1}{r} M\right) > r^n \frac{1}{e^n}.
\]
11.5 Back to Latin squares

**Definition 11.17.** For $k \leq n$, a $k \times n$ Latin rectangle is a $k \times n$ matrix of $n$ symbols such that every symbol appears at most once in each row and in each column. $L(k,n)$ denotes the number of $k \times n$ Latin rectangles.

Obviously, $L(n,k) \leq (n!)^k$ and $L(n,1) = n!$.

**Exercise 11.18.** Prove: $L(n,2) \sim \frac{[n!]^2}{e}$.

**Theorem 11.19.** Let $k < n$. Every $k \times n$ Latin rectangle can be extended to a $(k+1) \times n$ Latin rectangle.

**Theorem 11.20.** Every regular bipartite graph of degree $\geq 1$ has a perfect matching.


Suppose we have $k \times n$ Latin rectangle $L$. Define a bipartite graph $G$ on $n + n$ vertices as follows. An edge goes from $(i,1)$ to $(j,2)$ if the $i$-th column in $L$ does not contain the symbol $j$. This way we obtain a regular bipartite graph (of degree $n - k$). The perfect matchings of $G$ correspond to all possible extensions of $L$ to a $(k + 1) \times n$ Latin rectangle.

Therefore $L(n,k+1) \geq L(n,k)\left(\frac{n-k}{e}\right)^n$, implying

$$L(n) = L(n,n) \geq \left(\frac{n}{e}\right)^n \left(\frac{n-1}{e}\right)^n \left(\frac{n-2}{e}\right)^n \cdots \left(\frac{1}{e}\right)^n = \frac{(n!)^n}{e^n}.$$ 

**Exercise 11.22.** Prove: $n \ln(n!) - n^2 \sim n \ln(n!)$; therefore $\log L(n) \sim n^2 \log n$.

11.6 Orthogonal Latin Squares

**Definition 11.23.** Let $A$ and $B$ be two $n \times n$ Latin squares. $A$ and $B$ are orthogonal if the set $\{(a_{i,j}, b_{i,j}) \mid i, j \in [n]\}$ contains all the $n^2$ distinct pairs, i.e., no pair $(a_{i,j}, b_{i,j})$ is repeated.

Orthogonal Latin squares are related to Euler’s “36 officers” problem: There are 36 officers, 6 divisions and 6 ranks. The task is to assign a division and a rank to every officer so that no two officers are assigned the same combination of division and rank; and arrange the officers in a $6 \times 6$ array such that each division is represented in each row and each column and each rank is represented in each row and each column.

**Exercise 11.24.** Let $p$ be a prime. Prove that there exist $p - 1$ pairwise orthogonal Latin squares of order $p$.

**Exercise 11.25.** Prove previous exercise for $p$ a prime power.
Theorem 11.26 (Tarry’1900). Euler’s 36 officers problem does not have a solution, i.e., there does not exist a pair of orthogonal $6 \times 6$ Latin squares. (Proof: tedious.)

Exercise 11.27. Prove that for every $n$ there are at most $n - 1$ pairwise orthogonal $n \times n$ Latin squares.

Exercise 11.28. Show that the following are equivalent: (a) there exists a set of $n-1$ pairwise orthogonal $n \times n$ Latin squares, (b) there exists a projective plane of order $n$.

Exercise 11.29. Prove: $\binom{n}{k} \geq \left( \frac{n}{k} \right)^k$.

Exercise 11.30. Prove: $\binom{n}{k} < \left( \frac{en}{k} \right)^k$. Find an elegant solution, like the proof of $n! > \left( \frac{n}{e} \right)^n$.

Exercise 11.31. Prove: $\left( \frac{en}{k} \right)^k > \binom{n}{k} + \binom{n}{k-1} + \cdots + \binom{n}{0}$.

Exercise 11.32. In $\mathbb{R}^n$ find $cn^2$ points with only two pairwise distances.

The vertices of a regular pentagon are an example of a set in $\mathbb{R}^2$ with only two pairwise distances.

Exercise 11.33. Prove: Any six points in the plane span at least three different distances.
12.1 Constructive Proofs of Negative Results in Ramsey Theory

Recall that we discussed the following results:

- $4^n \to (n+1,n+1)$, proved by Erdős-Szekeres,
- $2^{n/2} \not\to (n+1,n+1)$, showed by Erdős using a probabilistic proof of existence.

So far we have seen only one constructive proof of a negative result, the trivial example observed by Turán showing that $n^2 \not\to (n+1,n+1)$.

**Theorem 12.1 (Zsigmond Nagy, 1973).** \binom{n}{3} \not\to (n+1,n+1) can be proved constructively.

Nagy’s construction defines the graph as follows. The vertices are all the 3-element subsets of $[n]$. Sets $A$ and $B$ are adjacent if $|A \cap B| = 1$. The Erdős-deBruijn Theorem, a special case of Fisher’s inequality for sets with intersection size 1, implies that Nagy’s graph does not contain a clique of size $n+1$. By the Oddtown Theorem the graph does not contain an anticlique of size $n+1$.

The best known explicit construction is by Frankl and Wilson.

**Theorem 12.2 (Frankl - Wilson).** Let $\varepsilon > 0$. For sufficiently large $n$ one can construct a graph of at most $n(1-\varepsilon) \frac{\ln n}{\ln \ln n}$ vertices with no clique or anticlique of size $n+1$.

Let $p$ be a prime. The vertices in the Frankl - Wilson graph are all the subsets of $[2p^2 - 1]$ of size $p^2 - 1$. Sets $A$ and $B$ are adjacent if $|A \cap B| \equiv -1 \pmod{p}$.

**Claim 12.3.** The Frankl - Wilson example proves that

$$\binom{2p^2 - 1}{p^2 - 1} \not\to \binom{2p^2 - 1}{p - 1} + 1.$$
The proof of the claim is based on two theorems on extremal set theory:

**Theorem 12.4 (Ray-Chaudhuri - Wilson, 1975).** Fix $k$ and let $l_1 < \cdots < l_s < k$. If $A_1, \ldots, A_m \subseteq [n]$ are sets of size $k$ such that $|A_i \cap A_j| \in \{l_1, \ldots, l_s\}$ for every $i \neq j$, then $m \leq \binom{n}{s}$. 

**Exercise 12.5.** Prove that the Ray-Chaudhuri - Wilson Theorem is tight, i.e. find $\binom{n}{s}$ sets with $s$ different intersection sizes.

Frankl and Wilson generalized the Ray-Chaudhuri - Wilson Theorem:

**Theorem 12.6 (Frankl - Wilson, 1981).** Let $p$ be a prime and let $l_1, \ldots, l_s \in \mathbb{Z}_p$ be such that $k \not\equiv l_1, \ldots, l_s \mod p$. If $A_1, \ldots, A_m \subseteq [n]$ are sets of size $k$ such that $|A_i \cap A_j| \in \{l_1, \ldots, l_s\}$ mod $p$ for every $i \neq j$, then $m \leq \binom{n}{s}$. 

A clique in the Frankl - Wilson graph corresponds to a set system $A_1, \ldots, A_m$, $|A_i| = p^2 - 1$ for every $i$, such that $|A_i \cap A_j| \in \{p - 1, 2p - 1, \ldots, p(p - 1) - 1\}$. Thus the Ray-Chaudhuri - Wilson Theorem implies that $m \leq \binom{2p^2 - 1}{p - 1}$.

An ant клиque corresponds to a set system $B_1, \ldots, B_m \subseteq [2p^2 - 1]$, $|B_i| = p^2 - 1 \equiv -1 \mod p$ for every $i$, such that $|B_i \cap B_j| \in \{0, 1, \ldots, p - 2\}$ mod $p$. By the Frankl - Wilson Theorem, $m \leq \binom{2p^2 - 1}{p - 1}$.

### 12.2 Bipartite Ramsey Theory

We define a bipartite version of the Erdős-Rado arrow.

**Definition 12.7.** We say that $a \sim (b, c)$ if every bipartite graph $G$ with $a$ vertices contains a bipartite clique $K_{b, b}$ or the complement $\overline{G}$ contains a bipartite clique $K_{c, c}$.

**Exercise 12.8.** Prove: $4^n \sim (n + 1, n + 1)$.

**Exercise 12.9.** Prove: $2^{n/2} \not\sim (n + 1, n + 1)$.

**Hint.** Probabilistic proof of existence.

### 12.3 Hadamard Matrices

**Theorem 12.10 (Hadamard’s Inequality).** Let $A \in M_n(\mathbb{R})$, i.e. $A$ is an $n \times n$ matrix over $\mathbb{R}$. Then

$$|\det(A)| \leq \prod_{i=1}^{n} \|a_i\| ,$$

where $a_i$ is the vector in the $i$-th row of $A$ and $\|a_i\| = \sqrt{\sum_{j=1}^{n} a_{i,j}^2}$ is its norm. The equality holds if and only if there exists a zero row or if the rows are pairwise orthogonal.
Definition 12.11. An Hadamard matrix is a $\pm 1$-matrix with all rows orthogonal.

The Sylvester matrices are $2^k \times 2^k$ matrices defined by the following matrix recurrence:

\[
H_0 = \begin{pmatrix} 1 \end{pmatrix} \\
H_{k+1} = \begin{pmatrix} H_k & H_k \\ H_k & -H_k \end{pmatrix} \quad \text{for } k > 0
\]


Definition 12.13. $n$ is an Hadamard number if there exists a $n \times n$ Hadamard matrix.

Exercise 12.14. If $n$ is an Hadamard number then $n \leq 2$ or $n$ is divisible by 4.

Exercise$^+$ 12.15. If $p \equiv -1 \pmod{4}$ is a prime power, then $p + 1$ is an Hadamard number. Hint. Quadratic residues.

Exercise 12.16. If $H$ is an Hadamard matrix then $H^T$ is also an Hadamard matrix. Hint. Examine $HH^T$.

Exercise 12.17. Construct a matrix with all rows orthogonal such that the columns are not orthogonal.

Definition 12.18. Let $M = (m_{i,j})$ be a $n \times n$ matrix and let $I, J \subseteq [n]$. A rectangle is a submatrix of $M$ corresponding to rows defined by $I$ and columns defined by $J$. The discrepancy of the rectangle is defined as

\[
\text{disc}_{I,J}(M) = \left| \sum_{i \in I, j \in J} m_{i,j} \right|.
\]

Theorem 12.19 (Lindsay's Inequality). Let $a = |I|$ and $b = |J|$. If $H$ is an Hadamard matrix then

\[
\text{disc}_{I,J}(H) \leq \sqrt{nab}
\]

Note 12.20 (Back to bipartite Ramsey numbers.). Let $M$ be an incidence matrix of a bipartite graph, with zeros replaced by $-1$. The rectangle corresponding to a bipartite clique or anticlique of size $t$ has discrepancy $t^2$. By Lindsay's Inequality, $t^2 \leq t\sqrt{n}$, so $t \leq \sqrt{n}$. Therefore

\[ n^2 \not\leq (n+1, n+1). \]

The proof of Lindsay’s inequality is based on the following facts:

Theorem 12.21 (Cauchy-Schwarz Inequality). Let $a, b \in \mathbb{R}^n$. Then

\[ |a \cdot b| \leq \|a\| \|b\|, \]

where $\cdot$ denotes the standard inner product.
Definition 12.22. A matrix $K$ is orthogonal if $K^T = K^{-1}$.

Lemma 12.23. If a $n \times n$ matrix $K$ is orthogonal, then $\|Kx\| = \|x\|$ for every vector $x \in \mathbb{F}^n$.

Proof: [of Lindsay’s Inequality] Let $e_S \in \mathbb{F}^n$ be the characteristic vector of the set $S \subseteq [n]$. Then

$$\sum_{i \in I, j \in J} m_{i,j} = e_I^* M e_J$$

If $H$ is Hadamard then $\frac{1}{\sqrt{n}} H$ is orthogonal. By Cauchy-Schwarz and the above lemma,

$$\text{discr}_{I,J}(H) = |e_I^* H e_J| \leq \|e_I\| \|H e_J\| = \sqrt{a} \sqrt{n} b$$

Exercise 12.24. Prove: If $a$ and $b$ are Hadamard numbers then $ab$ is a Hadamard number.

Conjecture 12.25. If $n$ is divisible by 4 then $n$ is a Hadamard number.

Definition 12.26. The upper density of a set $A \subseteq \mathbb{N}$ is

$$\limsup_{n \to \infty} \frac{|A(n)|}{n},$$

where $A(n) = \{x \in A \mid x \leq n\}$. The lower density is

$$\liminf_{n \to \infty} \frac{|A(n)|}{n}.$$

We say that $A$ has density $\gamma$ if $\gamma$ is both the lower and the upper density.

Exercise 12.27. Construct a set with lower density 0 and upper density 1.

Exercise 12.28. The upper density of Hadamard numbers is $\leq 1/4$.

The Conjecture 12.25 would imply that the density of Hadamard numbers is $1/4$.

OPEN PROBLEM: Is the (upper) density of Hadamard numbers positive? The Sylvester matrices example shows that there are infinitely many Hadamard matrices. The quadratic residue Hadamard matrices give asymptotic density $\frac{3}{2\ln n}$. However, the density of this set is still 0.
Chapter 13

13th day, Tuesday 7/20/04 (Scribe: Charilaos Skiadas)

13.1 The Gale-Berlekamp Switching game

In this game there are 2 players and they play on a $n \times n$ checkerboard. There are exactly two moves in this game. In the first move, player 1 assigns a sign (+1) to every square. Then, the second player switches the signs in any rows and columns he wants. After he finishes, they sum all numbers in the final configuration and look at the absolute value of the result. This is the amount the 1st player pays to the 2nd player. So naturally, the 1st player wants to minimize this number, while the 2nd player wants to maximize it. We want to know what the value of this game is, at least asymptotically in $n$. The value is the amount that the player 2 is guaranteed to receive, or equivalently the largest amount that the player 1 will have to pay, assuming both players follow optimal strategies. This value is certainly no more than $n^2$.

Player 2 can easily achieve $n$ in the following way: Make everything in the first row a plus, by switching columns. Then, if the sum is $a$, we could switch the first row to get sum $a - 2n$. Since $\max\{|a|, |a - 2n|\} \geq n$, we see that the value of the game is at least $n$. The question is: Can he achieve more?

Let us first make some observations: First of all, the order in which player 2 makes the switches does not matter, so we can assume that he switches the columns first, and then switches the rows. After he has switched the columns, then his best strategy is to switch the rows with negative sum, so that all the rows will have nonnegative sum. In particular, his row-switching moves are completely determined (except for the rows with zero sum, where switching makes no difference). If player 2 could arrange it so that many row sums are large, then he could improve his win.

So our goal should be to switch the columns so as to maximize the absolute value of the sum on an "average" row. Our goal is to achieve approximately $c\sqrt{n}$ gain per row. So this would give $c' n^{3/2}$ total gain. The answer is to randomly decide whether to switch a column or not. The main observation is that a random sequence of $n$ pluses and minuses has expected absolute sum $\Theta(\sqrt{n})$: 

\[ \sum_{j=1}^{n} a_j = \Theta(\sqrt{n}) \]
CHAPTER 13.

Assume for simplicity that \( n \) is even. We see that the probability

\[
P(\text{coin flips, } k \text{ heads}) = \left( \frac{n}{2} \right) \frac{n}{2^n}.\]

Asymptotically,

\[
\frac{n}{2^n} = \frac{n!}{(\frac{n}{2})!^2} = \cdots \approx \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{n}} = \frac{c}{\sqrt{n}},
\]

were we used Stirling’s formula to estimate the factorials. So

\[
P(\# \text{heads within } \frac{n}{2} \pm k) = \left( \frac{n}{2} - k \right) + \cdots + \left( \frac{n}{2} + k \right) \cdot \frac{1}{2^n}.
\]

There are precise estimates for this sum, but for our purposes, a rough estimate is enough. Since the middle binomial coefficient is the largest one, we see that the above sum is less than

\[
\frac{(2k + 1) \left( \frac{n}{2} \right)}{2^n} \sim (2k + 1) \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{n}}.
\]

As long as \((2k + 1) \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{n}} \leq \frac{1}{2}\), i.e., as long as \( k < c\sqrt{n} \), where \( c := \frac{\sqrt{\pi}}{4\sqrt{2}} \), the above probability is less than \( \frac{1}{2} \). Then \( P(\# \text{heads is within } \frac{n}{2} \pm c\sqrt{n}) \) is asymptotically at most \( \frac{1}{2} \).

(Note: The Central Limit Theorem, says that asymptotically, the binomial coefficients fit on a bell curve, and can hence obtain a better value for the constant above.)

So if in our game the 2nd player randomly switches the columns, then we expect at least roughly half of the rows to have absolute sum greater than \( c\sqrt{n} \).

This means that on average, at least \( \frac{n}{2} \) of the rows have this advantage. But this means that it is possible for player 2 to achieve this advantage, which would give him overall at least \( \frac{1}{2}cn^{3/2} \) total gain.

The question now arises: Can we recommend a strategy for player 1 to force the total to be no more than that order? One could argue in a similar way with a random choice, but in this case there actually is a deterministic strategy. The idea is to use Hadamard matrices. Recall that a Hadamard matrix is a matrix with entries only \( \pm 1 \), where the rows are orthogonal. As a consequence, all columns are orthogonal. We know that such matrices exist when \( n = 2^k \) and when \( n = p + 1 \), where \( p \) is a prime and \( p \equiv -1 \) mod 4. If \( H \) is a Hadamard matrix, then \( H \) provides a good strategy for player 1: Any switch of rows or columns gives again a Hadamard matrix. Then Lindsay’s lemma tells us that a \( a \times b \) submatrix has sum of entries less than or equal to \( \sqrt{na\overline{b}} \) in absolute value. For \( a = b = n \), we get that the total sum of the entries in a Hadamard matrix is no more than \( n^{3/2} \). Hence, in this case player 2 can’t achieve a total gain of more than \( n^{3/2} \).

If \( n \) is not itself a Hadamard number, just take a Hadamard number \( \ell \) larger than it, no bigger than \( 2n \). (Doable, since between \( n \) and \( 2n \) there is always a power of 2.) Then
13.2. THE RAY-CHAUDHURI - WILSON THEOREM

take a Hadamard matrix of size $\ell \times \ell$, and choose your matrix to be any $n \times n$ submatrix of that. Then any row and column changes that player 2 performs, we can assume that they are performed to the whole matrix. By Lindsay’s lemma we then get that the sum is no more than $\sqrt{\ell n^2} < \sqrt{2}n^{3/2}$. (Note: If we use Hadamard matrices of size $p + 1$, we can actually get $1 + \varepsilon$ as the constant instead of $\sqrt{2}$.) Notice that there is a gap between the constants we have obtained in the deterministic upper bound and the probabilistic lower bound for the value. While this gap can be reduced, the instructor believes it has not been closed: the lower bound differs from the upper bound by $cn^{3/2}$.

13.2 The RAY-CHAUDHURI - WILSON theorem

See separate handout.
Chapter 14
14th day, Thursday 7/22/04 (Scribe: Ivona Bezáková)

14.1 Points in general position

Exercise 14.1. Find a curve $f : \mathbb{R} \rightarrow \mathbb{R}^n$ mapping any $n$ distinct points into $n$ points in “general position”. In other words, if $t_1, \ldots, t_n \in \mathbb{R}$ are distinct, then $f(t_1), \ldots, f(t_n)$ are linearly independent.

Recall the definition of the Vandermonde determinant and the corresponding closed form expression

$$V_n(t_1, \ldots, t_n) := \begin{vmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^{n-1} \\ 1 & t_2 & t_2^2 & \cdots & t_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_n & t_n^2 & \cdots & t_n^{n-1} \end{vmatrix} = \prod_{i \neq j}(t_i - t_j)$$

Then a solution to the above exercise is given by $f(t) = (f_0(t), \ldots, f_{n-1}(t)) = (1, t, t^2, \ldots, t^{n-1})$.

14.2 Pattern in proofs of inequalities using the Linear Algebra Method

The pattern can be summarized as follows. We have objects $a_1, \ldots, a_m \in \Omega$, where $\Omega$ is some abstract domain. We want to obtain an upper-bound on $m$. We define functions $f_1, \ldots, f_m : \Omega \rightarrow W$, where $W$ is a vector space, satisfying the “diagonal condition”

$$f_i(a_j) = \begin{cases} \neq 0 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

The set of functions $\Omega \rightarrow W$ forms a vector space in $W^\Omega$. The diagonal condition implies linear independence of $f_1, \ldots, f_m$. If in addition we find another set of functions $g_1, \ldots, g_k$
such that $f_1, \ldots, f_m \in \text{Span}(g_1, \ldots, g_k)$ then $m \leq k$ follows by the Fundamental Fact of Linear Algebra. A weaker condition suffices for this.

**Claim 14.2.** Suppose the $f_i$ satisfy “triangular condition”:

$$f_i(a_j) = \begin{cases} \neq 0 & \text{if } i = j \\ 0 & \text{if } i > j \end{cases}$$

Then the $f_i$ are linearly independent. (Note that we know nothing about $f_i(a_j)$ when $i < j$.)

**Exercise 14.3.** Prove Claim 14.2.

**Claim 14.4.** If $W = \mathbb{F}$ and the matrix $(f_i(a_j))_{i,j=1}^m$ is nonsingular then the $f_i$ are linearly independent.

**Exercise 14.5.** Prove Claim 14.4.

**Exercise 14.6.** Assuming $W = \mathbb{F}$, prove that the following statement is false: If the $f_i$ are linearly independent then the matrix $(f_i(a_j))_{i,j=1}^m$ is nonsingular.

The following theorem is a nonuniform version of the Ray-Chaudhuri–Wilson’s Theorem (the sets do not have to be equal size.)

**Theorem 14.7 (Frankl-Wilson).** Let $A_1, \ldots, A_m \subseteq [n]$ be such that $(\forall i \neq j) \left( |A_i \cap A_j| \in \{\ell_1, \ldots, \ell_s\} \right)$. Then

$$m \leq \binom{n}{s} + \binom{n}{s-1} + \cdots + \binom{n}{0}$$

The proof is a modification of the proof of the uniform case (all the $A_i$ are of the same size.) We cannot use the same functions $f_i(x)$ to prove the linear independence. The original $f_i$ were defined as

$$f_i(x) = \prod_{t=1}^s (v_i \cdot x - \ell_t),$$

where $x, v_i \in \mathbb{R}^n$, $v_i$ is the incidence vector of $A_i$. Notice that $f_i(v_i) \neq 0$ only if $|A_i| \notin \{\ell_1, \ldots, \ell_s\}$. Therefore we define $g_i(x)$ as follows

$$g_i(x) = \prod_{t: \ell_t \in A_i} (v_i \cdot x - \ell_t)$$

Then the $g_i$ satisfy

$$g_i(v_j) = \begin{cases} \neq 0 & \text{if } i = j \\ 0 & \text{if } |A_i \cap A_j| < |A_i|, \text{ i.e. } A_j \nsubseteq A_i \end{cases}$$

If we reorder the $A_i$ so that $A_j \nsubseteq A_i$ for $i > j$, then the condition on $g_i$ is a triangular condition. Therefore we could conclude that the $g_i$ are linearly independent. Such an ordering exists: it suffices to order the $A_i$ by their size — let $|A_1| \leq |A_2| \leq \cdots \leq |A_m|$. From this point on the proof follows along the lines of the uniform case.
14.3 Additional Exercises in Probability Theory

Exercise 14.8. A poker hand is a set of five cards. We say that the poker hand has a pair if there are two cards of the same kind (two Kings, or two 9s, for example). Compute the probability that a poker hand has at least one pair.

Definition 14.9. Two random variables are uncorrelated if $E(XY) = E(X)E(Y)$.

Exercise 14.10. Let $c$ and $d$ be constants. Prove: If $X$ and $Y$ are uncorrelated then $X + c$ and $Y + d$ are also uncorrelated.

Exercise 14.11. Suppose $X_1, \ldots, X_k$ are non-zero pairwise uncorrelated random variables with $E(X_i) = 0$. Prove: $X_1, \ldots, X_k$ are linearly independent (over $\mathbb{R}$).

Exercise 14.12. If $X_1, \ldots, X_m$ are random variables with $E(X_i) = 0$ then there exist pairwise uncorrelated random variables $Y_1, \ldots, Y_k$ with $E(Y_i) = 0$ such that $\text{Span}(X_1, \ldots, X_m) = \text{Span}(Y_1, \ldots, Y_k)$.

The process of finding $Y_1, \ldots, Y_k$ is called factor analysis in statistics. The $Y_i$ are called factors of the $X_i$.

Exercise 14.13. Prove: If there exist $k$ independent non-constant random variables over a probability space $\Omega$, then $|\Omega| \geq 2^k$.


(a) Prove: If $X_1, \ldots, X_k$ are non-constant pairwise independent random variables over $\Omega$, then $|\Omega| \geq k + 1$.

(b) Prove that this inequality is tight for $k = 2^\ell - 1$. 
Chapter 15

15th day, Friday 7/23/04 (Scribe: Eric Patterson)

15.1 Bases for Vector Spaces of Polynomials

Suppose there are $n$ cookies to distribute among $k$ children. How many ways are there to distribute the cookies? First, consider the number of ways to distribute the cookies so that each child gets at least one cookie. If we suppose the $i$-th child gets $x_i$ cookies, then we can solve the problem by looking for a solution to the equation $\sum_{i=1}^{k} x_i = n$ such that the unknowns are integers $x_i \geq 1$, $i \in \{1, \ldots, n\}$.

If the children are in order and we put the cookies in an order, all we have to do is insert partitions among $n$ objects. The first set of cookies ends at the first inserted partition, and this set goes to the first child. The second set of cookies begins at the first inserted partition and ends at the second inserted partition, and this set goes to the second child, and so on. For $n$ cookies, there are $n-1$ slots between cookies, and for $k$ children, there are $k-1$ partitions that we must insert, so we get $\binom{n-1}{k-1}$ ways to distribute the cookies.

Suppose now that we want to allow the possibility that some children might not get a cookie. Then we are trying to solve $\sum_{i=1}^{k} y_i = n$ with $y_i \geq 0$. By giving one “dummy” cookie to each child before we distribute the cookies, we can reduce this to the previous problem. Instead of $n$ cookies, however, we need $n+k$ cookies, so then we get $\binom{n+k-1}{n-1}$ ways to distribute the cookies.

Now, suppose we look at homogeneous polynomials in $r$ variables of degree $d$. They form a vector space with a basis of all monic monomials of degree $d$ in the $r$ variables $x_1, \ldots, x_r$. To compute the dimension of this space, we need to count those monomials.

Denote the number of such monomials by $H(r, k, \mathbb{F})$, where $\mathbb{F}$ is the field over which we are taking these polynomials. This number is the same as the number of solutions in the second cookie problem, but now we have degree $d$ instead of $n$ cookies, which we must divide between $r$ variables instead of $k$ children. Thus, we get $\binom{d+r-1}{r-1}$ monomials of degree $d$ in $r$ variables.

Now let $P(r,d,\mathbb{F})$ be the space of polynomials of degree $\leq d$ in $x_1, \ldots, x_r$. The condition here is that we are looking for $a_i \geq 0$ such that $\sum_{i=1}^{r} a_i \leq d$. We know how many ways we
can do this for each \( j \leq d \), so the total dimension is just the sum. To compute \( \sum_{j=0}^{d} j^r \begin{pmatrix} r \end{pmatrix} \), we can use Pascal’s triangle. In Pascal’s triangle, the sum of two adjacent entries is the entry between them on the next line. If we look at the entries \( j^r \begin{pmatrix} r \end{pmatrix} \) in Pascal’s triangle, we can use the preceding fact to calculate \( \sum_{j=0}^{d} j^r \begin{pmatrix} r \end{pmatrix} = \begin{pmatrix} r+d \end{pmatrix} \).

As an alternative method, we can use a “dummy” child \( x_{r+1} \) to get the remaining cookies not distributed. That is, if we have \( x_1^a \cdots x_r^a \) such that \( \sum_{i=1}^{r} a_i \leq d \), we can insert \( x_{r+1}^{d-\sum_{i=1}^{r} a_i} \) to get a monomial in \( r+1 \) variables of degree \( d \). Conversely, given a monomial of degree \( d \) in \( r+1 \) variables we can replace the \( r+1 \)st variable by 1 to get a monomial of degree \( d \) in \( r \) variables. This directly gives \( \begin{pmatrix} r+d \end{pmatrix} \) as \( P(r,d,F) \).

### 15.2 Projective Representation of a Graph

**Definition 15.1.** A **projective representation of a graph** \( G \) consists of the following information. Let \( W \) be a vector space over a field \( F \). To every vertex \( x \) in \( G \), we assign a subspace \( U(x) \leq W \) in such a manner, that \( U(x) \cap U(y) \neq \{0\} \) iff \( x \sim y \).

We want to minimize the dimension of the space \( W \) where a graph has such a representation.

**Definition 15.2.** The **projective dimension** of a graph \( G \) over a field \( F \) is

\[
\text{pdim}_F(G) = \min \{ \dim W : \text{ a projective representation of } G \text{ in } W \text{ exists} \}.
\]

For example, for the empty graph \( K_n \), \( \text{pdim}_F(K_n) = 0 \). For the complete graph \( K_n \), \( \text{pdim}_F(K_n) = 1 \) by having each vertex correspond to the whole space.

For a cycle, say a 5-cycle, we produce a projective representation in \( F^4 \). Assign nonzero vectors to each edge, and take the subspace corresponding to a vertex to be the subspace spanned by the vectors assigned to the incident edges. Since the subspaces associated to adjacent vertices will share a nonzero vector, the subspaces will have nonzero intersection. To prevent subspaces corresponding to nonadjacent vertices from having nonzero intersection, we need that any four of the vectors assigned to the edges be linearly independent, i.e., we need the five vectors \( u_1, \ldots, u_5 \) assigned to the edges to be in general position.

We know we can produce as many points in general position in our space as the number of elements in our field. So \( \text{pdim}_F(C_n) \leq 4 \) if \( |F| \geq n \).

Now look at a perfect matching graph with \( n \) vertices \( \frac{n}{2} K_2 \). If the field \( F \) has enough elements (at least \( \frac{n}{2} - 1 \)), its projective dimension is 2: send each pair to a different line in \( F^2 \).

**Theorem 15.3.** Over any field, every graph has a projective representation

**Proof:** Consider a vector space with basis a set of vectors in bijection with the set of edges on the graph. Then assign each vertex to the subspace spanned by the vectors corresponding
to the incident edges. Actually we can do better: if $\Delta =$ maximum degree, then $\dim W = 2\Delta$ suffices because all we need is, from a vector space of that dimension, to pick the vectors so that they are in general position. This we can do as long as $|F| \geq m$, where $m$ is the number of edges.

So, in fact we have shown that:

**Theorem 15.4.** \[ \text{pdim}_F G \leq 2\Delta. \]

In particular, for a bipartite graph, we get that its projective dimension is less than or equal to $n$.

**Exercise 15.5.** Improve this bound by a lot (at least to $o(n)$).

**Exercise 15.6.** Give a logarithmic upper bound on the complement of a perfect matching, i.e., show $\dim \left( \frac{n}{2}K_2 \right) \leq O(\log n)$.

Now, can we guarantee a lower bound on the projective dimension? (This question is related to sub-linear space complexity).

Given a machine with a big read-only table with $n$ bits, and small read-write tables, give example of something that can’t be computed in sub-linear space ($o(n)$).

The best lower bounds we have are $\log(n)$. We can prove though that almost all graphs have at least $\sqrt{n}$ lower bound.

**OPEN PROBLEM** (Concept introduced by Pudlák-Rödl): Construct an explicit family of graphs with projective dimension greater than $n^c$, where $c > 0$.

### 15.3 The Number of Zero-Patterns of a Sequence of Polynomials

See separate handout.
16.1 Random Variables

Recall the definition of a random variable: Given a probability space \((\Omega, P)\), where \(\Omega\) is the sample space and \(P\) the probability distribution over the sample space, a random variable is simply a function \(X : \Omega \rightarrow \mathbb{R}\). The expected value of \(X\) is \(E(X) = \sum_{x \in \Omega} X(x)P(x)\), i.e., it is a weighted average of the values of \(X\). Notice that

\[
E(X) = \sum_{x \in \Omega} X(x)P(x) = \sum_{y \in \mathbb{R}} yP(\{X = y\}).
\]

Given an event \(A \subseteq \Omega\), its probability is

\[
P(A) = \sum_{x \in A} P(x).
\]

The indicator variable of event \(A\) is the function \(\theta_A : \Omega \rightarrow \{0,1\}\) defined by

\[
\theta_A(x) = \begin{cases} 
1 & \text{if } x \in A \\
0 & \text{if } x \notin A 
\end{cases}.
\]

16.2 Independence

The expected value of \(\theta_A\) is \(E(\theta_A) = P(A)\). We say that two events \(A, B\) are independent, if \(P(A \cap B) = P(A)P(B)\). Three events \(A, B, C\) are independent, if \(P(A \cap B \cap C) = P(A)P(B)P(C)\) and they are pairwise independent. In general \(A_1, \ldots, A_k\) are independent, if for every \(I \subset [k]\) we have that

\[
P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i).
\]
The random variables $X_1, \ldots, X_k$ are said to be independent, if for any $a_1, \ldots, a_k$

$$P(X_1 = a_1 \text{ and } \ldots \text{ and } X_k = a_k) = \prod_{i=1}^{k} P(A_i).$$

**Exercise 16.1.** If $X_1, \ldots, X_k$ are independent random variables, then for any $I \subset [k]$ the sub-collection $(X_i : i \in I)$ is also independent.

**Exercise 16.2.** The events $A_1, \ldots, A_k$ are independent if and only if their corresponding indicator variables are independent.

Notice: $\theta_A = 1 - \theta_A$, $\theta_{A \cap B} = \theta_A \theta_B$, and these two also give us: $\theta_{A \cup B} = 1 - \theta_{A \cap B} = 1 - \theta_A \theta_B = \ldots = \theta_A + \theta_B - \theta_A \theta_B$.

**Exercise 16.3.** If $X_1, \ldots, X_k$ are independent, then $E\left(\prod_{i=1}^{k} X_i\right) = \prod_{i=1}^{k} E(X_i)$.

Let’s see what this means for the indicator variables: $E\left(\prod_{i=1}^{k} \theta_{A_i}\right) = \prod_{i=1}^{k} E(\theta_{A_i}) = E(\theta_{\cap A_i})$, in other words $P\left(\bigcap_{i=1}^{k} A_i\right) = \prod_{i=1}^{k} P(A_i)$. So this is true almost by definition for indicator variables.

**Exercise 16.4.** If $X_1, \ldots, X_k$ are random variables, then there exist polynomials in $k$ variables, $f_1, \ldots, f_m$ such that for each $i$, $Y_i := f_i(X_1, \ldots, X_k)$ is an indicator variable, the corresponding events are disjoint, and $(\forall j)(X_i \in \text{span}(Y_1, \ldots, Y_m))$.

**Exercise 16.5.** If $X_1, X_2, X_3, X_4$ are independent random variables, then $\sqrt{X_3^2 + \frac{1}{X_4^2+1}}$, $e^{X_1}$, $\cos(X_2)$ are independent.

In general, if we start with a number of random variables and split them in groups, and for each group we create a new random variable by using any function of the variables in that group, then those resulting random variables are independent.

### 16.3 Conditional Probability

For $B \neq \emptyset$ and any $A$, we define the conditional probability $P(A|B)$ as $P(A \cap B)/P(B)$. It is easy to see that if $A$, $B$ are independent and $B$ is nonempty, then $P(A|B) = P(A)$. The advantage of using the notion of independence instead of this last equality is that it doesn’t require us to exclude the case $B = \emptyset$ and it shows clearly that the notion of independence is symmetric.
16.3. CONDITIONAL PROBABILITY

Define the conditional expectation of a variable $X$ given the event $B$ as: 
\[ E(X|B) = \sum_{x \in B} X(x)P(\{x\}|B) = \sum yP(\{X = x\}|B) \]

Given $X,Y$ two random variables, what should $E(X|Y)$ mean? It would have to be a random variable: $Z := E(X|Y)$. It is defined by $(z \in \Omega) \mapsto Z(z) = E(X|Y = Y(z))$. Let’s see what this does when $X,Y$ are independent. Then
\[ Z(z) = \sum_{\{x|Y(x) = Y(z)\}} \frac{X(x)P(x)}{P(\{Y = Y(z)\})} \]
This is equal to \( \sum_t P(\{X = t\}|\{Y = Y(z)\}) \). If $X$ and $Y$ are independent, then it is further equal to \( \sum_t P(\{X = t\}) = E(X) \). So if $X$ and $Y$ are independent, then $Z$ is going to be just a number, the expected value of $X$.

Exercise 16.6. Show $E(X|X) = X$

We say that $X,Y$ are uncorrelated, if $E(XY) = E(X)E(Y)$. We know by Exercise 16.3 that if $X,Y$ are independent then they are uncorrelated.

Let us provide a counterexample for the converse: Let $\Omega = \{a,b,c\}$ and $P$ be the uniform probability distribution. Let $X$ take values \{1,0,-1\} at \{a,b,c\} respectively, and let $Y = X^2$, so it takes values \{1,0,1\} at \{a,b,c\} respectively. Then we have that $XY = X$ and $E(X) = 0$, so they are uncorrelated. However, $P(\{X = 0\}) = P(\{Y = 0\}) = \frac{1}{3}$, and $P(XY) \neq P(X)P(Y)$.

Exercise 16.7. If $|\Omega| = 2$, then uncorrelated random variables are also independent.

Exercise 16.8. If $X_1, \ldots, X_k$ are independent and not constant, then $|\Omega| = n \geq 2^k$.

Indeed, each variable can take at least two values. For each choice of a value $a_i$ for every variable $X_i$, we get a set \{X_1 = a_1, \ldots, X_k = a_k\} with positive probability, hence nonempty. There are at least $2^k$ such sets, and they are all disjoint.

In particular, to get many independent events, we need a large sample space.

Definition: $A$ is called a trivial event, if $A = \emptyset$ or $A = \Omega$. So for nontrivial events we always have: $0 < P(A) < 1$.

Corollary 16.9. If $A_1, \ldots, A_k$ are independent non-trivial events, then $n \geq 2^k$

Proof: use their indicator variables. \qed

If we only require pairwise independence, then how small can the size of the sample space be?
**Theorem 16.10.** If $X_1, \ldots, X_m$ are pairwise independent and non-constant, then $m \leq n - 1$.

**Proof:** Recall that if $X, Y$ are independent, then $X + c$ and $Y + d$ are independent, so without loss of generality we can assume that $(\forall i)(E(X_i) = 0)$. We claim that under this condition, and if the $X_i$ are pairwise uncorrelated, the $X_1, \ldots, X_m$ are linearly independent over $\mathbb{R}$ (as functions from $\Omega$ to $\mathbb{R}$). Recall that $\mathbb{R}^\Omega$ denotes the space of functions $\Omega \rightarrow \mathbb{R}$.

Since the dimension of this space is equal to $|\Omega|$, we get our result, since our functions lie in the kernel of the non-zero functional $E$ (the hyperplane consisting of the random variables with expected value 0). Another way to argue this last step is to add the function $X_0 = 1$. Then $X_0, \ldots, X_m$ are linearly independent. (They are still uncorrelated)

Now, the uncorrelated condition tells us that $E(X_i X_j) = 0$ iff $i \neq j$, since $E(X^2) > 0$ unless $X$ is zero. This shows by the standard argument that the $X_i$ are linearly independent.

As a consequence, if $A_1, \ldots, A_m$ are nontrivial events, then $m + 1 \leq n$. This is actually tight for infinitely many values of $n$: Suppose $n = 2^k$, and let $\Omega = \mathbb{F}_2^k$. Then any subspace of dimension $k - 1$ gives an event with probability $\frac{1}{2}$. So for each $u \in \mathbb{F}_2^k \setminus \{0\}$, $P(u) = \frac{1}{2}$. We need to know that these events are pairwise independent. If $u_1 \neq u_2$, then their span has dimension 2 since they are linearly independent (they can't be parallel), so $P(u_1 \cap u_2) = \frac{1}{4}$.

More generally, let $q$ be a prime power, $\Omega = \mathbb{F}^k_q$, and let $u_1, \ldots, u_m$ be elements in $\Omega$. Notice that $P(u_i) = \frac{1}{q}$ if $u_i \neq 0$.

**Exercise 16.11.** Show that $u_1^\perp, \ldots, u_m^\perp$ are independent events iff $u_1, \ldots, u_m$ are linearly independent vectors.

**Exercise 16.12.** Find out how this example over $\mathbb{F}_2^k$ relates to the Sylvester matrix.

**Exercise 16.13.** Find $n - 1$ pairwise independent events of probability $\frac{1}{2}$ over a sample space of size $n$ for every Hadamard number $n$.

**Exercise 16.14.** Show that if there exist $n - 1$ pairwise independent events of probability $\frac{1}{2}$ over a uniform probability space of size $n$, then $n$ is a Hadamard number.

**Exercise 16.15.** Show that for infinitely many $n$ there exist $\frac{n}{2}$ 3-wise independent non-trivial events over a sample space of size $n$.

**Exercise 16.16.** Show that if $X_1, \ldots, X_m$ are 4-wise independent nontrivial random variables, then $\binom{m}{2} \leq n$. 
Chapter 17

17th day, Monday 8/2/04 (Scribe: Charilaos Skiadas and Eric Patterson)

17.1 Puzzles

The Monty Hall Paradox.

On a game show, there are three closed doors. There is a car behind one door. Behind each of the other two doors, there is a goat. You select one of the doors, and the game master opens a different door, behind which is a goat. Then you are offered a choice to stay with your choice or to switch. The chance that we picked the car initially is $\frac{1}{3}$, so the chance that we get the car if we stay with our choice is $\frac{1}{3}$. Thus, the strategy that switches gives us a $\frac{2}{3}$ chance of picking the car.

The 2 Envelope Paradox

You get two envelopes. Each of them contains some money. One of the envelopes has twice as much money as the other. You are allowed to open one of the envelopes, see how much is in it, and then choose which envelope to pick. No matter what is in the envelope that you opened, you would expect that the other envelope would give you a larger expected amount of money. But you do not need to open the envelope to determine that. By this reasoning, we should keep switching back and forth between the two envelopes to increase the amount of money we expect to get.

Moral: be careful with the notion of expectation and probability space.

17.2 Statistical Independence vs Linear Independence

**Problem 17.1.** Can we create $m$ 3-wise independent non-trivial events, such that $n = |\Omega| = 2m$? ($m = 2^k$)

For pairwise independent events, we know that $n \geq m + 1$. Following the ideas in the pairwise case, define $\Omega := \mathbb{F}_q^d$ and suppose $v_1, \ldots, v_t$ are vectors in $\Omega$. We have shown that $v_1^\perp, \ldots, v_t^\perp$
are independent events iff the \( v_i \) are linearly independent. If \( v_i \neq 0 \), \( P(v_i^\perp) = \frac{1}{q} \) because the number of elements of a hyperplane is \( q^{\ell-1} \). Now \( U := v_1^\perp \cap \cdots \cap v_t^\perp = \text{span}(v_1, \ldots, v_t)^\perp \) had dimension \( \ell - \dim(U) = \ell - r \) where \( r \) is the rank of \( (v_1, \ldots, v_t) \). So the probability of \( U \) is \( \frac{1}{q^r} \). If the events are independent, then this has to be equal to \( \frac{1}{q^t} \). Hence \( r = t \), and the \( v_i \) are linearly independent. Conversely, if the \( v_i \) are linearly independent, then \( r = t \) and the events are independent.

To construct 3-wise independent events, we would need to construct vectors that are 3-wise linearly independent. In \( F_2^\ell \), we need to find \( 2^{\ell-1} \) 3-wise independent vectors, so \( n = 2^\ell \) and \( m = \frac{n}{2} \). Take an affine hyperplane not passing through 0 (a shift of a subspace). For example, we could put 1 in the first coordinate and either 0 or 1 in the other coordinates for a total of \( 2^{\ell-1} \) elements of \( \Omega \).

Claim: These vectors are 3-wise linearly independent.

Taking any three of the vectors, we would need to show that any nontrivial linear combination cannot be 0. Since the only elements of \( F_2 \) are 0 and 1, nontrivial combinations are sums of one, two, or all three of the vectors. This means (i) no one of them is 0, (ii) no two of them add up to 0, and (iii) no three of them add up to 0. The vectors we chose begin with coordinate 1, so they are not zero. If \( v + v' = 0 \), then \( v = -v' = v' \), but we assumed that we did not take two identical vectors. Any three of them add to a vector with first coordinate 1, so the sum is not equal to 0. If \( m \) is not a power of 2, we can still get that \( n < 4m \) by taking the smallest \( \ell \) such that \( m < 2^\ell \).

17.3 Algorithmic Application

A **Boolean function** is a function \( \{0, 1\}^n \to \{0, 1\} \). A **Boolean formula** is a formula composed of literals (variables and their negations) using AND and OR operations. For instance, \( \overline{x}_1 \lor (x_2 \land x_1 \land \overline{x}_2) \). A **disjunction** is an OR of literals. A **CNF** (conjunctive normal form) is an AND of clauses, each of which is a disjunction.

**Exercise 17.2.** Every Boolean formula can be represented as a CNF.

A 3-CNF formula is a formula in which every clause has 3 literals. To evaluate a Boolean formula on some substitution of values of the variables, recall that an OR of two variables is zero if and only if both variables are zero, and an AND of two variables is one if and only if both variables are one. An assignment of values to the variables in a Boolean formula **satisfies** the formula if the substitution of the values returns 1.

**Theorem 17.3.** Satisfiability of 3-CNF formulae is NP-complete.

Claim: For a 3-CNF formula, there always exists a substitution that satisfies at least \( 7/8 \) of the clauses.
Let $n$ be the number of variables and $m$ be the number of clauses.

Hint 1: Flip coins for the value of each variable; that is, make a random substitution. If each set of values is an element of $\Omega$, then $|\Omega| = 2^n$.

Hint 2: Find the expected number of satisfied clauses.

Let $X$ be the number of satisfied clauses, so $X$ is a random variable from the uniform probability space $\Omega$. We can write $X = \sum_{i=1}^{m} \vartheta_{i}$, where $\vartheta_{i}$ is the indicator of the event that the $i$th clause is satisfied. Then the expected value is the sum of the expected values of the $\vartheta_{i}$. Since the $\vartheta_{i}$ are indicator variables, $E(\vartheta_{i})$ equals the probability that the clause $C_{i}$ is satisfied. By the rules for evaluating Boolean formulae, the probability is $\frac{7}{8}$ that a disjunction of 3 literals with random variables is satisfied. Therefore, the expected number of satisfied clauses is $\sum_{i=1}^{m} \frac{7}{8} = \frac{7m}{8}$. Hence there exists an outcome $x$ such that $X(x) \geq \frac{7}{8}m$. So there exists a substitution that satisfies at least $\frac{7}{8}$ of the clauses.

If we want to find such a substitution deterministically in polynomial time, we should notice that we only used the fact that the variables occurring in a clause are 3-wise independent. Hence we can replace our space of outcomes with a space of size less than $4^n$ by the construction for finding 3-wise independent vectors above. This gives us a tiny fraction of all the substitutions, but the expected number of satisfied clauses is the same. Now we can simply try everything in this space, which will finish in quadratic time.

Moral: It is worth fighting for small sample space.

Recall the following exercise: if $X_1, \ldots, X_m$ are 4-wise independent nonconstant random variables, then $n = \left| \Omega \right| \geq \binom{m}{2}$.

**Proof:** Without loss of generality, we can assume that the expected value of each random variable is 0. The space of random variables has dimension $\dim \mathbb{R}^{\Omega} = n$. We want to construct $\binom{m}{2}$ random variables that will be linearly independent in this space. Look at the pairwise products of the $X_i$.

**Exercise 17.4.** Prove that the $\binom{m}{2}$ products $X_iX_j$ for $i \leq j$ are linearly independent.

### 17.4 Algebraic Coding Theory

Question: what is the maximum number of $k$-wise linearly independent vectors in $\mathbb{F}_q^\ell$?

A **codeword** is a sequence $(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}_q^n$. When you transmit the codeword, there is some noise, that is, some of the entries in the codeword might change value. If we have a given set of codewords from $\mathbb{F}_q^n$, we want to be able to distinguish them from one another even after some reasonable amount of noise interference. The **Hamming distance** between two codewords is the number of substitutions that must be made to change one into the other. If the Hamming distance is large among the codewords, the codewords will still be distinguishable after some interference. The **code space** is a subspace $U \leq \mathbb{F}_q^\ell$. We want $\dim U$ to be large (i.e., lots of codewords) and the Hamming distance between codewords to be at least something.
Exercise 17.5. Establish what this something is and relate it to the $k$ above, i.e., establish a connection between the two data above.
Chapter 18
18th day, Wednesday 8/4/04 (Scriber: Charilaos Skiadas)

18.1 Projective dimension of a graph, continued

Definition 18.1. A projective representation of a graph $G$ over a field $\mathbb{F}$ is a vector space $W$ over $\mathbb{F}$ and an assignment of subspaces to vertices $i \mapsto U_i \leq W$ such that $i \sim j$ (adjacent vertices) $\Rightarrow U_i \cap U_j \neq \{0\}$.

Projective dimension of a graph: $\text{pdim}_\mathbb{F}(G)$, $G$ a graph, $\mathbb{F}$ a field, defined as the smallest possible dimension of a projective representation of $G$. We showed that if $|\mathbb{F}| \geq m$, $m$ the number of edges, then the projective dimension is $\leq 2 \deg_{\text{max}}$, twice the maximum degree of the graph. We did it by associating to every edge of a graph a vector in a vector space of dimension $2\deg_{\text{max}}$ in such a way, that the vectors are in general position. Then we assign to each vertex the subspace spanned by the vectors corresponding to edges attached to that vertex. If two vertices are adjacent, then the spaces corresponding to them both contain the vector of the edge joining them. If we take the vectors to be in general position in a space of dimension $2\deg_{\text{max}}$, then any $2\deg_{\text{max}}$ of them will be linearly independent. So we get a representation. Recall that the vectors of the form $\{(1, \alpha, \alpha^2, \ldots, \alpha^{k-1}) : \alpha \in \mathbb{F}\}$ are in general position in $\mathbb{F}^k$.

Theorem 18.2. For any field $\mathbb{F}$, almost all graphs have projective dimension at least $c \sqrt{\frac{n}{\log n}}$.

"Almost all" means the lower bound is true for all but $o(2^\binom{n}{2})$ graphs on a given set of $n$ vertices, so the probability that a random graph on $n$ vertices does not have this property tends to 0 as $n \to \infty$.

It is an open question whether almost all graphs have large projective dimension over every field. Large here means greater than $n^c$, where $c > 0$ is a constant. Not even $(\log n)^{1+\epsilon}$ is known.

For the proof, we will need the theorem on zero-patterns. Recall that given $m$ polynomials $f = (f_1, \ldots, f_m)$ of degree $\leq d$ in $n$ variables over the field $\mathbb{F}$, then by substituting $\alpha \in \mathbb{F}^n$
in them, we get an element in \( F^m \). Replacing nonzero entries with a star gives us a “zero pattern,” i.e., a string in \([0, \ast]^m\). Let \( Z(f) \) be the number of zero-patterns for \( f \).

**Theorem 18.3 (RBG(2000)).** If \( d \geq 2 \), then \( Z(f) \leq \left( \frac{\text{end}}{n} \right)^n < \left( \frac{\text{end}}{n} \right)^n \).

The bound \( \left( \frac{\text{end}}{n} \right)^n \) holds for \( d = 1 \) also. We will use this to estimate the projective dimension of a random graph.

**Lemma 18.4.** If \( \text{pdim}_p(G) = k \) and \( |F| \) is large enough, then there exists a projective representation over \( F \) of dimension \( 2k \) such that for every \( i \), the dimension of every subspace representing a vertex has dimension \( k \).

**Proof:** Let \( U_i \) correspond to the vertex \( i \). Each of them has dimension at most \( k \). Add new vectors that are linearly independent (do it in a large space) such that each space has dimension \( k \). Then the dimension of this space is \( \leq k + nk \). Now each \( U_i \) has dimension \( k \), and the intersections are still the same.

**Exercise 18.5.** The new \( U_i \) represent \( G \).

We need to make a random projection to a \( 2k \)-dimensional space. Suppose \( W \) is a space over \( F \), \( F \) not too small, and let \( U_1, \ldots, U_n \leq W \), \( \dim U_i = k \) and let \( S \) be a space \( \dim S = 2k \). Let \( \phi : W \to S \) be a random map (i.e. a random \( 2k \times \dim W \) matrix). Then the dimension of \( f(U_i) \) remains the same, and the \( \dim \phi(U_i) \cap \phi(U_j) = \dim \phi(U_i \cap U_j) \), i.e. these spaces avoid the kernel of the map, whose dimension is \( \dim W - 2k \). Any random \( k \)-dimensional subspace, is likely to avoid it. It is also likely to avoid every \( U_i + U_j \), which guarantees that the dimensions of the pairwise intersections of the \( U_i \) and the \( \phi(U_i) \) are the same.

**Lemma 18.6.** If \( |F| \) is large enough, then given \( W, U_i \) as above, there exists \( K \leq W \) of dimension \( \dim W - 2k \) such that \( K \cap (U_i + U_j) = \{0\} \)

**Exercise 18.7.** Find out how large \( |F| \) has to be.

Project using this “generic” subspace as a kernel, onto a \( 2k \)-dimensional space. This gives the desired representation, proving Lemma [18.4].

Suppose now that we look for a projective representation as in the Lemma. Every subspace of it can be represented by a \( k \times 2k \) matrix, where the rows form a basis of this subspace. Then the intersection condition is that if we put the two matrices on top of one another, then the resulting matrix is singular if and only if the corresponding vertices are adjacent. Now if we are looking for such a projective representation, then all entries are variables \( t_{i,r,s} \). The determinants of the above matrices, (that we want to be 0 or not zero according to adjacency of the corresponding vertices), are polynomials in the \( t_{i,r,s} \), where \( 1 \leq i \leq n \), \( 1 \leq r \leq k \), \( 1 \leq s \leq 2k \). So we have \( 2k^2n \) variables. The conditions are that certain polynomials are zero in some cases and nonzero in some other cases. So this corresponds to a zero-pattern of these
18.2. THEOREM OF MILNOR-THOM

\binom{n}{2} polynomials. A zero entry corresponds to an edge, and \ast corresponds to a non-edge. So the number of distinct graphs that can be represented in such a space is exactly equal to the number of zero-patterns.

Suppose at least an \(\varepsilon\) fraction of the \(2\binom{n}{2}\) graphs can be so represented. Then

\[
\left(\frac{e\binom{n}{2}2k}{2k^2n}\right)^{2k^2n} \geq Z\left(\frac{1}{2} \geq \varepsilon 2^{\binom{n}{2}}\right),
\]

so

\[
\left(\frac{en^2k}{2k^2n}\right)^{4k^2n} > \varepsilon 2^{n(n-1)},
\]

\[
\left(\frac{2n}{k}\right)^{4k^2n} > \varepsilon 2^{n(n-1)},
\]

\[
(2n)^{4k^3} > \varepsilon 2/n^{2n-1}.
\]

Taking logs, we get that

\[4k^2 \log_2(2n) > (n-1) - \frac{2}{n} \log(1/\varepsilon)\]

so that

\[4k^2 > \frac{n-1}{1 + \log_2 n} - \frac{2}{n} \log(1/\varepsilon).
\]

Assuming \(\varepsilon > e^{-n}\), we get that asymptotically \(k \approx \sqrt{n/(\log n)}\).

This finishes the proof of the Theorem 18.2.

Over the real numbers, there is a much stronger result. There we have the concept of a sign-pattern. Again given polynomials as above, whenever we plug in something and get all nonzero numbers, then we get a sign pattern by looking at whether the number is positive or negative. Then we have the theorem:

**Theorem 18.8 (Warren(1968)).** *The number of sign patterns is* \(< \left(\frac{4md}{n}\right)^n\).

(As before, \(m\) is the number of the polynomials, \(d\) is the degree, and \(n\) is the number of variables.) An application of this: If you cut up \(\mathbb{R}^n\) with hypersurfaces, each region of the complement corresponds to a sign-pattern.

### 18.2 Theorem of Milnor-Thom

A real algebraic set is the set of common roots of a set of polynomials in real affine space.

**Theorem 18.9 (Milnor (1964), Thom(1965)).** *Let \(V \subseteq \mathbb{R}^n\) be the set of common roots of the polynomials \(f_1, \ldots, f_m \in \mathbb{R}[x_1, \ldots, x_n]\), where \(\deg f_i \leq d\). Then the number of connected components of \(V\) is \(\leq d(2d-1)^{n-1} \leq (2d)^n\).*
In fact, the same bound holds for the sum of the Betti numbers, which gives much more (i.e., it counts higher-dimensional holes). There are computer science applications of these higher Betti numbers too. Notice that the bound does not depend on the number of polynomials, but it depends exponentially on the number of variables.

**Exercise 18.10.** Show Warren’s theorem using the Milnor-Thom Theorem.

**Exercise 18.11.** From Warren’s theorem deduce the theorem on number of zero-patterns for the reals and complex numbers, or any field of characteristic 0 (with a different constant than e).
Chapter 19

19th day, Friday 8/6/04 (Scribe: Charilaos Skiadas)

19.1 Matrix rigidity (Valiant 1978)

Let $A$ be a matrix over a field $\mathbb{F}$. Our aim is to explore how we can reduce its rank with changing as few entries as possible. For that, denote by $R_r(A)$ the minimum number of entries that need to be changed in $A$ in order to obtain a matrix of rank less than or equal to $r$. In other words, this is equal to $\min\{\text{weight}(C) : \text{rk}(A - C) \leq r\}$ where the weight of a matrix is the number of nonzero entries. We are particularly interested in the case when $r = \Theta(n)$.

**Exercise 19.1.** For the identity matrix we have $R_r(I) = n - r$ (Prove that fewer changes do not suffice!)

**Proposition 19.2.** For every $n \times n$ matrix $A$ we have $R_r(A) \leq (n - r)^2$.

For instance, this would tell us that $R_{n/2} \leq n^2/4$. For the proof of the proposition, we can without loss of generality assume that $\text{rk}(A) \geq r$, so we can assume that $A$ has a nonsingular $r \times r$ minor $B$, and let us for simplicity assume it appears in the upper left corner. If we focus on the first $r$ rows of $A$ for the moment, we see that the columns of $B$ are linearly independent. Let $A'$ be the $r \times n$ matrix consisting of the first $r$ rows of $A$. Hence, all of the other columns of $A'$ can be written as linear combinations of the columns of $B$. So there are numbers $c_{i,j}$ such that $a_{i,j} = \sum_{k=1}^{r} a_{i,k} c_{k,j}$ for $i = 1, \ldots, r$ and $j = r+1, \ldots, n$. Now, if we simply change the entries in the $(n-r) \times (n-r)$ bottom right corner, and redefine them according to the above equation (we define $a'_{i,j} = \sum_{k=1}^{r} a_{i,k} c_{k,j}$ for $i, j = r+1, \ldots, n$), then this equation now hold for all $j$, in other words all the columns of $A$ are linear combinations of the first $r$ columns. These first $r$ columns are linearly independent, since the columns of $B$ are. This gives us a matrix with rank $r$, and we had to make $(n-r)^2$ changes, hence the proposition is proved.

**Theorem 19.3.** Over infinite fields, almost all matrices satisfy $R_r(A) = (n - r)^2$.

The meaning of “almost all” is that while $\dim M_n(\mathbb{F}) = n^2$, the “dimension” of the set (“variety”) of matrices with $R_r(A) < (n - r)^2$ is strictly less than $n^2$. 

85
Think of all \( n \times n \) matrices, \( M_n(\mathbb{F}) \), and let \( D_r \) be the subset consisting of all matrices of rank \( r \) or less. Then this is defined by a set of algebraic equations, namely all \((r+1) \times (r+1)\) minors are equal to 0. There are \( \binom{n}{r+1}^2 \) such equations. Such a set, an “affine variety”, has an intuitive notion of dimension. We want to find what that is for \( D_r \). Notice that all of the equations above are independent since, for any \((r+1) \times (r+1)\) minor, the matrix which is the identity on that minor and 0 everywhere else satisfies all but one of the equations.

A better way to measure dimension is how many independent directions there are for small changes: Suppose the rank of a matrix is \( r \), and that the top left \( r \times r \) minor is nonsingular. We would like to know how many parameters we can change slightly, while keeping the rank \( r \). We can freely slightly change all entries in the first \( r \) rows and in the first \( r \) columns and still keep the top left \( r \times r \) minor non-singular. But all the other \((n-r)^2\) entries are completely determined, if we want to keep the rank no more than \( r \). So we end up with \( r(n-r) = r(2n-r) = 2rn - r^2 = n^2 - (n-r)^2 \) free choices. So this is the dimension of \( D_r \). (\( D_r \) is a finite union of spaces of the above form, for the various choices of minors. But a finite union of sets of a given dimension still has the same dimension as them. This is only true over an infinite field!) Suppose now that we permit \( m \) entries to be changed before getting a matrix of rank \( r \). Then each of these entries increases the dimension by 1, giving us one more free variable, hence the whole process will increase the dimension overall by \( m \). Denote \( E_{r,m} \) to be the set of matrices \( B \) such that there exists a matrix \( A \in D_r \) with weight(\( B - A \)) \leq m. Then by the above discussion we have \( \dim E_{r,m} \leq \dim D_r + m \). So for \( m < n^2 - \dim D_r \), almost all matrices satisfy \( R_r(A) > mn \) since the set \( E_{r,m} = \{ A \mid R_r(A) \leq m \} \) has dimension \( \leq n^2 - \dim D_r + m < n^2 \). So for almost all matrices, \( R_r(A) \leq n^2 - \dim D_r = (n-r)^2 \), and since it was also less than or equal to it, it is actually equal to it. Let us emphasize again, that this only works over infinite fields.

**Exercise 19.4.** Prove that if \( \mathbb{F} \) is a fixed finite field, then almost all matrices satisfy

\[
R_2(A) > c \frac{n^2}{\log n}.
\]

“Almost all” here means that the proportion of matrices satisfying the inequality goes to 0 as \( n \to \infty \).

What we need is **explicit families** of matrices with high rigidity \((R_n(A) > n^{1+\varepsilon})\). This is not known for any fixed \( \varepsilon > 0 \). We have some examples of non-explicit families: For instance, a matrix with independent transcendental entries. In other words, a “generic matrix” is not explicit. Independent transcendental entries means that if \( f(a_{1,1}, \ldots, a_{n,n}) = 0 \) for some \( f \in \mathbb{Z}[a_{i,j}] \), then \( f = 0 \).

The big question here is: Can one construct matrices with high rigidity with integer entries, where the integer entries have at most polynomially increasing number of digits (i.e., the number of digits is at most \( n^c \) for some absolute constant \( c \))?

All nonzero Vandermonde matrices are candidates. A special case of interest is Vandermonde matrix \( V_n(1,\omega,\omega^2,\ldots,\omega^{n-1}) \), where \( \omega \) is a primitive \( n \)-th root of unity (this is known
19.1. MATRIX RIGIDITY (VALIANT 1978)

as the discrete Fourier Transform (DFT) matrix.) One way to generalize this is as follows:

Recall that a **character** of a finite abelian group is a homomorphism \( \chi : G \to \{ z \in \mathbb{C}, |z| = 1 \} \).

**Exercise 19.5.** Show that a finite abelian group of order \( n \) has exactly \( n \) characters.

The **character table** of an abelian group is an \( n \times n \) matrix with rows indexed by characters and columns indexed by elements of \( G \). The entries are evaluations of the characters on the elements.

**Exercise 19.6.** Show that the characters of a finite abelian group are orthogonal (i.e., the rows of the character table are orthogonal).

For example, if \( \mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z} \), all the characters are of the form \( \chi_i(a) = \omega^{ai} \), where \( \omega \) is as above. Then the above matrix is simply the character table of \( \mathbb{Z}_n \). Another example would be \( G = \mathbb{Z}_2 \times \cdots \mathbb{Z}_2 = (\mathbb{F}_2^k, +) \). Then \( |G| = 2^k \). If \( a, b \in G \), then \( \chi_0(a) = (-1)^{ab} \) is a character, so every element of \( G \) defines a character, and the character table in this case is the Sylvester matrix. More generally, all Hadamard matrices are expected to be good candidates for rigidity.

The best known result at this time is:

**Theorem 19.7 (S. Lokam).** \( R_{\pi/2}(P) > cn^2 \), for \( P \) the matrix made out of the square roots of the first \( n^2 \) prime numbers.

Unfortunately, this is not exactly explicit. The dimension of the extension of \( \mathbb{Q} \) by these numbers is \( 2^{n^2} \), so this field has exponentially large dimension.

**Exercise 19.8.** The square roots of all square-free numbers are linearly independent over \( \mathbb{Q} \).

This is the best we can say at the moment.

Open question: Give a nonlinear lower bound on \( R_{cn} \) for the generic Vandermonde matrix. Lokam’s proof gives \( R_{\sqrt{\pi/2}}(V_n(x_1, \ldots, x_n)) > cn^2 \).
Chapter 20

20th day, Monday 8/9/04 (Scribe: Charilaos Skiadas)

20.1 Matrix Rigidity continued

Matrix rigidity Let $A$ be a matrix over a field $\mathbb{F}$. Recall from last time that $R_r(A) = \min \{ \text{weight}(C) : \text{rk}(A - C) \leq r \}$, the rigidity function. We “proved” that if $\mathbb{F}$ is infinite, then for almost all $n \times n$ matrices $A$ we have that $R_r(A) = (n - r)^2$. So, if $r < (1 - \epsilon)n$ then $R_r(A) = \Omega(n^2)$ for almost all matrices. We don’t know any “explicit” examples of families of such matrices. But we know

**Theorem 20.1 (S. Lokam).** $R_{\sqrt{n}}(V_n(x_1, \ldots, x_n)) > cn^2$, where the $x_i$ are independent transcendentals.

For $r > 2\sqrt{n}$, even a $n^{1+\epsilon}$ lower bound is not known. We will see the idea for the proof by looking instead at the matrix in the following theorem:

**Theorem 20.2.** Let $P$ be the matrix with entries the square roots of the first $n^2$ primes. then $R_{\sqrt{n}^2} > cn^2$

We will use that the square roots of all square-free integers are linearly independent over $\mathbb{Q}$. The main tool in the proof is the Shoup-Smolensky invariant of a set of numbers. This is defined as follows: $S_t(a_1, \ldots, a_n) = \text{rk}_\mathbb{Q} \{a_{i_1} \cdots a_{i_t} : 1 \leq i_1 < \cdots < i_t \leq m\}$. For example, $S_2(\sqrt{2}, \sqrt{3}, \sqrt{5}, \sqrt{7})$ is the rank of $\{\sqrt{6}, \sqrt{10}, \sqrt{14}, \sqrt{15}, \sqrt{21}, \sqrt{35}\} = 6$, while $S_3(\sqrt{2}, \sqrt{3}, \sqrt{5}, \sqrt{7}) = 4$. If $A$ is an $n \times m$ matrix with entries $a_{i,j}$, then $S_t(A) = S_t(a_{1,1}, \ldots, a_{n,m})$. Also let $S_t(a_1, \ldots, a_n)$ is same as above, only with repetitions permitted. Obviously $S_t \leq S_{t'}$.  

**Lemma 20.3.** $S_t(AB) \leq S_t(A)S_t(B)$.

**Proof:** In $AB$, if we take the product of $t$ entries, then each such entry is a sum of products of pairs. So a $t$-wise product is going to be a sum of $t$-wise products from $A$ times $t$-wise products from $B$. So the space on the left hand side will be generated by those products. This proves the above bound.
Corollary 20.4. If \( A \in M_n(\mathbb{F}) \) with \( \text{rk}(A) = r \), then \( S_t^s(A) \leq \left( \frac{nr+t-1}{t-1} \right)^2 \).

Proof: The rank of \( A \) is less than or equal to \( r \) if and only if \( A \) can be represented as \( A = BC \), where \( B \) has \( r \) columns and \( C \) has \( r \) rows. Hence \( S_t^s(A) \leq S_t^s(B)S_t^s(C) \). Now each of those is no greater than the number of all \( t \)-wise products with repetition, which gives us the above binomial estimate. (Recall the cookies problem, where not everyone necessarily gets a cookie). In general, if \( B \) is an \( n \times r \) matrix, then \( S_t^s(B) \leq \left( \frac{nr+t-1}{t-1} \right) \).

Lemma 20.5. With \( P \) as in the theorem above (the one containing square roots of primes), suppose weight \( (C) = w \). Then \( S_t(P - C) \geq \left( \frac{n^2 - w}{t} \right) \).

The proof is immediate (using the fact that square roots of the square-free positive integers are linearly independent). So if \( w = R_t(P) \), we have that \( \left( \frac{n^2 - w}{t} \right) \leq S_t(P - C) \leq S_t^s(P - C) \leq \left( \frac{nr+t}{t} \right)^2 \). Setting \( t = nr \), we get that \( \left( \frac{n^2 - w}{nr} \right)^{nr} < \left( \frac{n^2 - w}{nr} \right) \leq \left( \frac{2nr}{nr} \right)^2 < 16^{nr} \). This gives us that \( \frac{n^2 - w}{nr} < 16 \). We find that \( w \geq n^2(1 - \frac{16n}{n}) \). So we have shown that \( R_t(P) \geq n^2(1 - \frac{16n}{n}) \), which proves the theorem, if we set \( r = \frac{n}{16} \).

Exercise 20.6. Prove with a similar argument Lokam’s theorem, that \( R_{\sqrt{n}/2}(V_n(x_1, \ldots, x_n)) > cn^2 \).

Exercise 20.7. Prove the same when \( x_i = p_i^{1/n} \).

Note, that \( \dim_{\mathbb{Q}} \mathbb{Q}(p_1^{1/n}, \ldots, p_n^{1/n}) = n^n \), so this is not very “explicit” either.

A linear code of length \( n \) and dimension \( k \) over \( \mathbb{F}_q \) is an \( k \)-dimensional subspace \( C \) in \( \mathbb{F}_q^n \). The information rate is \( \frac{k}{n} \). The idea is that our original message is in \( \mathbb{F}_q^k \), and is encoded using the encoding \( \mathbb{F}_q^k \cong C \). The minimum weight of \( C \) is the minimum of the weights of all \( x \in C \), \( x \neq 0 \). The Hamming distance of \( x, y \in \mathbb{F}_q^n \) is the weight of \( x - y \). For any \( x \neq y \) with \( x, y \in C \), \( \text{dist}(x, y) \geq \text{minweight}(C) \). Let \( d = \text{minweight}(C) \), the coding distance. In noisy transmission, up to \( d - 1 \) errors can be recognized, and up to \( \frac{d-1}{2} \) errors can be uniquely corrected. The goal of algebraic coding theory is to find codes with large information rate and large coding distance, and of course that are explicit.

Suppose \( C \leq \mathbb{F}_q^n \), \( \dim C = k \). Then \( \dim C^\perp = n - k \). A basis of \( C^\perp \) will be given by a \( n - k \times n \) matrix \( B \).

Exercise 20.8. \( \text{minweight}(C) \geq d \) iff the rows of \( B \) are \( d - 1 \)-wise linearly independent.

OPEN QUESTION: Do there exist good cyclic codes (cyclic codes are those where a cyclic permutation of the entries preserves the space.) Good means \( \frac{k}{n} = \Omega(1) \), \( d = \Theta(n) \).

In \( \mathbb{F}_q^n \), we have \( q \) vectors that were \( n \)-wise independent, namely the elements of the form \( (1, a, a^2, \ldots, a^{d-1}) \). This will give us a code in \( \mathbb{F}_q^n \). Then, \( \dim C^\perp = d \), rate is \( 1 - \frac{d}{n} \). If we do this over a large field, we get excellent rate and error-correction. These are good codes over large field. What we want though is a family of codes which is good over a fixed finite field. Most important are the binary codes, over \( \mathbb{F}_2 \).
Chapter 21

21th day, Wednesday 8/11/04 (Scribe: Ivona Bezáková)

Exercise 21.1. Let $u_1, \ldots, u_5, \frac{OA_i}{O}$ be five pairwise perpendicular unit vectors in $\mathbb{R}^3$ where $O$ is the origin and the vertices $A_i$ form a regular pentagon. Let us assume that all the $A_i$ have the same $x$-coordinate. Let $c = (1, 0, 0)$. Prove: $\cos(c, u_i) = \frac{1}{\sqrt{5}}$.

Hint. Spherical cosine formula: Let $(a, b, c)$ be a spherical triangle with arc lengths $a, b, c$ and let $A$ be the angle of $b$ and $c$. Then $\cos a = \cos b \cos c + \sin b \sin c \cos A$.

21.1 Linear Programming

Notation 21.2. Let $a, b \in \mathbb{R}^n$. We denote $a \leq b$ if $a_i \leq b_i$ for every $i \in [n]$.

Given is a real $k \times n$ matrix $A$, and vectors $b \in \mathbb{R}^k$, $c \in \mathbb{R}^n$. Let $x \in \mathbb{R}^n$ be an unknown vector. A linear program specifies the constraints on $x$ (a system of $k + n$ linear inequalities in $n$ unknowns): $Ax \leq b$ and $x \geq 0$. The goal, or the objective function is to maximize $c^T x$.

In a summarized form, here is the linear program and its dual:

<table>
<thead>
<tr>
<th>Linear Program (LP)</th>
<th>Dual Linear Program (DLP)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Constraints:</strong></td>
<td>$A^T y \geq c$</td>
</tr>
<tr>
<td>$Ax \leq b$</td>
<td>$y \geq 0$</td>
</tr>
<tr>
<td>$x \geq 0$</td>
<td>$\min b^T y$</td>
</tr>
<tr>
<td>$\max c^T x$</td>
<td></td>
</tr>
<tr>
<td><strong>Where</strong> $A \in \mathbb{R}^{k \times n}, b \in \mathbb{R}^k, c, x \in \mathbb{R}^n$</td>
<td>$A^T \in \mathbb{R}^{n \times k}, b, y \in \mathbb{R}^k, c \in \mathbb{R}^n$</td>
</tr>
</tbody>
</table>

Definition 21.3. The LP is feasible if there exists a solution. For a feasible LP we denote the optimum solution by $\text{opt}_{LP} := \max c^T x$.

Theorem 21.4 (LP duality). If $LP$ and dual $LP$ are feasible, then $\text{opt}_{LP} = \text{opt}_{dual LP}$.

Exercise 21.5. Prove the $\leq$ direction of the LP duality theorem.

Definition 21.6. Integer linear programming (ILP) looks for integral optima under linear constraints. In other words, all coordinates of $x$ must be integers.
Clearly,
\[
\text{opt}_{ILP} \leq \text{opt}_{LP} = \text{opt}_{dual\ LP} \leq \text{opt}_{dual\ LP}
\]

As an example we present the following ILP for \(\alpha(G)\), the size of the maximum independent set in a graph \(G\). There will be an unknown variable \(x_i\) for every vertex \(i\). The LP:
\[
\sum_{i \in C} x_i \leq 1 \quad \text{for every clique } C
\]
\[
x_i \geq 0
\]
\[
\max \sum_i x_i
\]

Integer optimum is exactly \(\alpha(G)\) (why?). We denote the fractional optimum by \(\alpha^*(G)\).

Dual variables are: \(y_C\) for every clique \(C\). The dual LP:
\[
\sum_{C \ni i} y_C \geq 1 \quad \text{for every vertex } i
\]
\[
y_C \geq 0
\]
\[
\min \sum_C y_C
\]

Integral optimum of the dual LP is the minimum clique cover (why?), i.e. the minimum number of cliques covering \(V(G)\). Notice that this is exactly the chromatic number of the complement of \(G\), \(\chi(G)\). Fractional optimum is denoted \(\chi^*(\overline{G})\).

Therefore,
\[
\alpha(G) \leq \alpha^*(G) = \chi^*(\overline{G}) \leq \chi(\overline{G})
\]

Notice that the inequality \(\alpha(G) \leq \chi(G)\) is trivial. However, using the linear programming method we get an interesting quantity in between.

**Exercise 21.7.** Find the values of \(\alpha(C_5), \alpha^*(C_5), \chi(\overline{C_5}), \chi^*(\overline{C_5})\).

**Exercise 21.8.** Work out \(\chi^*(\overline{C_7})\) from scratch (not using the above inequality).

## 21.2 Shannon Capacity of a graph

**Notation 21.9.** We denote the adjacency of two vertices \(i, j\) by \(i \sim j\). If \(i\) is adjacent or equal to \(j\), we write \(i \cong j\).

**Definition 21.10.** Let \(G, H\) be two graphs. The **strong product** of \(G\) and \(H\), denoted \(G \cdot H\) is defined as follows. The vertex set \(V(G \cdot H) := V(G) \times V(H)\), while the adjacency \((i, j) \sim (i', j')\) holds if \((i, j) \neq (i', j')\) and \(i \cong i'\) (in \(G\)) and \(j \cong j'\) (in \(H\)).

**Definition 21.11.** Let \(G\) be a graph and let \(G^n\) be its \(n\)-th strong power. **Shannon capacity** of \(G\), denoted \(\Theta(G)\), is defined as
\[
\Theta(G) = \lim_{n \to \infty} \sqrt[n]{\alpha(G^n)}
\]
Exercise 21.12. Prove: \( \alpha(G \cdot H) \geq \alpha(G) \alpha(H) \)

**Lemma 21.13.** Let \( f : \mathbb{N} \to \mathbb{R}^+ \) be a function satisfying \( f(nm) \geq f(n)f(m) \). Then \( \lim_{n \to \infty} \sqrt[n]{f(n)} \) exists and is equal to \( \sup_{n \to \infty} \sqrt[n]{f(n)} \).


**Corollary 21.15.** For every \( k \), \( \Theta(G) \geq \sqrt[2k]{\alpha(G^k)} \)

**Corollary 21.16.** \( \alpha(G) \leq \Theta(G) \)

Exercise 21.17. Prove: \( \Theta(G) \leq \chi(G) \)

The Shannon's capacity of \( C_5 \) is bounded by \( 2 \leq \Theta(C_5) \leq 3 \). We will find its exact value shortly.

**Open problem.** The exact value of \( \Theta(C_7) \).

**Lemma 21.18.** If \( g(G) \) is a graph function s.t.

1. for every \( G \), \( \alpha(G) \leq g(G) \)
2. \( g(G \cdot H) \leq g(G)g(H) \)

then \( \Theta(G) \leq g(G) \).

Exercise 21.19. Prove: \( \alpha^*(G \cdot H) = \alpha^*(G)\alpha^*(H) \)

Hint. \( \leq \) directly, \( \geq \) dually.

**Corollary 21.20.** \( \Theta(G) \leq \alpha^*(G) \)

**Corollary 21.21.** \( \Theta(C_5) \leq 5/2 \).

We will use a bootstrapping method to tighten the lower bound. By Corollary 21.13 we know that \( \Theta(C_5) \geq \sqrt[5]{\alpha(C_5^5)} \).

Exercise 21.22. Prove: \( \alpha(C_5^2) = 5 \)

**Theorem 21.23 (Lovász).** Let \( G \) be a vertex transitive and self-complementary graph, i.e. the automorphism group of \( G \) is transitive and \( G \cong \overline{G} \). Then \( \Theta(G) = \sqrt{n} \).

In particular, the theorem implies that \( \Theta(C_5) = \sqrt{5} \).

Exercise 21.24. Prove: If \( G \) is self-complementary, i.e. \( G \cong \overline{G} \), then \( \Theta(G) \geq \sqrt{n} \).
21.2.1 Lovász’s theta function (Lovász’s capacity)

**Definition 21.25.** *Orthonormal representation* of a graph $G$ is defined as follows. Every vertex $i$ is assigned a $d$-dimensional unit vector $u_i \in \mathbb{R}^d$, $\|u_i\| = 1$. If $i \neq j$, then $u_i$ and $u_j$ are perpendicular.

**Definition 21.26.** *Lovász capacity*, denoted $\theta(G)$ is defined as the minimum over all orthonormal representations $u_i$ of $G$ of the quantity $\min \|c\| = 1 \max_{i=1,...,n} 1/\langle c, u_i \rangle$.

**Theorem 21.27.** For every graph $G$,

1. $\chi(G) \leq \theta(G)$ and
2. $\theta(G \cdot H) = \theta(G)\theta(H)$.

Moreover, $\theta(G)$ can be approximated in polynomial time.

**Corollary 21.28.** $\Theta(G) \leq \theta(G)$

**Exercise 21.29.** Prove: $\alpha(G) \leq \theta(G) \leq \chi(G)$.

By the first exercise, $\theta(C_5) \leq \sqrt{5}$. Recall that $\sqrt{5} \leq \Theta(C_5) \leq \theta(C_5) \leq \sqrt{5}$ and therefore $\Theta(C_5) = \sqrt{5}$.

**Lemma 21.30.** If $u_1, \ldots, u_k$ are orthonormal, then $\|c\|^2 \geq \sum \langle c, u_i \rangle$.

**Exercise 21.31.** Prove the lemma.

To prove part 1 of Theorem 21.27 take a suitable subset of vertices, use the above lemma (a variant of the Pythagorean theorem) and the definition of Lovász’s capacity. To complete the proof of $\Theta(C_5) = \sqrt{5}$ it suffices to show the $\leq$ inequality in part 2 of Theorem 21.27. This will be done using tensor products.

**Definition 21.32.** *Tensor product* of two vectors $a \in \mathbb{R}^k$ and $b \in \mathbb{R}^\ell$, denoted $a \circ b$, is a $k\ell$-dimensional vector, where $(ik+j)$-th coordinate is $a_ib_j$.

**Exercise 21.33.** Let $a,b \in \mathbb{R}^k$, and $u,v \in \mathbb{R}^\ell$. Prove: $\langle a \circ b, a \circ v \rangle = \langle a, b \rangle \langle u, v \rangle$.

**Corollary 21.34.** If $a_1, \ldots, a_k$ is an orthonormal representation of $G$ and $u_1, \ldots, u_\ell$ is an orthonormal representation of $H$ then $a_i \circ u_j$ is an orthonormal representation of $G \cdot H$.

**Corollary 21.35.** $\theta(G \cdot H) \leq \theta(G)\theta(H)$
Chapter 22

22th day, Friday 8/13/04 (Scribe: Charilaos Skiadas)

22.1 0, 1-measures

Exercise 22.1. Prove the following theorem. (Hint: Use Zorn’s Lemma.)

Theorem 22.2. If Ω is an infinite set, then there exists a finite additive nontrivial 0, 1 measure on $2^Ω$, $μ : 2^Ω → \{0, 1\}$, such that $μ(Ω) = 1$. Nontrivial means that $μ(\{a\}) = 0$ for all $a$.

This is a hint to the infinite switch problem.

22.2 Sign-Rigidity (Paturi-Simon)

Let $M$ be a sign matrix, i.e., a matrix consisting of ±1 entries. A matrix $A$ of the same dimensions as $M$ realizes $M$ if the sign of $a_{i,j}$ is equal to $m_{i,j}$. The sign-rank of $M$ is the minimum rank of a matrix realizing $M$.

Exercise 22.3 (Alon-Frankl-Rõdl, 1984). Use Warren’s theorem to prove that almost all $n \times n$ sign-matrices have sign-rank greater than or equal to $\frac{n}{32}$.

Finding explicit matrices that satify the preceding example is a hard problem. No particular examples are known.

Conjecture 22.4. Hadamard matrices have sign-rank $\geq cn$.

For applications, all we need would be an explicit matrix with rank greater than $n^ε$ for some fixed $ε > 0$. The best known explicit bound was $Ω(\log n)$ until 2002.

Theorem 22.5 (Forster 2002). Let $X ⊆ \mathbb{R}^k$ such that $|X| ≥ k$ and the elements of $X$ are in general position (i.e., $k$-wise linearly independent). Then there exists an invertible $k \times k$ matrix $A$ such that $\sum_{x ∈ X} \frac{(Ax)(Ax)^T}{(Ax)^T(Ax)} = aI_k$, a constant multiple of the identity matrix.
If we assume that the theorem is true, the value of $a$ can easily be found by taking the trace on both sides (recall that the trace of the product of two matrices is independent of the order). This gives $a_k = \sum_{x \in X} 1 = |X|$

**Exercise 22.6 (Forster).** If $A$ realizes $M$, then $\text{rk}(A) \geq \frac{n}{\|M\|}$. Recall that $\|M\| = \max_{\|x\| = 1} \|Mx\| = \sqrt{\lambda_{\max}(M^TM)}$ by the spectral theorem.

**Corollary 22.7.** If $H$ is a Hadamard matrix, then $\text{sign-rank}(H) \geq \sqrt{n}$.

**Proof:** If $A$ realizes $H$, then $\text{rk}(A) \geq \frac{n}{\|H\|}$. Since $H^TH = nI$, we get that $\|H\|^2 = \lambda_{\max}(nI) = n$